



**HAL**  
open science

# Une étude discursive de la circulation des nouveautés langagières sur les réseaux sociaux numériques chinois et français : le cas de Sina Weibo et Twitter

Shan Gao

## ► To cite this version:

Shan Gao. Une étude discursive de la circulation des nouveautés langagières sur les réseaux sociaux numériques chinois et français : le cas de Sina Weibo et Twitter. Linguistique. Université Bourgogne Franche-Comté, 2021. Français. NNT : 2021UBFCC013 . tel-03556383

**HAL Id: tel-03556383**

**<https://theses.hal.science/tel-03556383v1>**

Submitted on 4 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**THÈSE DE DOCTORAT DE L' UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ  
PRÉPARÉE À L'UNIVERSITÉ DE FRANCHE-COMTÉ**

École doctorale n°592

Lettres, Communication, Langues, Arts (LECLA)

**Doctorat de sciences du langage**

Par

GAO Shan

**Une étude discursive de la circulation des nouveautés langagières  
sur les réseaux sociaux numériques chinois et français : le cas de *Sina Weibo* et *Twitter***

Volume I

Thèse présentée et soutenue à Besançon, le 17 mai 2021

Composition du Jury :

M. PRUVOST Jean,  
Mme MITROPOULOU Eleni,  
Mme CHEN Min,  
Mme CHAUVIN-VILENO Andrée,  
M. MADINI Mongi,

Professeure émérite à Cergy Paris Université,  
Professeure à l'Université de Haute-Alsace,  
MCF à l'Université de l'An'hui (Chine),  
Professeure émérite à l'Université de Franche-Comté,  
MCF HDR émérite à l'Université de Franche-Comté,

Président de Jury & Rapporteur  
Rapporteuse  
Examinatrice  
Directrice de thèse  
Codirecteur de thèse



*La langue française n'est point fixée et ne se fixera point. Une langue ne se fixe pas [...].*

*--Victor Hugo*



## REMERCIEMENTS

A l'issue d'une longue aventure scientifique, je voudrais exprimer mes profonds remerciements en premier lieu, à ma directrice de thèse Madame Andrée CHAUVIN-VILENO pour son précieux encadrement ainsi que ses conseils, sa patience. Je tiens à remercier très chaleureusement mon co-directeur de thèse Monsieur Mongi MADINI qui a alimenté ma réflexion scientifique et m'a guidé dans mes recherches depuis mon travail de Master. J'aimerais leur dire à quel point j'ai apprécié leur grande disponibilité, leur bienveillance et leur soutien pendant mes moments de difficulté.

Je tiens à également à remercier les membres de jury, Madame la professeur Eleni MITROPOULOU et Monsieur le professeur Jean PRUVOST qui m'ont fait l'honneur d'accepter de lire et d'évaluer mon travail. Ma reconnaissance va également à Madame Min CHEN qui a bien voulu participer au jury et partager ses connaissances indispensables.

J'aimerais exprimer toute ma gratitude à l'équipe scientifique du Laboratoire ELLIADD, et notamment à Madame Séverine EQUOY-HUTIN et Monsieur Alain LAMBOUX-DURAND qui ont partagé leurs précieuses remarques lors des comités de suivi de thèse. Je remercie également l'École doctorale LECLA pour son soutien.

Je tiens à remercier mes parents pour leurs soutiens indéfectibles. Je tiens à remercier mes chères amies chinoises Yuting YAO, Hanning GAO, Qian NIE, Zhao RUAN qui ont plus particulièrement assuré le soutien affectif. Mes remerciements vont également à Xiaoyu HUANG pour son enthousiasme à l'égard de mes travaux comme de la vie en général. Mes remerciements particuliers s'adressent à Nadine NEDONCELLE qui m'a aidée à intégrer dans une belle vie française (vive les épinards !!!).

Enfin, je voudrais remercier tous mes amis et mes collègues qui m'ont encouragée jusqu'à la fin de cette thèse.



# SOMMAIRE

REMERCIEMENTS .....	3
SOMMAIRE .....	5
INTRODUCTION GÉNÉRALE .....	7
PARTIE I : CONTEXTUALISATION.....	15
CHAPITRE I : Présentation du chinois : caractéristiques historiques, linguistiques et sociolinguistiques.....	17
1.1. Généralités sur la langue chinoise.....	18
1.2. Aperçu historique .....	21
1.3. Les dialectes chinois.....	27
1.4. Les caractéristiques du chinois contemporain.....	34
1.5. La saisie informatique du chinois.....	45
CHAPITRE II : Contexte numérique et nouveaux dispositifs médiatiques.....	49
2.1. L'Internet et le Web .....	50
2.2. Réseaux sociaux, RSN, médias sociaux, quelles sont les différences ?.....	55
2.3. Le microblog, en tant que RSN.....	72
2.4. Rôles des usagers et des chercheurs dans le contexte numérique .....	79
CHAPITRE III : Des néologismes aux nouveautés langagières en contexte numérique : définition de l'objet de recherche.....	85
3.1. Les métadiscours savants en question.....	86
3.2. Créations lexicales en contexte numérique .....	109
3.3. Nommer notre objet d'étude .....	127
PARTIE II : MÉTHODOLOGIE ET ANALYSES.....	131
CHAPITRE IV : Présentation des corpus et de la méthodologie de constitution des corpus	133
4.1. Questionnements et principales options relatives au corpus.....	133
4.2. Démarches préparatoires en vue de la collecte des corpus .....	136
4.3. Une démarche parallèle mais différenciée pour un double corpus .....	154
CHAPITRE V : Constitution et analyse des corpus de liste .....	159
5.1. Constitution du corpus de liste chinois .....	159
5.2. Procédures de constitution du corpus français .....	175
5.3. Analyse linguistique comparée des corpus de listes.....	180
CHAPITRE VI : Constitution d'un corpus de tweets en vue d'une analyse discursive : approche théorique et méthodologique.....	217
6.1. Éléments théoriques et méthodologiques pour le traitement des corpus RSN.....	217
6.2. La constitution du corpus de tweets chinois en 4 étapes.....	237
6.3. La constitution du corpus de tweets français .....	261
CHAPITRE VII : Analyse discursive de deux corpus de tweets .....	273
7.1. Analyse discursive d'un corpus de tweets chinois .....	273
7.2. Analyse discursive d'un corpus de tweets français .....	295
CONCLUSION GÉNÉRALE.....	321
BIBLIOGRAPHIE .....	327
LISTE DES TABLEAUX .....	341
LISTE DES FIGURES.....	343
TABLE DES MATIÈRES .....	347





# INTRODUCTION GÉNÉRALE

Dans cette recherche, nous nous intéressons à de nouveaux mots et expressions dont l'apparition et/ou la diffusion sont favorisés par le contexte numérique et participatif du Web. Cet intérêt est éveillé principalement par deux raisons personnelles liées à notre formation et à notre curiosité à l'égard du langage et de la communication.

En premier lieu, cette étude s'inscrit dans la continuité de notre recherche de Master en Sciences du langage dans laquelle nous avons étudié les initiatives créatives des internautes, notamment à travers leurs productions langagières. Nous avons travaillé en effet sur les stratégies de traduction employées par les internautes chinois pour transposer la langue et la culture françaises dans le cadre du sous-titrage amateur (le *fansub*) pour des productions audiovisuelles françaises ensuite diffusées en ligne. Nous avons spécialement choisi d'observer les situations langagières ou culturelles où les équivalents chinois n'existent pas. Nous avons alors remarqué que des nouveaux mots et expressions populaires en ligne (souvent créés par des individus) sont utilisés dans le sous-titrage pour créer une connivence avec les spectateurs qui sont eux-mêmes éventuellement des internautes. Face à ces créations verbales constatées, nous nous sommes demandée comment les traduire en français et quelles sont les créations comparables en France. Ces questions nous ont conduit au projet d'observer les nouvelles productions langagières en Chine et en France.

En deuxième lieu, un autre facteur qui joue un rôle non négligeable dans le choix de notre sujet d'étude, ce sont nos expériences parallèles à la recherche de Master. Nous étions très sensible au développement inouï et en nombre croissant des nouveaux mots et expressions dans la société chinoise, caractérisé, nous a-t-il semblé, par des créateurs multiples, une apparition imprévue, une utilisation fréquente voire virale sur plusieurs supports, que ce soit en ligne ou hors ligne. Ce sentiment de singularité fondé sur l'intuition et en relation avec nos propres pratiques d'internaute, et notamment d'utilisatrice du microblog, a été ensuite confirmé par les discours médiatiques dont nous avons eu connaissance, aussi bien dans la presse écrite (papier ou numérique) qu'audiovisuelle. Cette omniprésence de nouveaux phénomènes langagiers nous a incité à le mettre en lien avec les nouveaux moyens de communication offerts par les services de l'Internet et les pratiques personnelles dans le contexte numérique. Ces constats nous ont

orientée vers un premier ensemble de lectures d'ouvrages scientifiques, notamment de chercheurs chinois (Bouzidi 2010 ; Che 2012 ; Chen et Bai 2012 ; Chen 2001 ; Deng 2010 ; Issac 2011 ; Jiang 2006 ; Liu 2012 ; Wang 2012), et nous ont convaincue de l'actualité et de la pertinence de développer un questionnement qui permettrait de mieux comprendre un phénomène émergent, en l'observant dans deux contextes socio-culturels et linguistiques différents.

Ces éléments nous ont conduite à définir nos objectifs d'étude pour une recherche doctorale : comparer de nouvelles productions langagières en France et en Chine dans la supposition que les mêmes phénomènes surgissent dans les deux pays ; pouvoir observer un dispositif semblable ou similaire qui soit propice à la diffusion et la circulation en ligne. Pour ce deuxième objectif, nous avons choisi le microblog, en raison de sa vitalité, de la multiplicité de ses usagers, de ses fonctionnalités favorisant une communication rapide et relationnelle tout en ayant une influence sur la pratique linguistique, de par les contraintes d'expression exigée par le dispositif (par exemple la brièveté). En prenant en compte les situations d'utilisation des microblogs dans les deux pays, nous avons choisi d'observer respectivement *Sina Weibo* (plateforme chinoise) et *Twitter*, version française.

Une fois ces objectifs définis, nous avons cherché dans un premier temps à savoir comment nous pourrions situer ces phénomènes langagiers, et comment nous pourrions les dénommer, à quels concepts il convenait de se référer pour cerner les spécificités de ces nouvelles créations et pour appréhender le milieu numérique des échanges interpersonnels, le microblog au sein du Web et le Web lui-même au sein du réseau de l'Internet. Il nous est apparu que trois contextes étaient à envisager : le contexte numérique d'origine de la plateforme de microblog, le contexte d'expansion en ligne et hors ligne, le contexte scientifique particulièrement celui des Sciences du langage avec son outillage théorique et méthodologique.

Les services de l'Internet ouvrent de grandes possibilités d'expression et permettent ainsi à chaque individu de devenir un créateur potentiel dans le domaine langagier. Les pratiques des internautes évoluent également avec les différents produits numériques, enrichissant des situations d'énonciation. Quand les individus interviennent dans les processus de création, la créativité devient spontanée et imprévue et cela rend difficile de retracer la source d'un mot ou d'une expression donnés. Le gain de popularité de ces productions est également remarquable comme en témoigne la haute fréquence d'utilisation. Avec la communication instantanée, un

nouveau mot peut être connu par un grand public et beaucoup repris en très peu de temps. Ce public ne se compose pas seulement d'internautes : le public des médias traditionnels, les consommateurs des productions culturelles (revues, livres, etc.) peuvent aussi être concernés par l'expansion.

Dans le contexte scientifique, nous avons remarqué en Sciences du langage et de l'information-communication des travaux de recherche sur les changements technologiques et leurs relations avec les phénomènes langagiers, les manières de lire et d'écrire, la création lexicale (Baron 2003 ; Chen 2017 ; Chu 2016 ; Crystal 2001 ; Liu et Shi 2015 ; Shen et Pu 2018 ; Souchier 1996 ; Zhou 2000). Une première revue de littérature en nous faisant prendre connaissance de perspectives diversifiées, nous a permis de saisir de manière plus précise notre sujet, mais l'hétérogénéité des travaux offrant une abondance de termes ne nous a pas permis de trouver une appellation unanime pour désigner notre objet de recherche, et nous avons trouvé la même difficulté en nous informant sur les études consacrées à la néologie et à la création lexicale. Nous avons toutefois pu faire une première distinction : les nouveaux mots et expressions émergés avec le développement du numérique regroupent d'un côté les termes créés sur l'Internet et ses services (par ex. : les réseaux sociaux numériques) – c'est-à-dire s'y référant –, et d'un autre côté, les nouvelles formes utilisées dans la communication interpersonnelle notamment sur l'Internet, – c'est-à-dire en ligne. Notre étude s'intéresse à la deuxième catégorie. Pour l'explorer nous avons mobilisé la lexicologie en nous centrant sur les procédés de création et les travaux consacrés aux néologismes (Allam-Iddou 2015 ; Bernhard *et al.* 2018 ; Boulanger 2003 ; Dury *et al.* 2014 ; Guilbert 1973 ; Humbley *et al.* 2007 ; Lecolle 2012 ; Lombard et Huyghe 2020 ; Mejri 2011 ; Pruvost et Sablayrolles 2016 ; Sablayrolles 1996 ; Sablayrolles 2019 ; Steuckardt 2016 ; Tallarico 2020) dans le but d'étudier les caractéristiques des nouveautés et de faire en premier lieu une comparaison sur le plan linguistique. Au terme de l'examen de notions plus complexes que nous le pensions, nous avons fait le choix de dénommer notre objet « nouveautés langagières numériques ».

Pour observer et analyser la circulation de ces nouveautés langagières sur des microblogs il nous est apparu indispensable d'une part de mieux connaître le dispositif et d'autre part de réunir un corpus.

Les lectures de travaux en Science de l'information et de la communication (Boyd et Ellison 2007 ; Coutant et Stenger 2010 ; Ellison et Boyd 2013 ; Ellison et Thierry 2011 ; Girard et

Fallery 2009 ; Zammar 2012) nous ont permis de situer le microblog comme réseau social numérique, et ainsi de prendre en compte les spécificités du discours numérique qui ont une intense relationnalité, des traits linguistiquement inédits et une forte imprévisibilité sur le plan discursif selon Paveau (2017:10). Nous avons donc décidé de ne pas rester au seul plan linguistique mais d'aller vers une analyse discursive adaptée au numérique en nous inspirant de l'approche écologique préconisée par Paveau (2012 ; 2013 ; 2017).

La problématique centrale de cette recherche concerne donc l'analyse des caractéristiques linguistiques et discursives de nouveautés langagières en contexte numérique via le dispositif du microblog, dans une perspective comparative sur le plan linguistique, communicationnel et socioculturel.

Pour aborder cette comparaison, nous devons résoudre plusieurs difficultés relatives notamment à l'identification, à la quantité du corpus et nous avons peu à peu affiné les démarches de constitution en tenant compte de l'environnement discursif et technologique. Un ensemble de questions s'est posé concernant le corpus, sa constitution et les procédures d'analyse s'agissant d'une circulation très abondante en ligne.

Tout d'abord, notre objectif étant de collecter des nouveautés langagières perçues comme telles et attestées, nous pensons important de passer par la caution d'un métadiscours autorisé afin d'échapper au risque de subjectivité. Dans la mesure où les pratiques communicationnelles et institutionnelles concernant la langue ne sont pas les mêmes, nous serons amenée à différencier les démarches de sélection d'échantillons que nous voulons étudier. En effet notre démarche qualitative ne vise pas à constituer de grands corpus mais à dégager des traits significatifs à partir d'échantillons. Nous travaillons à partir de résultats de palmarès de mots et expressions, issus d'activités médiatiques valorisées en Chine et à partir des nouvelles entrées des dictionnaires de langue citées par certaines organes de presse de référence en France. Nous avons défini une période d'observation de neuf années (2009-2017) suffisamment longue pour témoigner d'une évolution, et en relation avec le moment où nous avons entamé notre recherche. Cette période est également délimitée en prenant en considération la mise en service de *Sina Weibo* en 2009 et une modification technique importante sur *Twitter* (la suppression de la limite des 140 signes dans le message) en 2017.

Nous avons vu la nécessité de constituer principalement deux sortes de corpus, l'un regroupe

un échantillon de nouveautés langagières qui servent à dégager les aspects de nouveauté dans la formation et les éventuelles caractéristiques, l'autre comporte des cas d'utilisations collectés sur microblog, à savoir les tweets contenant les nouveaux mots et expressions.

À partir des résultats de collecte obtenus (175 pour le chinois et 371 pour le français), nous établissons un premier corpus d'entrées lexicales, des nouveautés langagières que nous dénommons « corpus de listes ». Ensuite, notre perspective est d'observer les caractéristiques de ces nouveautés langagières dans le flux discursif des plateformes de microblog *Sina Weibo* et *Twitter*. Ces plateformes sont caractérisées par une production de données massives et continues, par une multiplicité d'interactions construites autour de diverses manifestations relationnelles. Nous sommes donc en présence d'un corpus en expansion, extrêmement volumineux. Nous optons dans le cadre d'une démarche qualitative en quête d'éléments significatifs, pour la constitution d'un échantillon de tweets à partir du corpus de listes des nouveautés langagières. Nous avons établi ainsi le deuxième corpus dit « corpus de tweets » avec 70 tweets chinois et 70 tweets français collectés, ces tweets comportant dans leur message certaines des nouveautés langagières du corpus de listes.

Recueillir et analyser des corpus hétérogènes issus de contextes différents sur le plan linguistique, culturel, technologique et discursif, en permettant tout de même une comparaison qui est un de nos objectifs, exige une adaptation régulière de la méthodologie et une réflexion sur cette dernière. Nos démarches et nos pistes de réflexions développées tout au long de ces démarches sont détaillées dans la thèse.

La constitution, l'observation et l'analyse du corpus ont nécessité de construire une méthodologie permettant de prendre en compte non seulement la matérialité des nouveautés langagières mais aussi l'environnement technologique et relationnel du dispositif sur lequel la construction du sens des nouveautés langagières se produit. Pour cela, cette étude s'inscrit dans une perspective interdisciplinaire.

La réflexion sur notre problématique nous a incité à élaborer les questions suivantes : quels sont les procédés de formation que nous pouvons constater dans les deux langues ? Quelles sont les pratiques favorisées par les fonctionnalités technologiques du microblog ? Quelles sont les influences de cette nouvelle modalité d'expression sur la production lexicale ? Quels rôles les relations construites autour d'un compte jouent-elles sur la transmission du message dans la

communication ? Quelle est la nature des évolutions sémantiques et formelles ?

Notre hypothèse principale est que la création et la circulation des productions langagières dépendent de la pratique des internautes et de l'environnement du dispositif sur lequel apparait cette pratique, quels que soient les pays et les langues.

Même si cet enrichissement langagier devient en plus en plus banal en ligne, voire dans la communication quotidienne et médiatique suite à une expansion d'utilisation, il reste toujours un sujet émergent dans le domaine de recherche. Malgré l'attention des chercheurs dès le début de 21<sup>e</sup> siècle, les travaux portant sur ce phénomène sont assez épars et il y a à ce jour très peu d'études qui visent à relier ces phénomènes langagiers en Chine et en France parmi les travaux chinois que nous avons consulté. De ce point de vue, nous espérons que notre recherche apportera des nouveaux éléments en proposant une analyse comparative fondée sur un corpus original.

Notre recherche se présente en deux volumes : le premier volume présente les contextes linguistiques, culturels et sociotechniques, le cadre conceptuel, la méthodologie et les analyses des corpus que nous avons constitués. Le deuxième volume est réservé aux annexes composées par les références périphériques, les corpus d'étude et les documents de présentation et d'analyse de ces corpus.

Le premier volume est divisé en deux parties avec au total sept chapitres. La première partie contient les trois premiers chapitres, qui ont pour objectif de contextualiser notre recherche. Dans le premier chapitre, nous tendons de lever le brouillard sur la langue chinoise en retraçant brièvement son histoire et en mettant en relief ses caractéristiques linguistiques et sociolinguistique, dans une perspective comparative. Le deuxième chapitre est consacré à éclaircir les notions portant sur le numérique, à savoir les notions de l'Internet et du Web, les réseaux sociaux, les réseaux sociaux numériques et les médias sociaux afin de mieux cerner le contexte communicationnel. Le troisième chapitre est crucial pour cerner notre objet de recherche à partir d'une tentative de mise au point sur plusieurs notions linguistiques relatives à la création lexicale, proposées par des chercheurs venant des pays différents et de plusieurs disciplines ; les néologismes qui constituent notre socle principal d'une part, le langage de l'Internet (pour reprendre la traduction d'une expression chinoise) d'autre part, retiennent notre attention ; nous présentons aussi la typologie que nous retiendrons pour l'analyse de ce que

nous appelons les « nouveautés langagières ».

La deuxième partie qui englobe les quatre chapitres restants est consacrée à exposer les démarches méthodologiques en vue de la collecte des corpus et à développer les analyses. Le quatrième et le cinquième chapitre décrivent respectivement la délimitation puis l'analyse des corpus de liste (un échantillon de nouveautés langagières) alors que le sixième et le septième sont consacrés à la constitution et l'analyse des corpus de tweets. Pour ces chapitres, l'ordre de traitement est le même : nous commençons par le corpus chinois et poursuivons avec le corpus français.





## **PARTIE I : CONTEXTUALISATION**



# CHAPITRE I

## Présentation du chinois : caractéristiques historiques, linguistiques et sociolinguistiques

### Introduction

Notre recherche vise une étude comparative sur les nouveautés langagières dans le champ de la communication numérique, portant sur un corpus français et un corpus chinois. Même si notre objectif n'est pas une comparaison des langues, ni une comparaison exclusivement linguistique des discours, la perspective est toutefois contrastive et il nous paraît nécessaire de consacrer un chapitre à la présentation de la langue chinoise, dont nous sommes locutrice native.

En effet la langue chinoise, complètement différente de la langue française, et bien qu'elle soit de plus en plus enseignée en France<sup>1</sup>, reste le plus souvent mystérieuse pour ceux qui n'en sont pas locuteurs et les caractéristiques en semblent floues et complexes. Malgré les échanges culturels, communicationnels ou technologiques, cette distance lointaine n'est toujours pas réellement diminuée. De ce fait, pour préparer aux analyses ultérieures, une présentation succincte de la langue chinoise sera proposée sans aucune prétention à l'exhaustivité, et uniquement dans le but de clarifier certains phénomènes.

Après un bref historique, nous adopterons un angle plus spécifiquement linguistique et sociolinguistique afin de dégager les caractéristiques et les variétés de cette langue. Enfin, nous aborderons la question de l'écriture surtout sous l'angle de la saisie informatisée, en relation directe avec notre domaine d'analyse, la pratique des réseaux sociaux.

---

<sup>1</sup> Les sites de référence :

<https://fcae.fr/pdf/Etat-du-chinois-DEC2015.pdf>;

[https://www.lemonde.fr/education/article/2013/04/23/l-enseignement-du-chinois-en-plein-boom-en-france\\_3164677\\_1473685.html](https://www.lemonde.fr/education/article/2013/04/23/l-enseignement-du-chinois-en-plein-boom-en-france_3164677_1473685.html).

## 1.1. Généralités sur la langue chinoise<sup>2</sup>

La langue chinoise est l'une des plus anciennes langues du monde avec le sumérien, l'hébreu, l'égyptien ancien, le grec, le latin etc. L'écriture ossécaille (甲骨文 jiǎ gǔ wén), environ 3000 ans avant J.-C.<sup>3</sup>, est considérée comme la toute première écriture chinoise. Dans son histoire, le système graphique de nature idéogrammatique et non phono-grammatique a été adopté et adapté par plusieurs langues comme par exemple le Japonais et le Coréen. Aujourd'hui, à part la Chine, l'usage de l'écriture chinoise est aussi constaté dans les autres pays comme par exemple Singapour. La langue chinoise est non seulement très ancienne et bien documentée, mais elle reste aussi une langue encore parlée et écrite actuellement, c'est-à-dire une langue vivante à forte dynamique.

De nombreux récits mythologiques évoquent l'origine du chinois, parmi lesquels, la légende de Cang Jie 仓颉造字 (cāng jié zào zì) répandue pendant la période des Royaumes combattants (de 475 av. J.-C. à 221 av. J.-C.) est la plus connue. On pense que 仓颉 (cāng jié) Cang Jie, qui est au service de 黄帝 (huáng dì) l'Empereur Jaune, père mythique de la civilisation chinoise qui aurait régné de 2717 av. J.-C. à 2599 av. J.-C., serait l'inventeur légendaire de l'écriture chinoise. Son apparence est décrite comme celle d'un homme ayant quatre yeux, caractéristique biologique des hommes sages en Chine. Cela lui permet d'observer plus soigneusement les choses et ensuite de se procurer des inspirations pour la fabrication de l'écriture chinoise. On raconte que Cang Jie aurait compris en observant les empreintes des animaux sur le sol (en particulier les pattes des oiseaux), comment noter les idées et les choses au moyen de signes distinctifs.<sup>4</sup>

Bien évidemment, la langue chinoise comme toutes les autres langues ne peut avoir été créée par une seule personne. L'apparition et le développement de cette langue sont en relation avec l'histoire des peuples à travers l'histoire. Désigner cette langue par l'expression « le chinois » ne doit pas faire oublier que la langue chinoise actuelle a connu différentes étapes dans son

---

<sup>2</sup> Pour cette présentation générale nous avons eu recours à des références encyclopédiques en ligne, notamment : <https://www.un.org/zh/observances/chinese-language-day>  
[https://www.chinadaily.com.cn/chinawatch\\_fr/2018-02/23/content\\_35726294.htm](https://www.chinadaily.com.cn/chinawatch_fr/2018-02/23/content_35726294.htm)  
<https://www.pourlascience.fr/sd/histoire-sciences/lecriture-chinoise-mythes-et-realites-4494.php>.

<sup>3</sup> Dans ce travail, toutes les dates sont exprimées par rapport au repère chronologique chrétien qui a un usage courant et international.

<sup>4</sup> Voir : [https://chine.in/guide/cang-jie\\_2374.html](https://chine.in/guide/cang-jie_2374.html).

évolution et qu'elle est bien différente du chinois archaïque tant sur le plan morphologique que sur le plan phonétique.

Dans un premier temps, il nous est nécessaire de clarifier certaines appellations autour de la langue chinoise. Nous pouvons souvent constater la désignation sous un nom commun « chinois » les diverses langues chinoises, y compris le chinois simplifié (简体中文, jiǎn tǐ zhōng wén), le chinois traditionnel (繁体中文, fán tǐ zhōng wén) et le pinyin (拼音, pīnyīn). Les deux premiers concernent en fait deux systèmes d'écriture idéographique différents. Il est à noter que la notion chinoise équivalente à ce système d'écriture est 中文 (zhōng wén). Quant au pinyin, il s'agit d'un système de transcription de l'écriture chinoise simplifiée.

Le chinois est également connu sous un autre nom occidental « le mandarin ». Il désigne depuis le 16<sup>e</sup> siècle (Calvet 1987:163) la langue officielle de la Chine. Aujourd'hui, il désigne spécialement la langue officielle dont l'usage s'est développé à partir des années 1950 pour faciliter la communication dans tout le pays. Le nom chinois correspondant proposé par le gouvernement est 普通话 (pǔ tōng huà) dont la traduction littérale est « la langue commune/générale ». La prononciation du mandarin est basée sur le dialecte du Nord. Deux étapes principales ont été décisives pour sa formation :

1. la simplification graphique des caractères (l'écriture simplifiée a donné le chinois simplifié);
2. la transcription phonétique des caractères (le pinyin). (voir *infra*)

Dans les domaines scientifiques chinois, notamment celui de la linguistique, les chercheurs adoptent l'appellation 汉语 (hàn yǔ), la langue des Hans, pour désigner le chinois. Actuellement en Chine, à part le mandarin, il y a aussi de nombreuses langues régionales ou dialectes. Ces dialectes qui existent depuis plus longtemps que le mandarin ont toujours une importance non négligeable dans la vie quotidienne des Chinois. Les chercheurs ont pris conscience de ces variétés linguistiques et définissent la langue chinoise en usage ainsi : au sens large, le chinois englobe à la fois le mandarin et les autres variétés de la langue. Ces dialectes variés recourent au même système d'écriture que le mandarin, mais se prononcent bien différemment.

La langue chinoise appartient à la famille des langues sino-tibétaines<sup>5</sup> alors que le français appartient à la famille<sup>6</sup> des langues romanes au sein des langues indo-européennes. Contrairement à la plupart des langues indo-européennes, la langue chinoise est aussi une langue tonale. Les tons (marqués à l'écrit par différents signes diacritiques placés au-dessus des lettres) sont une partie indispensable de la structure phonétique de la langue chinoise. Les caractéristiques spécifiques de la langue chinoise les plus marquantes résident encore dans ses signes écrits, une écriture idéographique comportant des caractères, mais aussi dans sa morphosyntaxe. Les caractères sont monosyllabiques et invariables. Il n'y a pas de distinctions de catégories de genre, de nombre, de temps, d'aspect ou de mode, comme c'est le cas notamment en français. Mais l'ordre des mots ou expressions dans une phrase influence la relation grammaticale entre les unités, et le sens. Le français est une langue flexionnelle alors que le chinois est une langue isolante ce qui n'est pas sans conséquence sur la formation des néologismes comme nous le verrons dans nos analyses.

Suite au développement de la population, aux différents contacts et mouvements migratoires à l'origine d'une importante diaspora, ainsi qu'à l'affirmation de la puissance économique et politique du pays, la langue chinoise est devenue la plus parlée du monde. Selon les statistiques sur *Ethnologue*<sup>7</sup>, en 2020, le nombre des locuteurs natifs du mandarin atteignait 897 millions, soit presque 13 fois la population française<sup>8</sup>. Ces locuteurs sont présents dans 14 pays du monde.

Une langue vivante est toujours en évolution et elle est en rapport étroit avec les phénomènes sociaux. Ainsi face à l'essor de l'Internet, les modes de communication interpersonnelle ont changé et s'adaptent aux technologies de l'époque actuelle. Dès lors les méthodes d'écriture

---

<sup>5</sup> « La famille sino-tibétaine est la deuxième famille la plus importante du monde en termes de nombre de locuteurs après l'indo-européen. Elle comprend plus de 300 langues, parlées en Chine, en Birmanie, en Thaïlande, au Bhoutan, au Népal, dans les Etats du nord-est de l'Inde (Arunachal pradesh, Nagaland, Manipour, Mizoram, Meghalaya, Tripura, Assam) ainsi que du nord-ouest (Himachal pradesh, Ladakh), au Bangladesh et enfin au Pakistan (Baltistan). Elles présentent une importante diversité typologique, comprenant à la fois des langues isolantes telles que la plupart des dialectes actuels du chinois et des langues à la morphologie verbale foisonnante telles que le limbu » (Jacques 2006).

<sup>6</sup> La classification des langues selon des familles renvoie à une relation génétique supposée entre des langues, souvent avec une proximité géographique. La typologie des langues repose elle sur des critères de parenté structurelle.

<sup>7</sup> Source: « Summary by language size » sur Ethnologue: <http://www.ethnologue.com/statistics/size>.

<sup>8</sup> Il ne s'agit pas ici de comparer le nombre des locuteurs du mandarin et du français. L'usage de la langue française n'est toutefois pas limité à la France car il s'agit d'une langue internationale francophonie est représentée sur tous les continents. Le nombre de locuteurs du français dans le monde est de 75.9 millions.

sont appelées à innover. Avec des claviers d'ordinateur, des tablettes ou des smartphones presque identiques, comment saisit-on le français et comment saisit-on le chinois quand on est usager des réseaux sociaux numériques ? La saisie informatique de la langue chinoise a aussi connu des étapes d'évolution suite aux développements langagiers et techniques et nous y consacrerons le dernier point de ce chapitre.

## 1.2. Aperçu historique

### 1.2.1. Plusieurs divisions en périodes

Les sinologues n'ont pas tous le même point de vue sur la division des périodes de développement de la langue chinoise. Nous allons en citer quatre puis procéderons à une comparaison de leur périodisation.

Shuxiang Lv, éminent linguiste et sinologue très apprécié en Chine, a présenté le développement du chinois en deux grandes périodes dans son livre (Lv 1984:89) :

1- **Le chinois archaïque** (古代汉语, gǔ dài hàn yǔ)

2- **Le chinois moderne** (近代汉语, jìn dài hàn yǔ)

Selon lui, la distinction de ces deux grandes périodes peut s'expliquer par deux styles de réalisation : le chinois littéraire (文言, wén yán) et le chinois vernaculaire (白话, bái huà). Le premier se caractérise par un style compact et obscur, enrichi de figures rhétoriques. Ce sont plutôt les lettrés qui peuvent le maîtriser. Un caractère peut être associé à un sens compact que nous devons exprimer par une phrase d'aujourd'hui. Le deuxième est un style explicatif proche de l'expression orale. Donc, il est plus facile à apprendre et à comprendre. La période où les premiers écrits en langue vernaculaire commencent à se former est considérée comme une ligne de démarcation dans l'histoire du chinois. Cette période peut remonter à la fin de la Dynastie Tang (618-907)<sup>9</sup>.

Li Wang<sup>10</sup>, l'un des fondateurs de la linguistique du chinois contemporain, en se basant quant à

---

<sup>9</sup> Voir Annexe 1: Tableau Chronologique des dynasties chinoises.

<sup>10</sup> Li Wang (1900-1986).



lui exclusivement sur la langue chinoise écrite, divise l'histoire du chinois en quatre périodes (Wang 1980) :

- 1- **Le chinois archaïque** (上古汉语, shàng gǔ hàn yǔ) s'est développé avant le 3<sup>e</sup> siècle. Sa formation peut remonter jusqu'à la dynastie des Shang (environ 1570 à 1045 av. J.-C.) où les textes sont gravés sur des plastrons de tortues et sur des os de bovidé, appelé aussi « 甲骨文 » (jiǎ gǔ wén, *litt.* « carapace os texte »). La fin de son développement se voit jusqu'à la dynastie des Qin (221-207 av. J.-C.) ;
- 2- **Le chinois médiéval** (中古汉语, zhōng gǔ hàn yǔ) se forme entre le 4<sup>e</sup> et 12<sup>e</sup> siècle, après une période de transition (entre le 3<sup>e</sup> et le 4<sup>e</sup> siècle) ;
- 3- **Le chinois moderne** (近代汉语, jìn dài hàn yǔ) s'est développé du 13<sup>e</sup> au 19<sup>e</sup> siècle ;
- 4- **Le chinois contemporain** (现代汉语, xiàn dài hàn yǔ) se développe au 20<sup>e</sup> siècle (après le Mouvement du 4 Mai 1919)<sup>11</sup>, qui n'aurait donc que cent ans d'histoire.

Le linguiste Bernhard Karlgren, fondateur de la sinologie comme discipline scientifique en Suède, s'est appuyé pour sa part sur la phonologie chinoise pour diviser l'histoire du chinois en 5 périodes (Karlgren 1915). Il écrit :

En effet, nous verrons que les fan-tsc'ie ne donnent que des bilabiales, tandis qu'il y a des indices que la série dentilabiale s'est développée dès l'époque Souei<sup>12</sup>. Pour ne courir aucun risque, disons qu'il s'agit, pour les fan-tsc'ie, d'une langue entre 500 et 600 apr. J.C. Nous pourrions appeler cette langue **l'ancien chinois**. Les phases précédentes pourraient être appelées **le chinois archaïque** et **le proto-chinois** ; les phases postérieures **le chinois moyen** — celui des tables de rimes; et **le vieux mandarin** — celui du *Hong-wou teheng yun*.

Ce qui nous donne les 5 périodes suivantes :

- 1- Le chinois archaïque
- 2- Le proto-chinois

---

<sup>11</sup> Il s'agit d'un mouvement nationaliste en Chine, mené par des étudiants et intellectuels contre les prétentions de l'empire du Japon sur la Chine. Il marque l'émergence en Chine d'une conscience patriotique. Il renforce en même temps le mouvement de la nouvelle culture qui avait commencé quelques années avant 1919 (1915-1921). Cette révolution culturelle remet en question l'héritage culturel chinois face à la confrontation de la civilisation occidentale. Un enseignement pluridisciplinaire ouvert sur l'occident est principalement poursuivi. Un courant anti-wenyan est ainsi né. L'établissement d'une nouvelle langue nationale est proposé avec l'idée de marquer linguistiquement l'existence d'une « nouvelle » Chine.

<sup>12</sup> Dynastie des Sui (581-618).

- 3- L'ancien chinois (entre 500 et 600 ap. J.-C)
- 4- Le chinois moyen
- 5- Le vieux mandarin

On retrouve une division en 5 périodes chez le sinologue japonais Tatsuo Ota (1991) :

- 1- le **chinois archaïque** (上古汉语, shàng gǔ hàn yǔ), englobe trois sous-périodes : celle des Dynasties Shang (environ 1570 à 1045 av. J.-C.) et Zhou (environ 11<sup>e</sup> siècle-770 av. J.-C.), Époque des Printemps et Automne, et des Royaumes combattants (770-221 av. J.-C), Dynasties des Han (Han de l'Ouest 206 av.j.-C - 24 ap. J.-C; Han de l'Est 25-220) ;
- 2- le **chinois médiéval** (中古汉语, zhōng gǔ hàn yǔ) s'est développé de la période des Trois Royaumes (220-265) jusqu'à la Dynastie du Sud et du Nord (420-589) ;
- 3- le **chinois ancien proche** (今古汉语, jīn gǔ hàn yǔ) commence sa formation durant la Dynastie Tang (618-907) et s'est développé jusqu'à la Dynastie Ming (1368-1644) ;
- 4- le **chinois moderne** (近代汉语, jìn dài hàn yǔ) a été formé et utilisé pendant la Dynastie Qing (1644-1911) ;
- 5- le **chinois contemporain** (现代汉语, xiàn dài hàn yǔ) a connu sa naissance après la Révolution chinoise de 1911.

Voici un tableau récapitulatif :

<b>Représentants des principales écoles de périodisation</b>	<b>Périodes</b>	<b>Nombre de périodes</b>
Lv 1984	chinois archaïque (avant 618 ap. J.-C) chinois moderne (à partir de 618 ap. J.-C)	2
Wang 1980	chinois archaïque (1570 av. J.-C à 207 av. J.-C) chinois médiéval (4 <sup>e</sup> siècle au 12 <sup>e</sup> siècle) chinois moderne (13 <sup>e</sup> au 19 <sup>e</sup> siècle) chinois contemporain (à partir du 20 <sup>e</sup> siècle)	4
Karlgren 1915	chinois archaïque proto-chinois ancien chinois (500 et 600 ap. J.-C) chinois moyen vieux mandarin	5
Ota 1991	chinois archaïque (1570 av. J.-C.-) chinois médiéval (220-589 ap. J.-C) chinois « ancien proche » (618-1644 ap. J.-C) chinois moderne (1644-1911 ap. J.-C) chinois contemporain (à partir de 1911 ap. J.-C)	5

**Tableau 1 : Les principales écoles de périodisation du développement de la langue chinoise**

En se basant sur les divers aspects du développement de la langue chinoise, ces linguistes ont établi des critères différents qui amènent inévitablement une divergence d'opinions. Lv souligne les différences de style d'expression alors que Wang prend en compte les changements des caractères eux-mêmes ; quant à Karlgren, il s'intéresse plutôt au développement de la phonologie chinoise ; alors que Ota, parle aussi des styles d'expression mais d'une manière plus détaillée. Toutefois ces divergences ne sont pas irréductibles. En effet si l'on met à part la dichotomie de Lv, les trois autres linguistes ont des divisions qui sont d'une part plus précises et d'autre part assez similaires. Certes, en choisissant ces sinologues, ce ne sont que les principales écoles de division parmi tant d'autres classifications que nous avons rappelées car il n'était pas possible ni opportun de toutes les énumérer ici.

Pourtant, à l'aide de ces périodisations, on peut avoir une vision globale de l'histoire de la langue chinoise, afin de définir le contexte d'étude de cette recherche. L'étude de l'histoire de la langue chinoise est si vaste qu'elle reste toujours un sujet inachevé. Nous pouvons retenir de notre approche sommaire que la notion de « langue chinoise » ne saurait désigner seulement une langue d'une certaine période, pas plus qu'elle ne renvoie à une langue unique.

La majorité des chercheurs considèrent qu'à partir du 20<sup>e</sup> siècle, une « nouvelle langue

chinoise » commence à se former. Dans notre recherche, nous nous référerons au chinois de notre époque, c'est-à-dire au chinois contemporain dont nous allons préciser la définition.

### 1.2.2. Le chinois contemporain (现代汉语, xiàn dài hàn yǔ)

Dans un premier temps, regardons la définition du chinois et du mandarin que propose le dictionnaire français *Le Petit Robert* dans son édition de 2012 : « le chinois : ensemble des langues parlées en Chine ; SPECIALT dans la région de Pékin (Mandarin\*) » ; « mandarin : langue chinoise moderne, véhiculaire, parlée dans le nord, le centre et le sud-est de la Chine ».

On peut faire plusieurs remarques à propos de ces définitions. Le terme « chinois » ne désigne pas une seule langue mais un ensemble de langues associées au territoire de la Chine. Parmi ces langues parlées en Chine, le mandarin tient une place importante. Il est qualifié de langue véhiculaire<sup>13</sup>, c'est-à-dire une langue qui n'est pas forcément la langue première des locuteurs concernés mais qui permet à des groupes, locuteurs de langues différentes, de communiquer et de se comprendre. Le mandarin est également considéré comme « moderne » mais ici l'adjectif ne s'oppose pas à contemporain.

Précisons tout d'abord qu'en Chine, le chinois contemporain est devenu aujourd'hui une discipline scientifique. Dans un manuel universitaire tel que *现代汉语* (Yang *et al.* 2010) qui sert d'ouvrage de référence pour les cours de licence sur la langue chinoise, cette notion linguistique est définie plus précisément. Au sens large, le chinois contemporain englobe la langue standard<sup>14</sup> des Hans (ethnie majoritaire)<sup>15</sup>, les variétés régionales (appelées dialectes), les variétés sociales, les variétés fonctionnelles, les variétés individuelles, etc. Au sens strict, il indique seulement la langue commune des Hans qui se base principalement pour la prononciation standard sur le parler de Pékin, pour le vocabulaire, sur les parlers du nord de la Chine, et pour la grammaire, sur les écrits vernaculaires. Il est nommé 普通话 (pǔ tōng huà)<sup>16</sup>

---

<sup>13</sup> En sociolinguistique, véhiculaire s'oppose à vernaculaire qui désigne la langue d'une communauté. Pour une définition de la véhicularité voir :

<http://www.wikilf.culture.fr/barometre2012/tmpl.php?data=doc/facteur/vehicularite&crt=2&type=default>.

<sup>14</sup> Il s'agit de la langue officielle en usage qui s'est développée depuis longtemps. Pour des raisons historiques et politiques, cette langue des Hans a été choisie comme une langue intermédiaire.

<sup>15</sup> Au total 56 ethnies différentes sont reconnues officiellement en Chine. Les Hans représentent 92% de la population, soit plus de 1300 millions de personnes, constituant l'ethnie majoritaire en Chine.

<sup>16</sup> Qu'on peut traduire comme langue générale ou langue standard.

qu'on traduit souvent en Occident par le nom « mandarin ». En somme, le chinois contemporain est la langue utilisée aujourd'hui par la nation chinoise.

D'un point de vue diachronique, le chinois contemporain se distingue des états antérieurs que sont le chinois archaïque, l'ancien chinois mais constitue aussi un héritage des langues chinoises des périodes différentes. D'un point de vue synchronique et non restrictif, le chinois contemporain contient le mandarin (la langue commune de la nation chinoise) et toutes les autres variétés. Il est clair que quand on parle actuellement de la langue chinoise, on fait référence plutôt au mandarin, la langue officielle en Chine. Cependant, la réalité linguistique en Chine est assez complexe car il existe effectivement une grande variété de langues qui sont considérées comme des dialectes régionaux. Ainsi, le mandarin lui-même s'est aussi développé à partir des parlers du Nord de la Chine.

Tout au long de l'histoire de la Chine, la normalisation linguistique est une préoccupation des autorités. Avec moins de 100 ans d'histoire, le chinois contemporain a tellement évolué qu'il se distingue fortement de la langue dont il est issu.

La normalisation du chinois contemporain date du début du 20<sup>e</sup> siècle. C'est un processus mis en œuvre à l'échelle nationale par les gouvernements chinois d'époques différentes. Si on prend l'année 1949 (fondation de la République populaire de Chine) comme ligne de démarcation, on peut diviser ce processus en deux grandes phases. C'est à travers ces étapes que se sont formés principalement le système de romanisation et la nouvelle écriture du chinois.

La première phase (avant 1949) est caractérisée par trois mouvements qui ont des incidences sur les plans phonétique, lexical et grammatical :

1. l'alphabet latin est suggéré pour transcrire phonétiquement la langue chinoise. Il conduit à la formation en 1913 d'une façon de transcription qui est connu sous le nom 老国音 (lǎo guó yīn, la traduction littéraire est que: vieille méthode de prononciation nationale) en s'opposant à 拼音 (pīnyīn). Ce dernier constitue la méthode récente ;
2. le parler du Nord devient la langue standard de toute la nation sur décision du gouvernement de l'époque vers l'année 1913 ;
3. le style vernaculaire (白话, bái huà) est proposé comme style d'écriture en remplaçant le

style littéraire (文言, wén yán). Cela influence la grammaire de l'écrit qui est standardisée ultérieurement.

Durant la deuxième phase (à partir de 1949), tout en continuant dans la même orientation, le gouvernement chinois a établi des critères plus unifiés sur les différents plans et ainsi fait naître le chinois qu'on connaît aujourd'hui. Il faut noter que dans la deuxième phase, les réformes, dans l'objectif de faciliter l'apprentissage du chinois à l'échelle nationale et de promouvoir l'alphabétisation, ne prennent pas les dialectes comme sujet de normalisation.

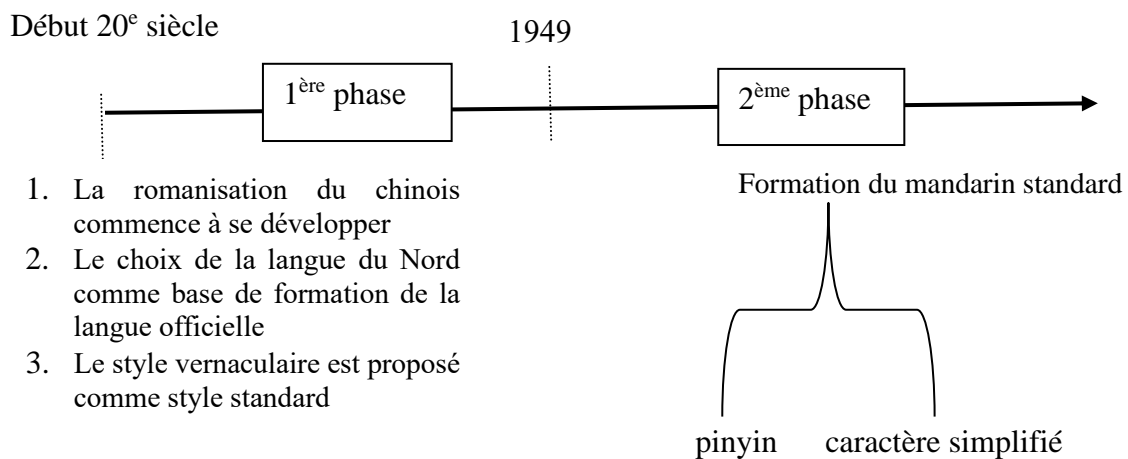


Figure 1 : Aperçu du processus de la normalisation de la langue chinoise

Comment a évolué le mandarin ? Quelles sont les relations entre les dialectes et le mandarin ? Cette diversité du chinois contemporain influence-t-elle la formation des néologismes ? Avant de répondre à ces questions, il est nécessaire de faire tout d'abord un rapide état des lieux de ces « dialectes » chinois.

### 1.3. Les dialectes chinois

Le terme de *dialecte* s'apparente à d'autres mots comme *langue*, *parler*, *patois*, *idiome*. Selon que l'usage est commun ou spécialisé leur définition n'est pas la même. Nous le comprenons en comparant plusieurs dictionnaires sur les entrées *dialecte* et *parler*.

Selon *Le Dictionnaire de l'Académie Française* (1986)<sup>17</sup>, *dialecte* se définit ainsi : « Variété régionale d'une langue ». Le substantif *parler* a pour acception : « Ensemble des façons de parler propres à une classe d'individus, à une région, et caractérisées par des tournures, des vocables et un accent particuliers. Le parler des écoliers. Les parlers locaux, méridionaux. ».

Il apparaît donc que *dialecte*, s'il est défini comme une variété régionale de la langue correspond à une variation diatopique et nous donne seulement un indice géographique. Dans la définition du *parler*, on peut constater deux niveaux de sens coordonnés et une ouverture à différents types de variation : le parler lié à une classe sociale et le parler concernant des moyens d'expression d'une région. Il est donc difficile de distinguer un dialecte d'un parler, ce dernier terme englobant le premier.

Voyons à présent les définitions du *Dictionnaire Le petit Larousse* (2012) :

**DIALECTE** : n.m. [...] Forme régionale d'une langue considérée comme un système linguistique en soi...SPECIALT système linguistique qui n'a pas le statut de langue officielle ou nationale, à l'intérieur d'un groupe de parlers.

**PARLER** : Ensemble des moyens d'expression employés par un groupe à l'intérieur d'un domaine linguistique. Dialecte, idiome, langue, patois. Les parlers régionaux.

L'élément d'information nouveau pour *dialecte* concerne son statut. Un dialecte s'oppose à une langue officielle, il n'a pas la même reconnaissance ni la même fonction sociale, et on suppose aussi qu'il n'a pas la même extension, en raison de l'opposition nationale/régionale. Le parler englobe toujours le dialecte puisque ce dernier appartient à un « groupe de parlers ». La définition de *parler* est toujours floue puisqu'il s'agit des façons de s'exprimer dans un milieu linguistique limité. Le rapport hiérarchique entre un dialecte et un parler n'est pas élucidé.

Un ouvrage spécialisé *Le Dictionnaire de linguistique* (Dubois *et al.* 2001) apporte un éclairage aux trois notions de *langue/dialecte/parler* en les différenciant tout en montrant la parenté entre eux :

**DIALECTE** : 1. Employé couramment pour dialecte régional par opposition à « langue », le dialecte est un système de signes et de règles combinatoires de même origine qu'un autre système

---

<sup>17</sup> Il s'agit de la 9<sup>e</sup> édition, en ligne :

[http://atilf.atilf.fr/Dendien/scripts/generic/showps.exe?p=main.txt;host=interface\\_academie9.txt;java=no](http://atilf.atilf.fr/Dendien/scripts/generic/showps.exe?p=main.txt;host=interface_academie9.txt;java=no).

considéré comme la langue, mais n'ayant pas acquis le statut culturel et social de cette langue indépendamment de laquelle il s'est développé [...]. 2. Le dialecte social est un système de signe et de règles syntaxiques utilisé dans un groupe social donné ou par référence à ce groupe.

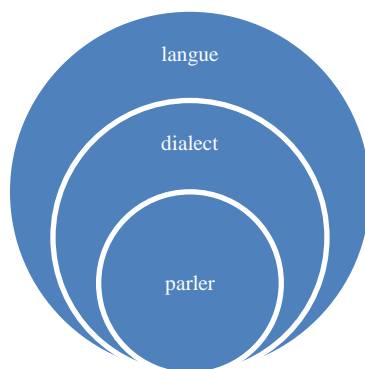
**PARLER** : 1. par opposition au dialecte [...] le parler est un système de signes et de règles de combinaison défini par un cadre géographique étroit (vallée, par exemple ou village) et dont le statut social est indéterminé au départ. Une langue ou un dialecte étudiés en un point précis sont donc étudiés en tant que parlers. 2. Le parler est une forme de la langue utilisée dans un groupe social déterminé [...]

Dans un premier temps, sont distinguées deux catégories pour l'entrée *dialecte* : le dialecte régional et le dialecte social (pour ce dernier on pourrait s'attendre au terme *sociolecte*). Ensuite, on peut constater, comme dans le dictionnaire précédent que ce qui différencie un dialecte et une langue, ce ne sont pas des critères linguistiques liés au système qui les constituent mais des facteurs extralinguistiques : les aspects géographique, culturel, social et politique sont déterminants. Une langue a une position plus dominante qu'un dialecte.

Quant à *parler*, on remarque que, tout d'abord, il est différent d'un dialecte, avec un rapport mentionné immédiatement : l'aire géographique du parler est assez petite mais son statut social n'est pas défini. Mais la suite de la définition remet cette distinction en cause puisque langue et dialecte peuvent être étudiés comme des parlers si c'est en relation avec un « espace précis ». Donc le parler serait plus une perspective d'étude qu'une catégorie bien circonscrite. La deuxième définition de parler qui le rattache à un groupe semble coïncider plus ou moins avec le dialecte social.

Si on reprend les définitions de *dialecte* et *parler*, on constate donc qu'elles sont similaires, ce qui nous fait rencontrer des difficultés à les différencier. La seule piste est l'extension géographique. Comme un parler est défini par un cadre étroit, on peut donc supposer qu'un parler est un dialecte plus limité. De ce point de vue, le rapport entre la langue, le dialecte et le parler peut être décrit par la figure suivante :





**Figure 2 : Le rapport interprété entre la langue, le dialecte et le parler**

Nous utiliserons le terme de dialecte en conservant l'idée que le dialecte consiste principalement en une variété géographique d'une langue donnée, qui n'a pas une position officielle dans la société. Dans le chinois contemporain, les dialectes sont considérés comme des variétés de la langue chinoise, et sont classés selon des divisions géographiques. Il est à noter que les dialectes chinois sont différents du mandarin tant sur le plan phonétique que sur le plan graphique. À part son statut standard, le mandarin est en réalité une des variétés de langue chinoise puisque sa prononciation est basée sur la langue du Nord. Pour son écriture, il s'agit d'un nouveau système qui est le résultat de la simplification des caractères chinois qui existent depuis longtemps. Or, les caractères chinois employés dans les dialectes ne connaissent pas de normalisation et ils sont la plupart du temps très peu connus au niveau national, s'appelant 生僻字 (shēng pì zì, *litt.* « mot peu fréquent »). Suite aux échanges fréquents entre les régions et au développement de l'Internet, les tendances à transcrire les dialectes par le chinois simplifié ou à introduire des caractères dialectaux dans le mandarin sont de plus en plus constatés.

Ce terme de dialecte correspond au terme chinois 方言 (fāng yán). 方 signifie « endroit, région » et 言 signifie « parole, parler ». Selon ce mot un dialecte est une variété langagière liée à un espace. À l'époque actuelle en Chine, les dialectes sont toujours très couramment utilisés dans la vie quotidienne, surtout à l'oral. Les locuteurs maîtrisent à la fois le mandarin et le(s) dialecte(s) et s'adaptent aux situations de communication en alternant la langue utilisée. Puis, les productions culturelles en dialectes ne se manquent pas : des programmes diffusés sur les chaînes télévisuelles locales, des spectacles, des feuilletons, des films voire des mêmes créés par des internautes. L'emploi des dialectes dans ces productions peut s'expliquer par les visées

suivantes : 1. pour rapprocher les locuteurs et les cultures locales et créer un sentiment de connivence ; 2. pour créer certains effets, comme par exemple un effet comique, souvent constaté dans des sketches ou films. Comme dans les autres langues, les dialectes chinois diffèrent selon les régions, mais le degré de variation des dialectes chinois est difficile à imaginer. Il se peut que les habitants de deux villes voisines ne se comprennent pas mutuellement à l'oral. Ces différences sont aussi importantes que celles entre le français et l'espagnol pour donner un exemple au sein des langues romanes. La complexité des réalités linguistiques, sociales et géographiques amène toujours des difficultés à catégoriser les dialectes. Les chercheurs chinois n'ont pas un avis unifié sur ce point. Nous adoptons dans la partie suivante les sept grandes catégories<sup>18</sup> proposées par le Ministère de l'Education (République de Chine) comme une source de référence :

---

<sup>18</sup> Voir : [http://www.moe.gov.cn/jyb\\_sjzl/wenzi/201612/t20161219\\_292430.html](http://www.moe.gov.cn/jyb_sjzl/wenzi/201612/t20161219_292430.html).

<b>Catégories de dialectes en Chine</b>	<b>Distribution géographique des dialectes</b>
<b>La langue du Nord</b> (北方方言, běi fāng fāng yán)	Fondements du mandarin ; utilisé dans les régions du Nord, notamment à Pékin et ses environs.
<b>Le WU</b> (吴方言, wú fāng yán)	Aussi connu sous le nom de dialecte de Jiangnan (江南方言, jiāng nán fāng yán), représenté par les parlers des villes à l'Est de la Chine, par exemple : Suzhou et Shanghai.
<b>Le XIANG</b> (湘方言, xiāng fāng yán)	Se nomme aussi le dialecte de Hunan (湖南方言, hú nán fāng yán), utilisé dans les régions Sud-Est de la Chine, représenté par le parler de Changsha, ville capitale de la province de Hunan.
<b>Le GAN</b> (赣方言, Gàn fāng yán)	Appelé également le dialecte de Jiangxi (江西方言, jiāng xī fāng yán), représenté par le parler de Nanchang. Les endroits concernés se situent au centre de la Chine.
<b>Le HAKKA</b> (客家方言, kèjiā fāng yán)	Représenté par le parler du district Mei de la province Guangdong. Il s'agit de l'un des dialectes du Sud de la Chine.
<b>Le YUE</b> (粤方言, yuè fāng yán)	Le parler de Guangdong, représenté par le cantonais et connu sous ce nom occidental de « cantonais », est principalement utilisé dans les provinces du Sud, y compris la province du Guangdong, du Guangxi, à Hongkong et à Macao.
<b>Le MIN</b> (闽方言, Mǐn fāng yán)	Il est beaucoup parlé par les habitants de Sud-Est de la Chine. Il s'appelle aussi le dialecte de Fulao (福佬话, fú lǎo huà), représenté par le parler de Fuzhou, une ville capitale de la province Fujian.

**Tableau 2 : Catégories et distribution géographique des dialectes en Chine**

Nous avons énuméré les sept catégories dans la colonne gauche du tableau. La colonne de droite est consacrée à des précisions sur l'aire géographique de ces dialectes ainsi qu'aux différentes dénominations. Selon la discussion précédente portant sur les rapports entre une langue, un dialecte et un parler, nous utilisons le terme parler ici comme subdivision des dialectes. C'est-à-dire que le parler est utilisé dans un espace plus petit que l'aire géographique du dialecte. Pour donner un aperçu visuel de la répartition géographique des dialectes, nous présentons une carte ci-dessous (Figure 3) :

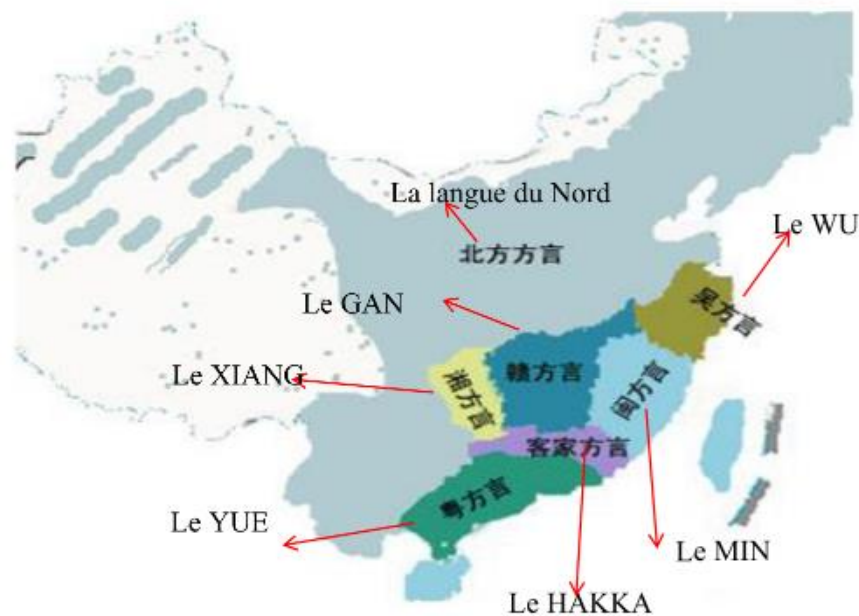


Figure 3 : Carte géographique des dialectes en Chine<sup>19</sup>

L'implantation, la diffusion, l'usage des dialectes et leurs relations entre eux et avec les langues reconnues comme telles sont dûs à des facteurs variés qui se combinent ou se contrecarrent : des facteurs sociaux, économiques, géographiques ou politiques. Le fait que la langue du Nord a été choisie comme le fondement de la langue officielle dépend principalement de raisons politiques. En effet le Nord a été le centre politique dans l'histoire de la Chine pendant près d'un millier d'années. Représenté par le parler de Pékin, la langue du Nord a joué le rôle d'un outil de centralisation, sous le nom de 官话 (guān huà), langue utilisée dans les milieux officiels. Son influence a été beaucoup plus grande que celle des autres dialectes dans tout le pays. Le mandarin se distingue des dialectes avant tout sur le plan phonétique, ensuite sur le plan lexical. Par exemple, pour exprimer « être fort en... », les expressions correspondantes dans le mandarin, le parler de Shanghai et le parle de Hefei sont les suivants : 厉害 (lì hài),

<sup>19</sup> Source : <http://dl.cau-edu.net.cn/kfzy/kcxx/zgctwhgl/yywzg/se01/slide/slide03.html>.

结棍 (jié gùn) et 过劲 (guò jìn). Les deux dernières expressions dialectales sont associées aux caractères simplifiés pour réaliser une transposition phonétique. À première vue, leurs prononciations n'ont rien de commun. Les différences phonétiques conduisent ensuite à un choix lexical distinctif. Pourtant, d'un point de vue macro, ils ont des règles de grammaires presque identiques que nous rappellerons *infra*. Pour cette dernière raison, les dialectes en Chine ne sont pas considérés comme des systèmes de langue indépendants mais des variétés du chinois contemporain. Par ailleurs, chaque catégorie de dialecte peut aussi être divisée en des sous-catégories qui peuvent continuer à être classées en des unités plus petites, par exemple : les parlers de Suzhou et Shanghai sont deux variétés du dialecte WU. Si on continue à prendre le parler de Shanghai comme exemple, celui utilisé dans le centre-ville est différent de ceux qui sont utilisés dans les banlieues, constituant ainsi des sous-variétés. Il est certain que le mandarin et les dialectes entrent en contact et s'influencent mutuellement au cours du temps.

Dans la présente recherche, nous nous concentrerons sur les caractéristiques du mandarin mais nous serons amenée à prendre en compte à la fois la présence et l'influence des dialectes sur les néologismes dans le contexte de la communication en ligne.

## **1.4. Les caractéristiques du chinois contemporain**

### **1.4.1. Caractéristiques grammaticales comparées**

Notre objectif ici n'est pas de présenter la grammaire du chinois mais de repérer les différences les plus saillantes entre le chinois et le français. Une différence remarquable de la langue chinoise par rapport au français se trouve au plan grammatical. La langue chinoise ne possède pas un système morphosyntaxique équivalent au français. Nous allons rappeler ses principales caractéristiques.

#### ***Les noms***

Les noms chinois ne comportent pas de distinction de genre ni de nombre. Par exemple, pour distinguer une étudiante d'un étudiant, en chinois, il faut s'appuyer sur les adjectifs : 一名女学生 (yī míng **nǚ** xué shēng), une étudiante ; 一名男学生 (yī míng **nán** xué shēng), un étudiant. 一 est le nombre un qui reste invariable. Les deux adjectifs 女 et 男 signifient

respectivement féminin et masculin. Comme le nombre, le nom 学生 reste morphologiquement identique. Dans cet exemple, il existe également un élément spécifique, le classificateur 名. Le classificateur, appelé 量词 (liàng cí) en chinois, comporte principalement deux catégories : classificateurs nominaux et verbaux. Le premier est obligatoire pour dénombrer des objets et le deuxième pour le nombre d'occurrences d'une action. Nous prenons ici les classificateurs nominaux comme exemple. Comme dans l'expression française « un verre d'eau », le classificateur est le mot « verre » qui est utilisé plutôt pour indiquer le contenant. En chinois, chaque objet doit être associé à un classificateur donné. Nous avons énuméré quelques exemples dans le tableau suivant :

Chinois	Classificateurs concernés	Traduction Française
一本书 yī běn shū	本	Un livre
一支笔 yī zhī bǐ	支	Un stylo
一块钱 yī kuài qián	块	Un euro
一张桌子 yī zhāng zhuō zi	张	Une table

**Tableau 3 : Exemples de classificateurs nominaux en chinois**

Ainsi, quand la quantité change, les noms restent toujours invariables :

1 支笔	2 支笔	3 支笔	4 支笔
un stylo	deux stylos	trois stylos	quatre stylos

**Tableau 4 : Exemple montrant l'absence de différenciation de nombre d'un nom chinois**

L'absence d'indication du nombre laisse plusieurs interprétations possibles :

笔 =	Un stylo
	Le stylo
	Des stylos
	Les stylos

**Tableau 5 : Exemple montrant la polysémie d'un nom sans déterminants numériques et articles**

## Les verbes

Les verbes ne connaissent pas de conjugaisons. Les temps s'expriment par des adverbes comme 已经 yǐ jīng (déjà), 将来 jiāng lái (dans l'avenir) etc. Mais les aspects sont marqués quant à eux par des particules aspectuelles comme 了 le (accompli, changement d'état), 着 zhe (duratif, progressif), 过 guò (expérience). Par exemple :

Exemples chinois	Traduction française
我吃过饭了	J'ai mangé.
我正在吃饭	Je suis en train de manger.
我将去吃饭	Je vais manger.

Tableau 6 : Exemples portant sur le temps et l'aspect du chinois simplifié

Le verbe ici est « manger » (吃饭, chī fàn). Les trois phrases chinoises présentent respectivement une action dans le passé, le présent et le futur. Nous pouvons constater que ce verbe est invariable dans toutes les phrases. L'expression du temps et de l'aspect s'appuie sur l'ajout des adverbes et les particules.

## L'ordre des mots

L'ordre des mots dans la phrase influence la structure et le sens entier. Par exemple :

会议中断 (Sujet + prédicat)	中断会议 (Verbe + complément d'objet)
Une réunion interrompue	Interrompre une réunion

Tableau 7 : Exemple portant sur le changement de structure influencé par l'ordre des mots

Dans la structure, 会议 (huì yì) signifie réunion ; 中断 (zhōng duàn) correspond à « interrompue » et « interrompre ». Dans l'exemple de la colonne de gauche, 中断 placé après le sujet, joue le rôle de prédicat et signifie « interrompue » alors qu'il change de fonction dans l'exemple de la colonne de droite après en place initiale de structure. Il sert de verbe qui signifie

« interrompre »<sup>20</sup>.

Voici un autre exemple :

没有/都/去	都/没有/去
Ne pas/tout/aller	Tout/ne pas/aller
On n'y est pas tous allé.	Aucun n'y est allé.

**Tableau 8 : Autre exemple portant sur le changement de structure influencé par l'ordre des mots**

Dans notre recherche portant sur les néologismes, c'est surtout le niveau lexical qui est concerné même si le corpus de tweets que nous analysons met en jeu aussi tous les aspects de la syntaxe. Nous nous réservons de revenir sur les points pertinents au fil des analyses et allons maintenant aborder les systèmes d'écriture du chinois.

#### **1.4.2. Les systèmes d'écriture**

Le mandarin actuel est connu pour son double système graphique : écriture idéographique et système pinyin. Contrairement au pinyin qui est un système de transcription récemment créé par la romanisation de la langue, la nouvelle écriture idéographique simplifiée est le résultat d'une intervention humaine sur une graphie qui a connu une très longue évolution dans l'histoire.

Nous allons aborder brièvement le développement de l'écriture chinoise dans l'histoire. Ensuite, nous nous intéressons aux deux systèmes d'écriture d'aujourd'hui, le chinois simplifié et son système de transcription, pinyin.

#### ***L'écriture idéographique et le chinois simplifié***

L'une des grandes différences entre le français et le chinois, réside dans leurs unités d'écriture respectives. L'écriture chinoise est marquée par les signes idéographiques, autrement dit les

---

<sup>20</sup> La différence correspond à l'opposition de sens actif/passif en français.



caractères.

Ces caractères, appelés en chinois 汉字 (hàn zì), ont connu tout au long de l'évolution historique, des mesures de rationalisation pour des raisons politiques et socioculturelles. La première mesure peut remonter jusqu'à la Dynastie Qin (221-207 av. J.-C). C'est tout d'abord l'empereur Qin (秦始皇, qín shǐ huáng) qui cherche à mettre en place une écriture unifiée et standard, la graphie 小篆 (xiǎo zhuàn). On compte à l'époque environ 3000 caractères normalisés. Ensuite, à partir de cette graphie, se sont développées diverses écoles de styles d'écriture. Les autorités suivantes ont choisi de privilégier certains styles comme une norme adaptée à leurs besoins. Par exemple, le style 楷书 (kǎi shū) est en usage dominant dans la Dynastie Tang (618-907). Les caractères formés ainsi jusqu'au 20<sup>e</sup> siècle servent de référence pour le chinois simplifié sur le plan graphique. L'effectif de ces caractères évolue aussi dans le temps. Selon le « Dictionnaire de caractères de kāngxī »<sup>21</sup> (康熙字典, kāng xī zì diǎn), le nombre de caractères enregistrés au cours de la Dynastie Qīng (1644-1911) atteint 47 035.

Les réformes récentes que l'écriture chinoise a connues après 1949 ont abouti aux formes actuellement en usage. Dans le but de réaliser un apprentissage plus rapide et efficace des caractères, le gouvernement de la Chine a entrepris deux démarches principales : diminuer le nombre de caractères et simplifier leurs formes complexes. C'est ainsi qu'est né ce qu'on appelle le **chinois simplifié**.

Selon la dernière version établie en 2013, le 通用规范汉字表<sup>22</sup> (Tableaux des caractères normalisés et usuels) contient au total 8105 caractères classés en trois catégories : la 1<sup>ère</sup> (3500) et la 2<sup>ème</sup> (3000) catégories sont constituées des caractères visant à satisfaire les besoins de l'édition, de la compilation et du traitement des informations. La 3<sup>ème</sup> catégorie (1605) est composée des caractères usuels liés à la vie quotidienne.

Un caractère est formé à partir des « traits » (笔画, bǐ huà). On distingue généralement cinq

---

<sup>21</sup> L'un des plus célèbres dictionnaires chinois dont la compilation avait été ordonnée par l'empereur 康熙 kāng xī de la dynastie Qīng (1644-1911 ap. j.-C) en 1710. Il fut publié en 1716 et nommé en l'honneur de l'empereur. Source : <https://www.zdic.net/zd/kx/>.

<sup>22</sup> Source: [http://www.gov.cn/zwjk/2013-08/19/content\\_2469793.htm](http://www.gov.cn/zwjk/2013-08/19/content_2469793.htm).

traits fondamentaux :

le trait horizontal (héng)	一
le trait vertical (shù)	丨
le trait jeté descendant de droite à gauche (piě)	丿
le point (diǎn)	丶
le trait brisé (zhé)	→

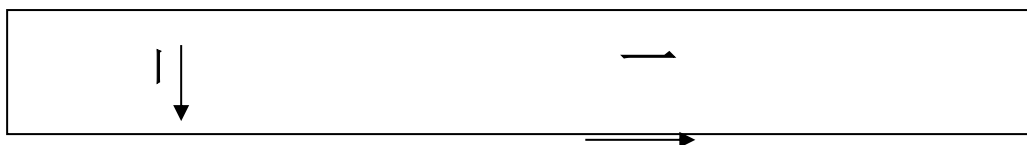
**Tableau 9 : Les traits fondamentaux de l'écriture idéographique du chinois simplifié**

Par la suite, ces traits se combinent pour former des radicaux 部首<sup>23</sup> (bù shǒu) qui sont l'unité de base des caractères chinois. Aujourd'hui, on compte au total 201 radicaux<sup>24</sup> selon le Ministère chinois de l'Éducation. Par exemple :

丿 + 丨 = 亻                      亻 + 尔 = 你

Les traits « 丿 » et « 丨 » se combinent en « 亻 » qui devient le radical du caractère « 你 ».

L'apprentissage de l'écriture chinoise consiste à mémoriser et pratiquer ces traits et radicaux. En plus, l'écriture chinoise impose des consignes. Il faut suivre absolument un certain ordre : les traits doivent être tracés de haut en bas, de gauche à droite...



À l'heure actuelle, l'écriture officielle de la Chine est constituée par des caractères simplifiés.

<sup>23</sup> Ce sont des éléments graphiques qui entrent dans la composition du caractère chinois. Il est à noter que le radical ne donne pas nécessairement d'indication sémantique.

<sup>24</sup> L'origine des radicaux remonte au lexicographe 许慎 (xú shěn) qui vécut sous la Dynastie des Hans de l'Est (25-220 ap. j.-C). Il a établi 540 radicaux dans son ouvrage 說文解字/说文解字 (shuō wén jiě zì). Cette liste de radicaux a été ramenée à 214 par 梅膺祚 (méi yíng zuò) dans 字汇 (zì huì) « Collection de caractères », en 1615. Ensuite, la liste a été entérinée par 康熙字典 (kāng xī zì diǎn) « Dictionnaire de caractères de kāngxī » en 1716. Ces 214 radicaux restent toujours la base du chinois traditionnel. Selon « Liste des radicaux des caractères chinois » publiée en 2009 par la Commission de la langue de l'État, le nombre des principaux radicaux chinois est de 201 : [http://www.moe.gov.cn/jyb\\_xwfb/xw\\_fbh/moe\\_2069/moe\\_2590/moe\\_2653/moe\\_2654/tnull\\_44668.html](http://www.moe.gov.cn/jyb_xwfb/xw_fbh/moe_2069/moe_2590/moe_2653/moe_2654/tnull_44668.html).

Néanmoins, l'écriture traditionnelle du chinois est toujours utilisée à Taiwan et Hongkong, à Singapour, etc. Une grande différence entre le chinois simplifié et le chinois traditionnel est qu'un caractère simplifié est composé de moins de traits que sa forme traditionnelle. Voici un exemple des deux graphies chinoises de l'expression hàn yǔ (langue des Hans) :

## 汉语 VS 漢語

À première vue, les différences formelles sont visibles. 汉语 appartient à l'écriture simplifiée et 漢語 à l'écriture traditionnelle. Le premier est composé par 14 traits au total (汉 5 + 语 9) alors que le deuxième en compte 28 (漢 14+語 14). Néanmoins, cela n'empêche pas quelqu'un qui pratique l'écriture simplifiée de reconnaître les caractères traditionnels puisque les idées représentées sont plus faciles à saisir dans la structure idéographique traditionnelle. La forme simplifiée masque souvent cette origine du caractère. Par exemple le caractère traditionnel 見 (jiàn, qui signifie voir), est composé par deux radicaux 目 et 儿. Le premier signifie *yeux* et le deuxième est associé à l'homme. L'action de voir est reliée aux yeux. Le radical 目 est placé au-dessus de 儿 pour indiquer l'emplacement des yeux dans le corps humain. Pourtant, cet aspect est moins explicite dans la forme simplifiée 见. Il perd une origine graphique de nature iconique en lien avec les yeux. Outre les différences formelles, les deux écritures n'adoptent pas identiquement certains signes de ponctuation. Par exemple, les signes de guillemets utilisés par le chinois simplifié est “” tandis que le chinois traditionnel utilise 「 」 .

Il convient également de mentionner ici la transcription des dialectes par les caractères<sup>25</sup>. Le mandarin a pris le parler de Pékin comme référence de prononciation comme nous l'avons dit *supra* et il est évident que le mandarin ne peut pas englober les sons de tous les dialectes. Donc, quand on veut transcrire les dialectes, on rencontre inévitablement un problème : pour certains, leurs structures phonétiques ne leur permettent pas d'association à un caractère donné. Pour résoudre ces cas-là, « 方言字 » (fāng yán zì) est apparu. Ce sont des caractères spécialement

---

<sup>25</sup> Le processus est le suivant : on transcrit d'abord les dialectes en pinyin par un rapprochement phonétique. Selon les résultats en pinyin, on cherche des caractères correspondants. Le son et le sens seront transposés en même temps dans les caractères finalement choisis.

réservés (souvent peu employés dans l'usage courant) ou fabriqués pour réaliser à la fois la transposition phonétique et sémantique des dialectes.

Si le mandarin favorise une intercompréhension sur le plan linguistique, le développement et les usages de l'Internet permettent de leur côté de briser la barrière géographique. L'emploi des dialectes n'est plus limité à un endroit précis et à des locuteurs donnés de ce fait. Comme les dialectes ont une importance non négligeable dans la communication quotidienne, cela nous amènera à observer si cette nouveauté apparaît, elle aussi dans la communication numérique et si la communication en ligne entre les locuteurs de régions différentes engendre de nouvelles caractéristiques parmi lesquelles les néologismes.

Nous allons maintenant aborder le deuxième volet des réformes qui ont transformé le système d'écriture en constituant une alternative aux caractères.

### *Le Hanyu pinyin (汉语拼音, hàn yǔ pīn yīn)*

Le 11 février 1958 la romanisation de l'écriture chinoise a été légalement approuvée. Dès lors ce système de transcription, appelé pinyin (拼音, pīn yīn) se basant sur les lettres latines s'est répandu à travers toute la Chine et ensuite a été reconnu à l'international. L'enseignement du pinyin est obligatoire dans le système éducatif chinois. Il est aussi devenu un outil aidant à apprendre le mandarin tant pour les Chinois natifs que pour les étrangers.

Le pinyin qui est une transcription phonétique des caractères chinois, n'a pourtant aucun lien avec la forme et la structure d'un caractère. Il nécessite un effort de mémoire exceptionnel pour associer un caractère avec son pinyin. Au sein du système, il y a au total **22** consonnes, **6** voyelles et **4** tons. Ces unités de prononciation peuvent constituer les unités de parole, les syllabes. Généralement, le caractère chinois est monosyllabique. Les syllabes basiques sont composées de consonnes et de voyelles selon certaines règles et elles sont à peu près au nombre de quatre cents dans le mandarin. Les règles de constitution sont simples :

- les initiales des syllabes sont entièrement consonantiques à part la consonne N qui peut être à la fois initiale et finale, par exemple : 南 (sud) est constitué comme suit : nán = n+an

- les modèles de constitution des finales d'une syllabe sont : voyelle + voyelle ou voyelle + consonne.

Autrement dit, de ce fait, les unités de prononciation du chinois sont composées de **21** initiales consonantiques (声母, shēng mǔ), **39** finales (韵母, yùn mǔ) et **4** tons.

Les initiales consonantiques	b p m f d t n l g k h j q x zh ch sh r z c s
Les finales simples	a o e i u ü
Les finales composées	ai ei ao ou ia ie iao iou (iu) ua uo uai uei (ui) üe
Les finales nasales	an en ang eng ong ian in iang ing iong uan uen (un) uang ueng üan ün
Les finale spéciques	er

**Tableau 10 : Tableau des unités graphiques de l'alphabet chinois (pinyin)**

Comme nous l'avons dit le mandarin est une langue à tons, ce qui est d'ailleurs le cas de beaucoup de langues dans le monde. Un ton se combine avec une syllabe en mandarin. C'est cet ensemble qui est associé ensuite à un caractère donné. Les tons servent de ce fait à distinguer les syllabes entre elles. Le changement d'un ton dans une syllabe conduit à l'association d'un caractère différent. Les 4 tons placés au-dessus de la voyelle servent à la distinction du sens des caractères. Grâce à ces 4 tons, les 400 syllabes basiques (environ) mentionnées *supra* peuvent former à peu près 1300 syllabes différentes qui sont associées chacune à un sens donné. Nous présentons dans le tableau suivant les éléments graphiques qui indiquent les quatre tons et donnons un exemple avec une même syllabe qui aboutit à quatre mots différents :

Le premier ton	-	Ex. : mā 妈 (mère)
----------------	---	-------------------

Le deuxième ton	/	<b>Ex. : má 麻 (chanvre)</b>
Le troisième ton	ˇ	<b>Ex. : mǎ 马 (cheval)</b>
Le quatrième ton	\	<b>Ex. : mà 骂 (insulter)</b>
NB : les lettres latines dans pinyin sont toujours en minuscule.		

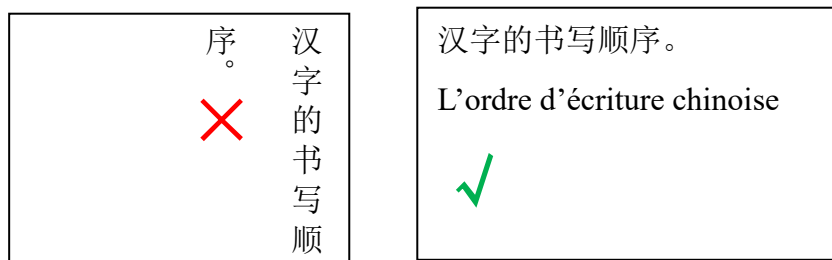
**Tableau 11 : Les quatre tons chinois avec la syllabe « ma »**

La syllabe « ma » associée à chacun des quatre tons de façon à illustrer les mots de sens différents, est un des exemples typiques et classiques utilisés pour décrire la fonction distinctive des tons en indiquant leurs graphies en pinyin. Il existe aussi un ton spécial, le ton neutre. Ce ton n'a pas de marque. Par exemple avec la syllabe « ma » cela donne : 吗 « ma » qui est une particule placée à la fin d'une phrase affirmative pour la transformer en une question. Les syllabes avec le ton neutre se prononcent plus légèrement. À partir du tableau 11 ci-dessus, on peut aussi constater que les tons se positionnent sur la voyelle. Dans le cas des finales composées, il se note sur la voyelle médiane, par exemple : Gāo.

L'utilisation du pinyin intervient plutôt dans l'apprentissage précoce du chinois et l'apprentissage de l'écriture simplifiée par la suite. Il faut aussi souligner que la création du système pinyin et sa vulgarisation depuis 1955 ont une histoire relativement courte. Par conséquent, il ne peut pas y avoir une influence sur toute la population. Aujourd'hui, ce sont les générations des personnes nées en 1970, 1980, 1990... qui sont à l'aise avec ce système. En outre, dans la vie quotidienne et à l'oral, ce sont toujours les dialectes qui occupent une place plus importante. Il faut aussi distinguer production et compréhension. Il se peut qu'un jeune Chinois maîtrise le mandarin et son propre dialecte ou seulement le mandarin. Il est aussi possible qu'une personne âgée de 40 ans ou plus ne sache parler que le dialecte mais comprenne bien le mandarin. De ce fait, on peut apercevoir que la normalisation de la langue chinoise sur le plan phonétique favorise l'intercommunication au sein de toute la Chine.

Le contraste formel et grammatical par rapport au français est fort. En effet les caractères chinois ne comportent aucune flexion. Ils n'ont pas de désinence verbale, de variation formelle. Même si le pinyin est fondé sur les lettres latines, il n'y a pas de différence entre majuscule et minuscule et l'écriture en pinyin est toujours en minuscules.

Certaines personnes croient que les Chinois écrivent de haut en bas, de droite à gauche (vignette à gauche de la Figure 4 ci-dessous<sup>26</sup>), mais ce n'est plus le cas du chinois simplifié. L'ordre de réalisation et la disposition de l'écriture chinoise sont aujourd'hui identiques à ceux du français (vignette à droite de la Figure 4) :



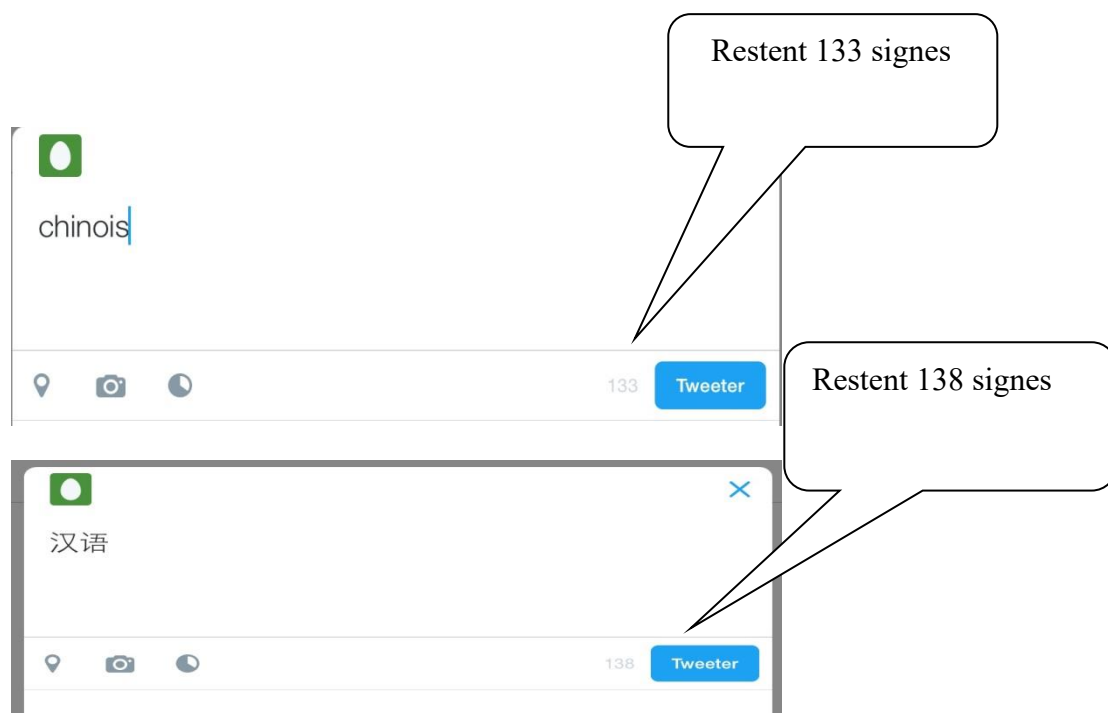
**Figure 4 : Disposition d'une des écritures japonaises (à gauche) et de l'écriture chinoise (à droite)**

À l'intérieur d'une phrase ou d'une expression, les caractères chinois se combinent sans espace. De ce fait, si une phrase française et une phrase chinoise comptent le même nombre des mots, la phrase chinoise sera probablement plus courte que la phrase française.

Le point le plus important pour notre recherche est que par exemple sur *Twitter*, un caractère chinois égale un caractère informatique alors que pour le français, c'est une lettre qui occupe un espace informatique. Dans un message limité à 140 signes, cette différence influence particulièrement le contenu et les moyens d'expression.

---

<sup>26</sup> Traditionnellement le japonais s'écrit de haut en bas et de droite à gauche, mais il est de plus en plus courant qu'il s'écrive et se lise de gauche à droite et de haut en bas. Ces deux dispositions sont en usage actuellement au Japon.



**Figure 5 : Interfaces de *Twitter* montrant les différences entre le comptage des signes des caractères chinois et d'un mot français**

### 1.5. La saisie informatique du chinois

Voyons à présent de plus près les questions relatives à la saisie informatique du chinois.

L'Internet et les technologies numériques bouleversent la vie quotidienne dans tous ses aspects. En changeant le mode de communication, ils ont modifié consécutivement le mode d'écriture. Du pinceau au stylo puis au clavier, de l'organisation verticale à l'organisation horizontale, la langue chinoise a ainsi rejoint les autres langues dans l'univers informatique. Réel ou représenté sur un écran, le clavier est l'outil d'écriture dominant à l'ère numérique. Comment tape-t-on les caractères chinois en ligne ? Les Français se posent probablement souvent cette question tant qu'ils n'ont pas l'occasion de l'observer directement. Certains pensent que les claviers chinois comportent des images de caractères sur chaque touche. Ils n'ont pas tout à fait tort... La saisie des caractères se fait à partir d'un clavier standard anglais (de type Qwerty) à l'aide de logiciels et il existe deux modes de saisie principaux.

Le mode de saisie Wubi (五笔字型部分输入法), ce qui signifie littéralement « 5 traits », est



un codage fondé sur les formes graphiques du chinois. Il a été créé en 1983. Avec ce codage, chaque touche du clavier est associée à certains traits ou traits composés. Donc il existe bien un clavier chinois mais il est virtuel et obtenu via un logiciel et non par un dispositif mécanique :



Figure 6 : Disposition des touches du mode de saisie Wubi

Pour maîtriser ce mode Wubi, les utilisateurs doivent mémoriser les règles de saisie et les pratiquer. Cela demande nécessairement une formation préalable. En plus, les utilisateurs de ce mode sont censés bien connaître l'écriture chinoise. Comme le pinyin se généralisait à peine dans les années 80, les usagers acceptaient cette méthode de saisie. Mais au fil du temps, la formation préalable et la nécessité de réciter les règles sont devenus de plus en plus des obstacles pour les choix.

La deuxième méthode de saisie, pinyin, est basée sur la transcription alphabétique. Si les utilisateurs ont bien appris le système pinyin à l'école, aucun apprentissage spécifique n'est nécessaire avec cette méthode. Cela permet une utilisation immédiate et simple. Aujourd'hui, les jeunes générations maîtrisant bien le pinyin occupent une proportion importante parmi les utilisateurs de l'Internet et des réseaux sociaux numériques. La saisie par pinyin est donc devenue la solution préférée des internautes.

Voici le mode d'emploi de cette méthode. On tape les pinyin correspondant à chaque caractère ou simplement les lettres initiales des caractères. Ensuite, le logiciel propose à l'utilisateur une liste lexicale au sein de laquelle choisir. Dans l'exemple ci-dessous (Figure 7), pour écrire « chinois », on n'a qu'à saisir son pinyin « hanyu » ou « h+y ».



**Figure 7 : Mode de saisie pinyin logicielle assistée de caractères chinois simplifiés**

Ce système a tellement évolué que le logiciel peut mémoriser la fréquence d'un caractère saisi et proposer une liste filtrée en reconnaissant les habitudes d'un utilisateur donné. Cette liste proposée vient de la base de données des vocabulaires du logiciel. Ainsi, le logiciel effectue une mise à jour instantanée qui vise à recueillir les mots ou expressions à la mode et donc les néologismes. Les utilisateurs peuvent profiter immédiatement de ces nouveautés dans la communication. De ce point de vue, les logiciels de saisie informatique jouent aussi un rôle positif dans la circulation des nouveautés langagières sur l'Internet.

## **Conclusion**

Au terme de notre rapide présentation de la langue chinoise, nous retenons qu'il nous faudra prendre en compte les caractéristiques du chinois contemporain, le mandarin, sans oublier les dialectes régionaux. Il est essentiel, dans la perspective de l'étude d'un réseau social numérique qui recourt à l'écrit et à l'image, de considérer les deux systèmes d'écriture relativement indépendants issus de la normalisation récente voulue par les instances politiques du pays, et qui ont à présent très bien pénétré dans les usages de la population : les caractères simplifiés et l'écriture phono-grammatique ou alphabétique en caractères latins, le pinyin. Les principales spécificités linguistiques du chinois et les grandes différences entre le français et le chinois auront évidemment une incidence sur l'analyse des nouveautés langagières que nous voulons mener et il nous faudra les mobiliser pour comprendre et éclairer les phénomènes qui se présenteront à notre observation. Enfin s'agissant de communication en ligne le mode de saisie informatique du chinois ainsi que les différences entre mode de saisie pour le français et pour le chinois ne sont pas indifférents. Le logiciel de codage créé pour le chinois et qui mémorise les mots nouveaux en est un exemple.

Après cette approche linguistique centrée sur le chinois qui vise à réduire la distance à cette langue et à faciliter l'intelligibilité de nos analyse ultérieures, nous allons dans le chapitre suivant aborder les aspects communicationnels et sociotechniques des réseaux.

## CHAPITRE II

### Contexte numérique et nouveaux dispositifs médiatiques

#### Introduction

L'Internet facilite énormément l'accès aux informations de toutes sortes et leurs partages. Ce faisant, un nouveau contexte de communication interpersonnelle s'élabore et de nouvelles habitudes de lecture et d'écriture s'imposent. Cette évolution technique et sociale majeure fait aussi apparaître de nouveaux termes qui renvoient à ces mutations et parmi lesquels il peut être difficile de se repérer. Les expressions comme Web 2.0, réseaux sociaux, médias sociaux, sites, plates-formes sont appliquées pour désigner des outils de communication avec lesquels nous sommes très familiers actuellement (*Facebook, YouTube, Twitter, Instagram, Wikipédia, Tiktok, Snapchat* etc.). Pour autant les distinctions entre ces termes largement diffusés et les réalités auxquelles ils réfèrent précisément ne sont pas toujours claires. La connaissance du contexte très diversifié de l'Internet semble à la fois évidente et brouillée. Des efforts venant de domaines variés (champs scientifiques tels que l'informatique, les sciences sociales, les sciences de l'information communication aussi bien que marketing...) sont déployés pour les définir. Dans ce chapitre, nous allons tenter de clarifier les notions et leurs désignations afin de mieux cerner le contexte communicationnel numérique. Premièrement, nous allons nous intéresser à la relation entre l'Internet et le Web et préciser ce qu'on entend par Web 2.0. Ensuite, nous aborderons les différences entre les réseaux sociaux, les réseaux sociaux numériques et les médias sociaux avant de nous questionner sur la catégorie à laquelle appartient le microblog, ce dernier dispositif étant au centre de notre étude. Nous aborderons en même temps une présentation des microblogs choisis (*Sina Weibo* et *Twitter*) pour notre observation ainsi qu'une revue de littérature autour de ces microblogs. Enfin, nous nous appuyerons sur les notions de « culture participative » et d'« observation participante » pour réfléchir successivement aux caractéristiques du contexte numérique et aux méthodologies d'analyse d'un tel contexte dans cette étude.

## 2.1. L'Internet et le Web<sup>27</sup>

### 2.1.1. Aperçu historique

L'omniprésence de l'Internet dans notre quotidien et la multitude des services du Web causent souvent des confusions et les deux termes sont utilisés à tort comme des synonymes. Nous allons nous intéresser dans un premier temps à une brève histoire de l'Internet et nous chercherons à comprendre sa nature et son fonctionnement. Dans un deuxième temps, nous allons nous pencher sur les caractéristiques du Web et ses différentes étapes de développement (Web 1.0, 2.0...). Nous essaierons de discuter les relations entre l'Internet et le Web 2.0 à l'issue de ce sous-chapitre.

Le mot « Internet » est la contraction de l'expression anglaise « *interconnected network* », qui signifie en français : réseaux interconnectés. La création de l'Internet dans les années 1970 est premièrement destinée à des objectifs militaires aux Etats-Unis<sup>28</sup>. L'Internet a ensuite connu une longue phase d'évolution technologique qui lui a permis l'expansion dont nous sommes à présent témoins. Cette phase de développement se manifeste principalement par un changement de méthodes d'accès aux informations, au moyen de nouveaux protocoles. Nous pouvons rappeler quelques événements phares qui dessinent le paysage de l'Internet. En 1969, l'utilisation de protocoles de transmission de codes par paquets permet de relier les ordinateurs à un réseau décentralisé, puis de réaliser l'interconnexion des ordinateurs, ce qui commence à former la maquette de l'Internet (le fameux APPAnet<sup>29</sup>). Dans les années 1970, la création du protocole TCP/IP<sup>30</sup> rend la communication par paquets – c'est-à-dire des ensembles de données acheminées en bloc – applicable à tous les systèmes de communications. Au 1<sup>er</sup> janvier

---

<sup>27</sup> Pour cette présentation générale nous avons eu recours à des références scientifiques (Bubaš 2001 ; Cardon 2019 ; Casilli 2011 ; Casilli 2012 ; Constantinides et Fountain 2008 ; Gervais 2007 ; Niel et Roux 2010 ; Vitali Rosati 2011 ; Vitali-Rosati 2014) et des ressources encyclopédiques :

<https://www.leconomiste.eu/decryptage-economie/551-retour-sur-20-ans-de-numerique-et-d-internet.html>;

<https://www.ladocumentationfrancaise.fr/dossiers/internet-monde/chronologie.shtml>;

<https://www.futura-sciences.com/tech/actualites/internet-internet-t-il-35-ans-aujourd'hui-4284/>;

<https://recherchemid.wordpress.com/2016/11/29/quest-ce-que-le-Web-2-0-histoire-caracteristiques-et-perspectives/>.

<sup>28</sup> Dans le cadre de la compétition sur le plan technologique durant la Guerre Froide entre les Etats-Unis et l'URSS.

<sup>29</sup> L'acronyme anglais d'Advanced Research Projects Agency Network.

<sup>30</sup> Il réunit les deux protocoles TCP (de l'anglais *Transmission Control Protocol*) et IP (de l'anglais *Internet Protocol*). Il s'agit un standard de communication entre deux processus pour faciliter le transfert de données sur l'Internet.

1983, TCP/IP a été appliqué comme le protocole unique sur APPAnet et on considère que c'est la date de naissance officielle de l'Internet. En outre, ce sont les deux inventeurs américains de ce protocole, Robert E. Kahn et Vinton Gray Cerf, qui utilisent en premier l'expression « l'Internet ». Jusqu'à aujourd'hui, ce protocole est toujours utilisé comme la base de toutes les communications de l'Internet. En parallèle, l'ouverture de l'accès au public des protocoles a été effectuée mais son utilisation restait principalement dans le domaine scientifique et universitaire en raison des restrictions techniques à l'époque. La recherche des informations sur le réseau était encore compliquée et peu efficace. Il faut attendre 1989, la naissance du *World Wide Web* (communément appelé le Web)<sup>31</sup> basé sur le protocole HTTP<sup>32</sup> pour que l'Internet s'étende au grand public en facilitant l'accès aux informations disponibles sur le réseau.

Actuellement, l'Internet permet à ses utilisateurs de se connecter à travers des supports numériques variés. Selon *Le Rapport Digital annuel* (2018), publié par la plate-forme de gestion des médias sociaux Hootsuite et l'agence digitale We Are Social<sup>33</sup>, le nombre d'internautes dans le monde a franchi 4 milliards (4.176 milliards), soit 55% de la population mondiale. L'Internet est considéré maintenant comme le réseau informatique ouvert au public dans le monde entier.

L'expression « l'Internet » est d'origine américaine. En France, selon l'Académie française, l'Internet est un nom commun défini comme suit :

Le réseau mondial associant les ressources de télécommunication et des ordinateurs serveurs et clients, destiné à l'échange de messages électroniques, d'informations multimédias et de fichiers. Il fonctionne en utilisant un protocole commun qui permet l'acheminement de proche en proche de messages découpés en paquets indépendants<sup>34</sup>.

Cette définition nous permet de comprendre le mécanisme de fonctionnement de l'Internet. Sans entrer dans le détail des technologies de support, nous pouvons saisir que ce réseau virtuel

---

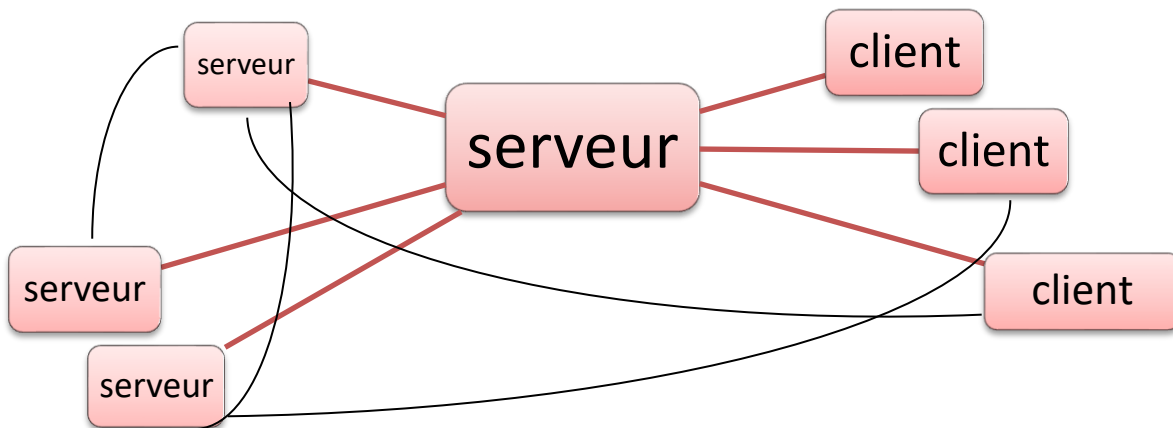
<sup>31</sup> Littéralement traduit en français comme « toile mondiale », il est inventé en 1989-1990 par Tim Berners-Lee suivi de Robert Cailliau.

<sup>32</sup> Sigle anglais de *HyperText Transfer Protocol*. Il s'agit d'un protocole de transmission permettant à l'utilisateur d'accéder à des pages Web par l'intermédiaire d'un navigateur.

<sup>33</sup> Disponible sur : <https://wearesocial.com/fr/blog/2018/10/digital-social-et-mobile-etat-des-lieux-sur-le-dernier-trimestre-2018> (consulté le 29 novembre 2019).

<sup>34</sup> Voir <https://www.dictionnaire-academie.fr/article/A9I1736>. Le nom commun implique de ne pas mettre l'initiale d'un mot en majuscule. Toutefois nous employons l'orthographe « l'Internet » en nous faisant référant à la source suivante : [https://www.lemonde.fr/big-browser/article/2016/04/04/faut-il-mettre-une-majuscule-a-internet\\_5991956\\_4832693.html](https://www.lemonde.fr/big-browser/article/2016/04/04/faut-il-mettre-une-majuscule-a-internet_5991956_4832693.html).

dispose d'une capacité de regrouper des ressources informatiques diverses. Au sein du réseau, la transmission de ces ressources se réalise par la communication entre les dispositifs de télécommunication, matériels ou logiciels, qui peut être divisés en deux catégories : les serveurs et les clients. Depuis un client, notamment les dispositifs informatiques employés par les internautes, tels que ordinateur personnel, tablette, téléphone, etc., nous envoyons des requêtes de recherche et nous recevons les réponses des serveurs – ces serveurs fournissent eux-mêmes les ressources et donnent aussi l'accès direct aux données provenant des autres serveurs –, et nous obtenons des résultats. Par exemple, du côté client, un utilisateur va rechercher un site Internet depuis son navigateur. Cette requête sera envoyée au serveur Web et on aura les résultats affichés suite aux réponses du serveur. Cette liaison du dispositif et de l'information constitue le réseau de l'Internet. Nous pouvons illustrer ce mécanisme à l'aide de la Figure 8:



**Figure 8 : L'architecture de l'Internet**

Pour nous résumer, l'Internet nous offre une infrastructure de communication virtuelle.

Le développement de l'Internet après les années 90 est étroitement lié à l'application du Web. Le service Web basé sur la technologie hypertexte<sup>35</sup> permet de lier entre eux des documents répartis sur l'Internet et de fournir en même temps une interface de navigation directe et pratique aux utilisateurs. Ainsi, les ressources multimédias aussi bien que les textes sont disponibles en ligne. Tout cela se réalise par le biais des navigateurs<sup>36</sup>. Ce qu'on peut obtenir à

---

<sup>35</sup> Fonction permettant d'établir des liaisons directes par des hyperliens entre éléments (texte, image...) de documents différents (Dictionnaire *Le Petit Robert* en ligne : <https://www.lerobert.com/>).

<sup>36</sup> En informatique, il s'agit d'un logiciel conçu pour consulter et afficher le Web.

travers les navigateurs, ce sont les ressources fournies par le Web. La mise en service du Web dépend de l'Internet et il est souvent considéré comme une application concrète de ce réseau informatique mondial. Comme nous l'avons indiqué ci-dessus, on doit aussi la propagation de l'Internet à l'application du Web. Mais même si l'Internet et le Web sont indissociables, ils sont différents tant sur le plan de la technologie fondamentale que sur celui de la fonction principale. Le Web constitue plutôt un intermédiaire entre l'Internet et ses utilisateurs tout comme les autres services (tels que l'email, la messagerie instantanée, etc.).

Jusqu'à aujourd'hui, le Web a connu 30 ans de développement. Ses différentes étapes d'évolution sont labellisées par les chiffres : 1.0, 2.0, 3.0, 4.0... pour indiquer l'ordre chronologique. Dans cette recherche, nous allons nous concentrer sur les différences entre Web 1.0 et Web 2.0 puisque les notions de 3.0 et 4.0<sup>37</sup> ne sont pas encore vraiment définies. Ces tournures sont employées pour décrire les futurs changements possibles du Web et non les réalités déjà opérationnelles. Nous considérons que le Web est toujours dans sa deuxième étape d'évolution, soit le Web 2.0.

### **2.1.2. Qu'est-ce le Web 2.0 ?**

Pendant une longue période, la première phase du Web, autrement appelée le Web 1.0, domine l'utilisation d'Internet. Ce terme est utilisé pour se différencier du Web 2.0. Durant cette étape 1.0, le Web avait pour rôle de diffuser des informations. Les utilisateurs ne pouvaient que lire ces informations. Il n'existait aucune possibilité d'y réagir. De ce fait, il est aussi appelé Web statique (Boisseau 2006). Le terme Web 2.0 est initialement utilisé en 2003 par Dale Douherty, cofondateur de la société O'Reilly Media lors d'une session de conférence entre O'Reilly et MediaLive International pour décrire une nouvelle ère d'Internet. Ensuite, l'entrepreneur dans le domaine de l'informatique, Tim O'Reilly rédige un long billet de blog<sup>38</sup> en 2005 pour clarifier la notion de Web 2.0 et ses caractéristiques. Ce billet a été considéré comme une

---

<sup>37</sup> L'inventeur du Web Tim Berners-Lee a décrit le Web sémantique comme une composante du Web 3.0 dans lequel « les ordinateurs sont capables d'analyser toutes les données du Web » (notre traduction) : <https://www.nytimes.com/2006/05/23/technology/23iht-Web.html>. Web 4.0 est considéré comme Web symbolique ou Web intelligent dont une définition est manquante.

<sup>38</sup> L'article fondateur de Tim O'Reilly : « *Qu'est-ce que le Web 2.0 : Modèles de conception et d'affaires pour la prochaine génération de logiciel* » : <https://www.internetactu.net/2006/04/21/quest-ce-que-le-Web-20-modeles-de-conception-et-daffaires-pour-la-prochaine-generation-de-logiciels/>.



référence pour définir le Web 2.0.

En nous appuyant sur ses interprétations, nous pouvons résumer ainsi : le Web 2.0 engendre une plateforme collaborative et participative au lieu d'un Web statique et linéaire. Par rapport à l'étape précédente, la façon d'exploiter le Web a été renouvelée. L'ère du CGU – le contenu généré par les utilisateurs – est advenue. Avec les nouveaux outils de communication (par exemple : *Wikipédia, Twitter, Facebook, etc.*), les internautes peuvent participer à l'activité de production et de partage d'informations en ligne. Cette communication interactive et sociale se distingue de celle du Web 1.0. Comme Girard et Fallery (2009) l'indiquent, les changements du Web 2.0 se reflètent dans deux dimensions :

- *une dimension technologique* : l'accès aux services devient facile ; les services sont accessibles aux différents supports, du smartphone à la tablette et maintenant les accessoires intelligents comme bracelet ou montre intelligents ; cela offre des interfaces intuitives et conviviales aux utilisateurs ;
- *une dimension sociale* : le Web 2.0 repose sur une culture participative (voir *infra*) ; la participation des utilisateurs est au centre de la production et de l'échange du contenu en ligne.

Cependant le débat autour de la signification de ce terme se poursuit. Il est aussi interprété sous les angles marketing, technologique et de design (Filip *et al.* 2008) : le contenu d'un site du Web 2.0 sera généré par les utilisateurs, dans ce cas-là, il y aura moins d'investissements sur sa mise en place et sa maintenance ; pour les techniciens, les sites disposent de nombreux points communs techniques ; côté design, les sites sont faciles à identifier par leurs caractéristiques graphiques.

La réalité de l'existence du Web 2.0. est aussi questionnée par des chercheurs qui critiquent son aspect marketing. Le Deuff a analysé ce terme sur le plan historique, philosophique et intellectuel dans son article disponible en ligne (Le Deuff 2007). Selon lui, le Web 2.0 est plutôt une idée recyclée qu'une notion innovante. Il n'incarne aucune révolution technologique mais s'inscrit dans un processus successif, celui d'une évolution lente de l'Internet. Cependant, Le Deuff montre que le Web 2.0 possède un objectif simple : « mettre l'utilisateur au centre et notamment au centre de l'information ». Quelles que soient les controverses liées à ce terme,

nous avons constaté que l'influence communicationnelle du Web 2.0 n'est pas négligeable. C'est cet aspect qui intéresse notre étude.

Il convient donc de concevoir l'Internet comme infrastructure et le Web comme un des services qui s'appuient sur ce grand réseau. Quant au Web 2.0, il peut être considéré comme une étape d'évolution du Web. Techniquement, l'Internet repose sur une communication entre les machines. La création du Web a mis au premier plan la communication homme-machine. À l'époque du 2.0, l'interaction entre les internautes basée sur la communication homme-machine prédomine. Autrement dit, l'Internet permet avant tout de réaliser une connexion mondiale alors que le Web, ensuite, favorise l'accès aux informations pour les usagers. Le Web 2.0 contribue, lui, à faciliter le processus de l'échange d'information (sous forme de plate-forme) et l'interactivité<sup>39</sup>.

## **2.2. Réseaux sociaux, RSN, médias sociaux, quelles sont les différences ?**

L'avènement de l'ère du Web 2.0 fait émerger des dispositifs de communication innovants et de ce fait de nouvelles pratiques se développent sur le Web avec *Facebook*, *Twitter*, *YouTube*, *Instagram*, etc. L'apparition de ces dispositifs s'accompagne de nouvelles catégories de classement comme « réseaux sociaux » et « médias sociaux ». Nous remarquons que ces deux expressions sont parfois employées sans distinction ; en effet la frontière entre ces catégories est floue en raison des similarités technologiques, des modes d'utilisation et des pratiques interactives des usagers. D'un autre côté, dans certains textes encyclopédiques en ligne qui définissent les dispositifs, les notions de « réseaux sociaux » et de « réseaux sociaux numériques » (RNS) ne sont pas distinguées.

Comme dans cette recherche, nous voulons observer la circulation des nouveautés langagières numériques sur deux sortes d'outils de microblogging (*Twitter* en France et *Sina Weibo* en Chine), nous nous posons la question de la catégorie d'appartenance de ces outils et nous nous proposons de clarifier ces catégories en deux temps : d'abord les réseaux sociaux par rapport aux RSN, ensuite les médias sociaux par rapport aux RSN. Au-delà d'une question de définition,

---

<sup>39</sup> Cela concerne à la fois l'aspect technique et l'aspect pratique. Nous nous intéressons plutôt à ce dernier aspect qui favorise les interactions entre les utilisateurs qui deviennent des participants actifs du réseau. Web 2.0 fournit des services centrés sur les usagers.

ce sera pour nous l'occasion de caractériser le contexte de communication et l'interactivité numérique.

### 2.2.1. Réseaux sociaux et réseaux sociaux numériques (RSN)

*Réseau* est un mot ancien qui s'est enrichi d'une évolution de sens depuis son apparition. Selon l'étude des origines du mot menée par le sociologue Mercklé (2004), son premier usage dans la langue française peut remonter jusqu'au 17<sup>e</sup> siècle pour désigner la matière tissu. Une extension de son sens explique ses occurrences dans le domaine médical pour nommer l'appareil sanguin aussi bien que le système nerveux, et dans le domaine géographique et des transports, pour désigner l'ensemble des chemins et des routes. À part ces désignations concrètes, des significations métaphoriques apparaissent qui mettent en relief l'entrelacement, la circulation et la représentation topologique. Comme le dit Mercklé « la notion de réseau [...] s'est progressivement abstraite des objets concrets qu'elle servait primitivement à désigner » (2004:2). Il a aussi souligné que « ces registres métaphoriques spécifiques » dans la langue courante peuvent être attestés à partir du 19<sup>e</sup> siècle.

Nous retrouvons ainsi cette définition de *réseau* dans le dictionnaire *Le Petit Robert* (2012) : « **Ensemble d'éléments reliés entre eux**.1.(1828) [...] réseau social : communauté d'individus liés par des centres d'intérêt, des goûts, des besoins communs. Les réseaux sociaux d'Internet. - **Site Web communautaire** ».

Nous pouvons constater que la notion de « réseau social » est employée pour désigner les relations sociales dans une entité donnée, c'est-à-dire que les liens interpersonnels sont au centre de cette notion. Effectivement, dans le domaine des sciences sociales, le terme « réseau social » (*social network* en anglais) a été utilisé pour la première fois de manière scientifique par l'anthropologue John A. Barnes (1954)<sup>40</sup> pour désigner l'ensemble des relations entre personnes ou au sein d'une communauté d'individus. Par la suite, nous trouvons les définitions

---

<sup>40</sup> « In parenthesis we may note that one of the principal formal differences between simple, primitive, rural or small-scale societies as against modern, civilized, urban or mass societies is that in the former the mesh of the social network is small, in the latter it is large. » (Barnes 1954:44). Nous le traduisons comme suit : « Entre parenthèses, nous pouvons noter que l'une des principales différences formelles entre les sociétés simples, primitives, rurales ou à petite échelle et les sociétés modernes, civilisées, urbaines ou de masse est que dans les premières, la maille du réseau social est petite, dans les secondes, elle est grande ».

de deux sociologues qui contribuent au développement de l'analyse des réseaux sociaux : Degenne (2011) considère que « les réseaux sociaux sont des formes sociales composées de relations ou d'interactions entre des personnes » et Forsé (2008) qu' « un réseau social est un ensemble de relations entre un ensemble d'acteurs ».

### 2.2.2. Les caractéristiques des RSN

Aujourd'hui, les usages de *réseau* évoluent notamment sous l'influence des nouvelles technologies. Le mot *réseau* existe bien avant l'Internet mais ce dernier met en place une nouvelle vision de la notion de réseau, le réseau informatique. En même temps, la notion de réseau social s'applique également dans le domaine de l'Internet pour se référer aux nouveaux outils de communication tels que *Facebook* ou *Twitter*, etc. A ce titre, la sociologue Boyd et la chercheuse en sciences de la communication Ellison ont proposé en 2007 la notion de *social network site*, appelée aussi SNS et traduite en français par « réseaux sociaux numériques » ou RSN. Par la suite, lors d'un entretien scientifique avec Thomas Stenger et Alexandre Coutant en 2011, Ellison considère que la définition qu'elle avait proposé avec Boyd en 2007 doit être renouvelée en raison « des changements techniques ou sociaux dans le paysage des réseaux socionumériques » (Ellison 2011:22). La définition donnée en 2011 est alors la suivante :

Un site de réseau social est *une plate-forme de communication en réseau* dans laquelle les participants 1) disposent de *profils associés à une identification unique* qui sont créés par une combinaison de contenus fournis par l'utilisateur, de contenus fournis par des « amis », et de données système ; 2) peuvent *exposer publiquement des relations* susceptibles d'être visualisées et consultées par d'autres ; 3) peuvent accéder à *des flux de contenus incluant des contenus générés par l'utilisateur* – notamment des combinaisons de textes, photos, vidéo, mises à jour de lieux et/ou liens – fournis par leurs contacts sur le site. (Ellison 2011:22) (version française traduite par Annike Thierry).

En réalité, la première définition du RSN (2007) est considérée comme fondamentale et ainsi beaucoup citée<sup>41</sup> par les chercheurs. Cependant, nous avons aussi constaté qu'il existe des critiques portant sur la confusion et le manque de précision de cette définition initiale.

Nous allons nous intéresser dans un premier temps à l'évolution de la définition élaborée par

---

<sup>41</sup> Selon google scholar (données de 04/04/2019), le nombre de citation est de 17117 : <https://scholar.google.com/citations?user=P0XTYBEAAAAJ&hl=zh-CN>.

les deux chercheuses. Ensuite, nous allons faire état des critiques autour de la définition initiale. En même temps, nous ferons une analyse de la deuxième définition par rapport à ces critiques. Nous proposerons enfin nos propres réflexions sur la définition des RSN.

Nous reprenons ici intégralement la définition initiale de Boyd et Ellison (2007) que nous traduisons de l'anglais comme suit :

Un service Web permettant aux individus de 1) construire un profil public ou partiellement public au sein d'un système donné ; 2) établir une liste des autres utilisateurs avec lesquels ils partagent une connexion ; 3) consulter et parcourir leurs listes de contacts et celles effectuées par d'autres personnes au sein du système. La nature et la nomenclature de ces contacts peuvent varier d'un site à l'autre<sup>42</sup>.

Cette définition permet de dégager trois caractéristiques propres aux réseaux sociaux numériques :

- un profil créé par les usagers ;
- une liste des relations (liste d'amis ou liste des abonnés ou d'abonnements) ;
- des fonctionnalités fournies par ces dispositifs permettant d'exploiter ces relations.

La condition préalable pour utiliser les services de RSN est d'avoir une identité digitale, c'est-à-dire que les usagers doivent s'inscrire sur tel service en disposant d'un compte unique. Cette identité se présente sous des formes variées selon les demandes concrètes des RSN : un pseudonyme, une adresse email ou un numéro de portable<sup>43</sup>. Il faut remarquer que le fait d'avoir un compte n'est pas toujours nécessaire pour certains RSN. Bien évidemment, sans disposer de cette identité virtuelle, on ne peut pas bénéficier de la totalité des services des RSN. Les

---

<sup>42</sup> « We define social network sites as Web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from site to site. ».

<sup>43</sup> En Chine, à partir de 2012, le système de noms réels (lors de la création d'un compte en ligne, les usagers sont tenus de fournir leurs informations d'identification authentiques, par exemple leur n° de téléphone) a commencé à s'appliquer, surtout dans le domaine téléphonique. La même année, les quatre compagnies de microblogging chinois se sont mises d'accord pour appliquer ce système. Donc, les usagers peuvent s'inscrire en entrant leurs numéros de portable. Ce numéro unique permet de s'associer à une identité réelle. Il faut signaler que si un numéro est utilisé comme le nom d'un compte, il ne sera pas affiché complètement afin de garantir les informations privées des usagers. Le profil associé à un numéro de portable est une option.

plateformes de contenus vidéos comme *Youtube*, *Youku* ou *Sohu*<sup>44</sup> sont représentatives de ce phénomène. La visualisation des vidéos ne demande qu'un simple clic mais on ne peut pas interagir avec les autres usagers (ex. : laisser des commentaires ou s'abonner à une chaîne) sans posséder un compte.

La fonctionnalité qui permet aux usagers d'établir une liste des relations constitue une autre propriété remarquable des RSN. Avec cette liste, les usagers peuvent décider de partager avec un groupe donné des informations et, *vice versa*, de recevoir celles qu'ils s'intéressent à connaître. La constitution de cette liste varie selon les services. Pour certains RSN, la catégorie « amis » ne sera possible qu'à travers un accord réciproque. Quand un usager envoie une invitation à un autre compte, il faut attendre que ce dernier accepte cette invitation pour qu'il puisse apparaître dans la liste de cet invitant. Une fois accepté dans la liste d'invités, le compte de cet usager va aussi être ajouté. C'est le cas de *Facebook* ou *LinkedIn*. Dans les autres cas, cette confirmation n'est plus obligatoire. On peut simplement cliquer sur « s'abonner » et « suivre les actualités d'un compte » sans demander un accord préalable. Cependant, cette liste est divisée en deux catégories : liste d'abonnés et liste d'abonnement. Pour les comptes d'abonnement, les contenus publiés par les abonnés ne seront pas disposés automatiquement sur leur page d'accueil. Par exemple, sur *Twitter*, la liste des relations est formée de cette manière. Ainsi, cette liste constitue le réseau social défini par les deux sociologues Ellison et Boyd (2013).

La troisième caractéristique cruciale des RSN est la possibilité de consulter et exploiter la liste des relations. Généralement, la liste des relations d'un compte est exposée publiquement en ligne, destinée aux membres de la liste ou à tout le monde selon les paramètres établis par l'utilisateur. À partir de cette liste, on peut connaître les interactions sociales d'un compte donné. Citons des exemples : les fonctionnalités comme « commenter » sur *Facebook* et « retwitter » sur *Twitter* permettent de visualiser et retracer le processus d'interaction. Quand un usager A commente ou retweete un contenu publié par un usager B, ce dernier va recevoir une notification du système. On peut donc savoir qui réagit à quel contenu et à quel moment. Le commentaire ou retweet apparaîtra dans les flux d'informations de ces deux comptes. Si ces comptes sont ouverts au public, cette interaction peut aussi être visualisée par les autres usagers au sein du

---

<sup>44</sup> Des plates-formes de vidéo chinoises proposent les mêmes fonctionnalités que *Youtube*.

système ou même à l'extérieur des RSN. Avec cette fonctionnalité de sociabilité, on peut aussi enrichir la liste en exploitant les listes des autres comptes. Sur *RenRen*, la version chinoise de *Classmates.com*, on peut simplement retrouver des camarades de classe à travers des relations en commun. En plus, suite à leur développement, les RSN ont lancé la technique dite « push » avec laquelle les usagers n'ont plus besoin de fouiller dans les listes pour retrouver des relations existant hors ligne ou établir un nouveau lien. À son tour, le système va proposer des comptes auxquels les utilisateurs peuvent s'abonner à partir de l'algorithme<sup>45</sup>. Cela dépend des comptes en commun ou des loisirs et intérêts indiqués préalablement par les usagers.

### 2.2.3. Evolution des RSN

Selon Boyd et Ellison, la définition de la notion de RSN (2011) s'est transformée vu les changements de la technologie et de la pratique sociale. Elles constatent que les RSN d'aujourd'hui ne sont plus dominés par les *profile-centric sites*<sup>46</sup> (Ellison et Boyd 2013:4). Autrement dit, le profil n'est plus qu'un espace distribué par le système où on entre un texte de présentation associé parfois à une photo d'image. Ce profil n'est pas non plus un élément important pour attirer les « amis ». La plupart du temps, ce profil ne change qu'au moment où on réalise une mise à jour.

Ellison et Boyd pensent qu'actuellement le profil consiste en une co-construction et a un champ élargi. Toutes les actions d'un usager et ses interactions qui se sont déroulées sur un RSN permettent de former un profil en temps réel. Quand le rôle du profil devient secondaire, le noyau des RSN se tourne plus vers le partage du contenu. Autrement dit, la mise en relation dépend moins de l'identité d'un usager, mais est beaucoup influencée par les informations publiées. En même temps, cette pratique de partage est enrichie par la possibilité de télécharger des photos, vidéos, musiques ou d'ajouter des technosignes ou technomots (Paveau 2015<sup>47</sup>) comme le hashtag, en plus des textes simples. Elles soulignent aussi que l'importance de la liste

---

<sup>45</sup> Dans le monde numérique, l'algorithme permet de regrouper des informations pour produire des résultats comme par exemple recommandation ou comparaison.

<sup>46</sup> Traduit comme « site centré sur le profil ».

<sup>47</sup> Il s'agit des « mots-consignes permettant d'accomplir des opérations en ligne », source : *Dictionnaire d'Analyse du Discours Numérique* (DADN) en ligne : <https://technodiscours.hypotheses.org/699>, consulté le 1er mars 2017. Paveau a d'abord publié en ligne sur son carnet de recherches un dictionnaire paru ensuite chez Hermann, dans la collection « Cultures numériques », *L'Analyse du discours numérique. Dictionnaire des formes et des pratiques*, en 2017.

des relations a été affectée par ces changements. Elle devient une fonctionnalité basique, moins décisive dans les RSN. Comme nous l'avons indiqué *supra*, avec la technologie de « push », la mise en relation n'est plus seulement basée sur la consultation et l'exploitation de la liste. Par l'algorithme, les RSN tendent à proposer des relations et des contenus « sur mesure » en étudiant la pratique des usagers en ligne.

Prenons l'exemple du microblog chinois *Sina Weibo* ; initialement, le contenu des comptes auxquels on s'abonne était affiché selon l'ordre chronologie inversé sur la page d'accueil. Dans ce cas-là, si on glissait de haut en bas, on pouvait visualiser tous les contenus. À partir de 2012, *Sina Weibo* a lancé « l'ordre intelligent ». Cet ordre est basé sur la fréquence d'interaction (l'intimité avec les autres comptes), la popularité des contenus (le nombre de consultations) et les intérêts des usagers. Par conséquent, on risque de rater certains contenus. Cette option est aussi critiquée par les usagers à cause de l'affichage répétitif d'un même contenu. Pour résumer, la définition (2011) privilège toujours les RSN comme une plateforme de communication tout en soulignant ses tendances à placer les interactions via le partage des informations, surtout par les flux, au centre des activités des RSN.

#### **2.2.4. Des notions en discussion dans la communauté des chercheurs**

À peine proposée la définition des RSN par Ellison et Boyd (2007), le sociologue Beer (2008) a exprimé ses interrogations – qui remettent en question leur définition – dans un article visant à communiquer avec ces deux chercheurs. Ses premières questions sont centrées sur la distinction entre *social network sites* (« réseaux sociaux numériques » en français) et *social networking sites* (traduit en français comme « réseautage social »). Ce dernier est surtout destiné à établir des relations entre des gens qui ne se connaissent pas. Or, Ellison et Boyd soulignent que ce qui rend les RSN uniques, ce n'est pas qu'ils permettent aux individus de se rencontrer mais plutôt qu'ils permettent aux utilisateurs d'articuler et de rendre visibles leurs réseaux sociaux. L'établissement des relations n'est pas le but mais un résultat de pratiques. À cet égard, Beer pense que le fait de distinguer les RSN des autres applications du Web en résumant leurs caractéristiques communes pose des problèmes. Pour lui, les RSN sont en phase de développement et les nouveautés apparaissent sans cesse. Bien évidemment, ces RSN possèdent des caractéristiques communes mais aussi des différences qu'il ne faut pas négliger. De ce fait, il considère que la définition d'Ellison et Boyd est trop vaste. Elle peut englober



d'une manière continue de nombreux nouveaux sites, une fois que certaines de leurs caractéristiques correspondent à la description. Il propose d'inclure les RSN sous une notion générale qui décrit le changement dans un contexte de pratiques (ex. : Web 2.0) et de différencier les RSN par catégories précises.

En citant les critiques de Beer (2008), Thelwall (2009) montre que la définition de Boyd et Ellison (2007) s'est focalisée sur le rôle de mise en relation des « amis » des RSN. Pourtant, même si certains sites possèdent les caractéristiques fondamentales des RSN (le profil, la liste, les flux d'informations), ils ne sont pas principalement destinés à construire un réseau de relations. Par exemple, il considère que *Youtube* correspond à la description de la première définition proposée par Ellison et Boyd mais que c'est plutôt un site dont l'objectif est la navigation. Nous pouvons aussi y ajouter des RSN récents comme *Instagram* qui sert à partager des photos et *Tiktok*, une plate-forme d'origine chinoise qui diffuse de courtes vidéos de 15 secondes. Ce sont des dispositifs qui ne sont pas destinés premièrement à la mise en relation des « amis ». Alors, ne sont-ils pas des RSN ?

Thelwall propose trois catégories de RSN pour différencier leurs objectifs :

1 – *Socialising SNS* (socialisation) a pour but de recréer les relations sociales existant souvent (mais pas toujours) déjà dans la vie quotidienne des usagers, ex. : *Facebook*, *RenRen*<sup>48</sup>, *Tencent Qzone*<sup>49</sup> ;

2 – *Networking SNS* (réseautage) sert à chercher et établir de nouvelles relations, ex. : *LinkedIn* ;

3 – (*Social*) *Navigation SNS* (navigation), concerne des services dont la destination première est d'offrir certaines informations spécifiques aux usagers : photos, vidéos, articles, etc. La liste de relations favorise donc la recherche de ces informations, ex. : *Youtube*, *Douban*<sup>50</sup>.

En nous référant à la définition (2011) des RSN, nous avons constaté que les RSN constituent toujours une notion vague bien que la communication orientée par le partage de contenu soit

---

<sup>48</sup> Un réseau d'Alumni qui permet de mettre en lien les camarades de classe ou les diplômés des mêmes établissements.

<sup>49</sup> La plate-forme proposée par la compagnie chinoise *Tencent* en 2005 offre des services multiples : écrire des blogs, écouter de la musique, partager des photos, etc. La consultation du contenu publié sur la plateforme est basé sur la mise en relation entre les utilisateurs.

<sup>50</sup> Un RSN lancé en 2005 qui sert à partager les informations et commentaires sur les livres, films, musiques, etc. Le contenu est fourni par les utilisateurs inscrits.

soulignée par les deux chercheurs pour prendre en considération les changements de pratique. Nous trouvons que la proposition de catégorisation de Thewall est pertinente vis-à-vis des confusions sur les objectifs initiaux des RSN.

Nous avons aussi remarqué que, en accompagnant la définition de 2011, Ellison et Boyd insistent toujours sur la nécessité de différencier *social networks sites* (réseaux sociaux numériques) de *social networks* (réseaux sociaux) et *social networking sites* (sites de réseautage social). Selon nos analyses précédentes, nous sommes d'accord sur le point que *social networks* est un terme sociologique pour désigner les relations des individus. Pourtant, nous nous questionnons sur la nécessité de distinguer le *network* et le *networking*. Si on reprend les catégories de RSN proposées par Thelwall, le réseautage – c'est-à-dire la recherche de nouveaux contacts – est considéré comme un des objectifs privilégiés des RSN. Ainsi, sur un RSN de socialisation ou de navigation, lorsqu'on établit la liste des relations, le phénomène de réseautage peut intervenir à condition qu'on s'abonne à des comptes d'inconnus. De plus, nous pensons que les tendances du développement des RSN d'aujourd'hui visent à une dimension mixte et uniforme. Les fournisseurs cherchent sans cesse à enrichir les fonctionnalités de leurs services en espérant englober le plus possible tous les objectifs dans une même application. De ce fait, nous trouvons que l'exclusion de l'objectif de réseautage des RSN n'est plus convenable actuellement et que *social networking sites* (réseautage social) est une sous-partie des RSN.

D'autre part, en considérant les changements technologiques et sociaux, nous nous demandons aussi si le mot « sites » (qui désigne en informatique l'ensemble des pages Web mises en ligne) dans l'expression anglaise *social networking sites* est toujours adéquat, puisque l'interface des applications sur smartphone et tablette est également introduite pour favoriser l'utilisation des RSN. Cette plate-forme de communication ne se limite plus aux sites Web en ligne.

Alexandre Coutant et Thomas Stenger (2009) ont quant à eux critiqué le fait que les usages des RSN ne sont pas précisés par Boyd et Ellison dans la première version de leur définition (2007). Ils ont donné de leur côté la proposition de définition suivante :

Les RSN constituent des services Web qui permettent aux individus : (1) de construire un profil public ou semi-public au sein d'un système, (2) de gérer une liste des utilisateurs avec lesquels ils partagent un lien, (3) de voir et naviguer sur leur liste de liens et sur ceux établis par les autres au sein du système, et (4) fondent leur attractivité essentiellement sur les trois premiers points et non sur une activité particulière. (Coutant et Stenger 2009:5)

En fait, cette définition est constituée à partir de la définition initiale de Boyd et Ellison (2007) en ajoutant un quatrième point. Cette propriété complémentaire permet d'associer les caractéristiques technologiques des RSN à leurs usages, c'est-à-dire qu'elle souligne le fait que l'objectif des RSN est axé sur la présentation du soi des usagers, surtout par l'interaction avec les autres. Distinguer cet usage spécifique des RSN aide à clairement identifier les dispositifs qui en relèvent. De ce point de vue en effet, les dispositifs organisés autour de centres d'intérêts ne doivent pas être considérés comme des RSN. Cela permet non seulement de distinguer les RSN des autres applications de communication, mais aussi de résoudre le problème de l'évolution. Si l'usage principal d'un dispositif ne correspond plus à cette définition du fait de son développement, ce dispositif sera exclu des RSN.

Nous pensons que cette précision sur les usages peut dans une certaine mesure diminuer le nombre des RSN et ainsi faciliter le travail d'identification. Mais elle ignore les pratiques individuelles différenciées des usagers. Si un dispositif est utilisé différemment par deux groupes, l'un centré sur des intérêts, l'autre centré sur la mise en relation, comment pourra-t-on les catégoriser ? D'autre part, Boyd et Ellison indiquent dans la définition (2011) que les activités des RSN sont centrées sur le partage des informations via l'interaction. Peut-on considérer cette action de partage comme une activité particulière ou une fonctionnalité basique ? Il faut souligner que, par rapport au paysage technologique et social de l'article, les services des RSN se sont multipliés. Ces dispositifs ont tendance à intégrer et mixer les fonctionnalités commercialement réussies. Avec l'augmentation du nombre d'utilisateurs et la diversification de leur identité, une seule sorte d'activité monotone et particulière sur un seul dispositif est peu probable. Si les dispositifs aux usages prescrits (partage des photos, vidéos comme *Instagram*, *Youtube*) sont différents des RSN au niveau de la technologie déployée, du support présenté, on peut dire qu'au niveau des usages, ils sont de la même lignée que les RSN. Nous pensons que les RSN d'aujourd'hui sont à usages évolués et mixtes.

Malgré le débat sur les caractéristiques des usages des réseaux sociaux numériques, nous constatons qu'un consensus se dégage : les réseaux sociaux numériques sont des moyens de communication en ligne par lesquels les utilisateurs s'expriment entre eux en disposant préalablement d'une identité virtuelle. À travers ce profil, on crée, on échange, on partage certaines informations. Cela permet d'entretenir des liens interpersonnels.

Il est actuellement clair aussi, qu'à la différence des réseaux sociaux qui désignent un ensemble de formes organisées de relations personnelles, les RSN sont marqués par leur dimension outillée. Cependant, les fonctions technologiques des RSN peuvent favoriser la formation des réseaux sociaux en ligne. De ce point de vue, nous approuvons ce que dit Degenne (2011:40) : « il ne faut pas confondre les outils de médiation et les réseaux eux-mêmes ». En résumé, le RSN désigne les services du Web – des outils de médiation – alors que les réseaux sociaux, constituent des systèmes formés par des relations entre des personnes (Pannier 2008:2).

### 2.2.5. Médias sociaux et RSN

Voyons maintenant une autre expression en la mettant en regard de « RSN ». Cette expression de « médias sociaux » est intimement liée au Web 2.0 et aux RSN. Comme ces deux derniers termes que nous nous sommes efforcée de préciser, elle reste assez floue à la fois pour les chercheurs et les publics qui bénéficient des services des médias sociaux. La confusion provient non seulement de ses relations avec les autres services similaires mais aussi de sa définition elle-même.

Ainsi, il est difficile de retracer l'emploi d'origine de ce terme. À travers une revue de littérature<sup>51</sup> autour de ce sujet, nous avons constaté que les articles ou les billets encyclopédiques ayant vocation à clarifier cette notion ne manquent pas. Nous allons nous intéresser à ces définitions venant de domaines différents et essaierons d'en récapituler les caractéristiques. En même temps, à partir des observations, nous questionnerons les relations entre les médias sociaux et les RSN pour les différencier.

Dans un premier temps, nous nous sommes appuyée sur l'outil de *Google Trends*<sup>52</sup> pour visualiser la fréquence des termes recherchés dans le moteur de recherche *Google*. En vue d'obtenir un résultat assez large, nous avons paramétré de la façon suivante : les régions de recherche concernent tous les pays ; l'expression anglaise *social medias* et française *médias sociaux* sont toutes deux prises en compte ; un troisième mot-clé concernant la recherche sur

---

<sup>51</sup> Nous avons eu recours aux références suivantes : (Anis 2006 ; Arènes 2015 ; Cardon 2011 ; Casilli 2012 ; Coutant et Stenger 2012 ; Gervais 2007 ; Sautédé 2011 ; Tisseron 2011 ; Yanoshevsky 2010).

<sup>52</sup> Outil permettant de connaître la fréquence d'un terme qui a été saisi dans le moteur de recherche Google. Les données peuvent être consultées en amont jusqu'à janvier 2004.

définition a été aussi ajouté avec pour but de faire la comparaison avec les termes recherchés. Selon les statistiques de la Figure 9 ci-dessous, nous pouvons faire quelques observations : 1) les recherches du terme existaient déjà en 2004 et ont eu tendance à augmenter dans l'ensemble ; 2) les recherches précises sur la définition du terme ne débuteraient qu'en 2011. Il faut souligner que ces résultats sont loin d'être exhaustifs vu les choix et les potentialités de composition des mots clés possibles. Mais cela peut nous fournir des pistes de réflexion que nous formulons ainsi : l'expression « médias sociaux » ou *social medias* a été proposée avant 2004 et a connu une certaine popularité, ce qui peut expliquer le pic du chiffre en 2004 ; l'augmentation des recherches montre que cette expression attire de plus en plus l'attention, et il est possible qu'il existe davantage de contenus qui lui sont relatifs ; le nombre moins important de recherches de définition peut être interprété de deux manières : soit les recherches directes des termes eux-mêmes fournissent suffisamment d'informations pour appréhender la notion de « médias sociaux » ; soit il n'existe pas véritablement de définition qui fasse consensus en raison des confusions ou des points de vue variés.

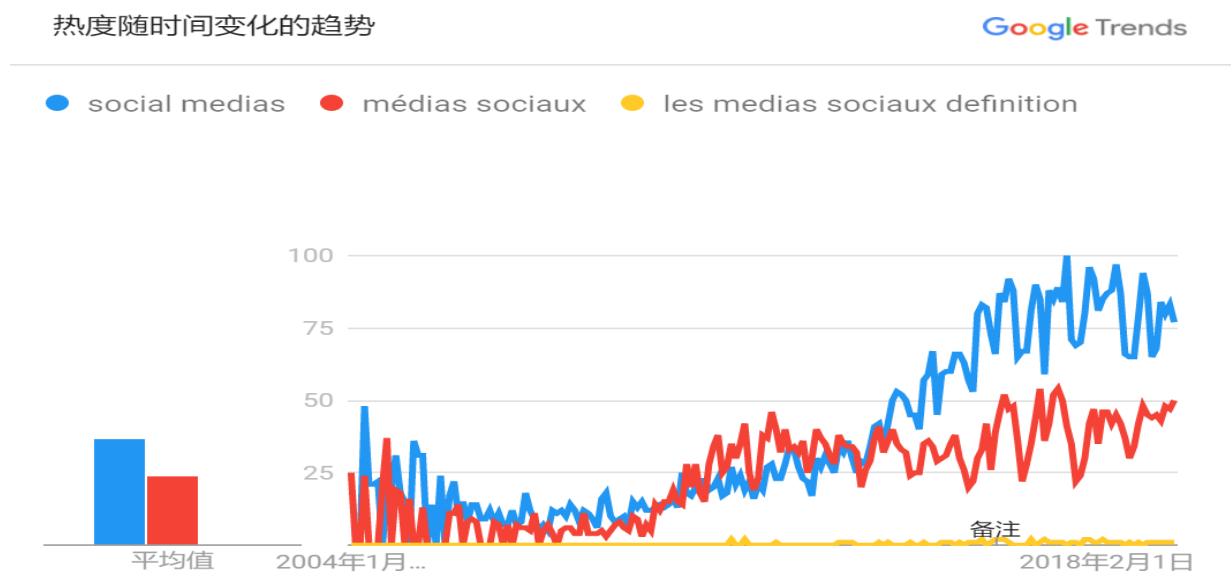


Figure 9 : Données de Google Trends

Après ces premiers éléments de réflexion, nous nous sommes tournée vers les ressources encyclopédiques particulièrement en ligne et les travaux scientifiques. Tout d'abord, nous avons

relevé que la définition de Cavazza (2018)<sup>53</sup> reconnu comme consultant de l'Internet, a été beaucoup reprise dans les billets d'information en ligne. Après avoir énuméré les diverses ressources expliquant la notion de « médias sociaux », il exprime son désaccord : il trouve que les définitions anglaises sont soit longues et mal organisées, soit trop courtes et imprécises et que les définitions françaises ne permettent pas de saisir l'essentiel du terme y compris *Wikipédia*, les billets des autres bloggeurs et des sites Web. Il donne sa propre définition sur son blog en ligne : « Les médias sociaux désignent un ensemble de services permettant de développer des conversations et des interactions sociales sur l'Internet ou en situation de mobilité ». Cette définition, par opposition aux médias traditionnels qui servent à la transmission en sens unique des informations, souligne la caractéristique interactive des médias sociaux. Elle montre en même temps que les médias sociaux consistent en « un ensemble de services » et non de technologies. En plus, les terminaux permettant d'utiliser les médias sociaux ne se limitent pas aux ordinateurs mais s'élargissent aux téléphones portables et aux tablettes.

Sous l'angle du marketing, une autre définition citée à maintes reprises est celle de Kaplan et Haelein<sup>54</sup> (2010) : « *Social Media is a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content* »<sup>55</sup>. Cette définition met l'accent sur l'influence décisive du Web 2.0 envers les applications de médias sociaux. Nous pouvons considérer que les médias sociaux sont des produits dérivés du Web 2.0. L'autre point important de cette définition est celui de la favorisation de production et de diffusion du contenu établi par les internautes.

Mais Coutant et Stenger, deux chercheurs dans le domaine des Sciences de la Communication sont critiques à l'égard de la définition de Kaplan et Haelein à cause du risque de confusions des notions de Web 2.0 et de médias sociaux. Ils proposent une approche sociotechnique :

---

<sup>53</sup> Une définition des médias sociaux : <https://fredcavazza.net/2009/06/29/une-definition-des-medias-sociaux/>, consulté le 28 mai 2019.

<sup>54</sup> Andreas Kaplan est spécialisé dans les médias sociaux et le marketing viral. Michael Haenlein s'intéresse aux études portant sur le gestion de la relation client et les médias sociaux. Ils sont tous deux enseignants-chercheurs en marketing à ESCP.

<sup>55</sup>Nous traduisons : « Les médias sociaux sont un groupe d'applications Internet qui reposent sur les fondements idéologiques et technologiques du Web 2.0, et qui permettent la création et l'échange de contenu généré par les utilisateurs ».

Les médias sociaux peuvent être définis comme des services l'Internet : 1. dont le contenu est très largement produit par les internautes utilisateurs (principe UGC : *user generated content*), 2. qui regroupent des configurations sociotechniques très variées en termes de dynamique de participation (par intérêt, par amitié) et de visibilité (nature et finalité des données publiées en ligne) (Coutant et Stenger 2013:115).

Outre le principe UGC, les auteurs considèrent les médias sociaux sous les angles de la participation et de la visibilité. Selon eux, la prise en compte à la fois des fonctionnalités technologiques des applications et des usages des utilisateurs permettent de repérer les différentes variétés de médias sociaux.

Parallèlement, nous avons constaté que le recours à l'étymologie est également une façon d'appréhender la définition des médias sociaux. Le « média » est un moyen de diffusion de l'information. En ajoutant la particularité de « social » (les activités des membres de la société), les médias sociaux ont un rôle d'intermédiaire favorisant les interactions sociales.

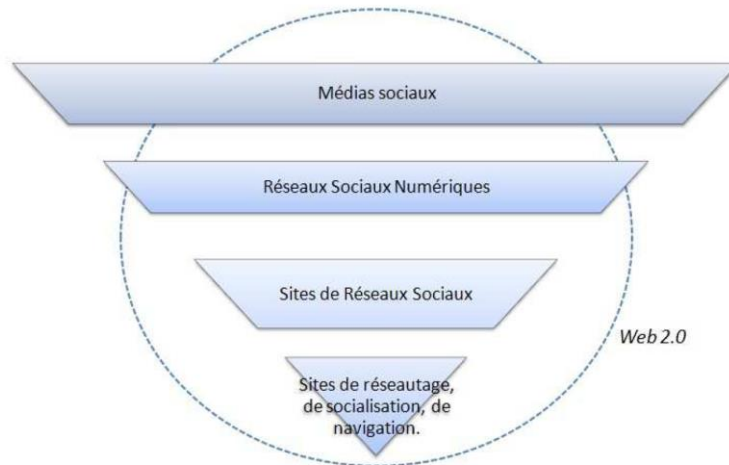
Nous pouvons conclure de ce qui précède que : 1) les médias sociaux sont reconnus comme des services bénéficiant des avantages technologiques du Web 2.0 et ont pour finalité de favoriser l'échange des informations ; 2) sur les médias sociaux, les contenus sont majoritairement produits par les utilisateurs.

Les médias sociaux et les RSN sont, les uns et les autres, des services émanant du Web 2.0. En tentant de clarifier la définition des médias sociaux, nous avons aussi observé quelques points de vue portant sur la relation entre médias sociaux et RSN. Nous les récapitulons ainsi :

1. les RSN constituent une sous-partie de médias sociaux ;
2. on ne fait pas de distinction entre médias sociaux et RSN ;
3. les médias sociaux sont une catégorie de RSN ;

Le premier point de vue est largement approuvé et circule tant en ligne que dans les travaux scientifiques. C'est ce point de vue auquel nous nous rangeons et que nous allons expliciter. Pour le vérifier, la question de la catégorisation des médias sociaux reste à résoudre. Dans la classification détaillée et assez complète de Cavazza (2018), les RSN sont considérés comme une des six catégories des médias sociaux, outre les outils de publication, de discussion, de partage, de collaborations, de messagerie. Cette classification est basée plutôt sur les

fonctionnalités. Girard et Fallery (2009) à partir de diverses études, élaborent une structure avec des relations hiérarchiques qu’ils schématisent de la manière suivante (Figure 10) :

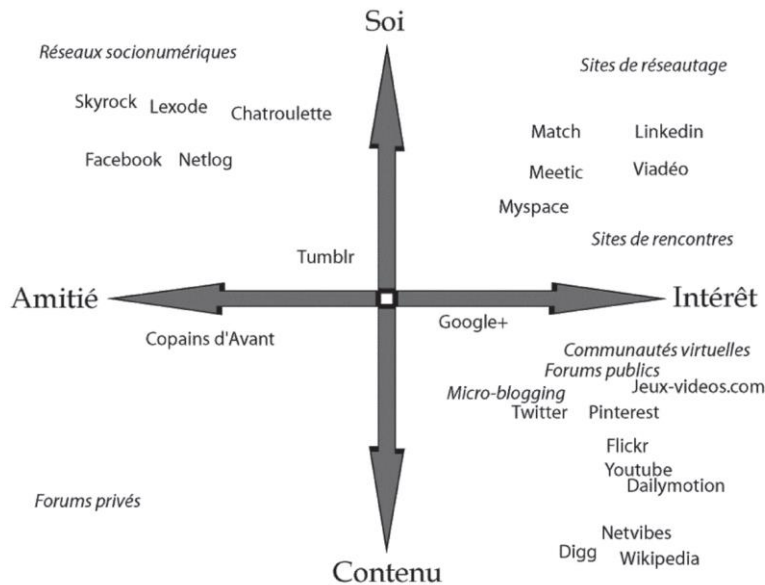


**Figure 10 : Relation entre médias sociaux et RSN selon Girard et Fallery (2009)**

D’après ces auteurs, le Web 2.0 regroupe un ensemble d’applications centré sur le partage d’information et la socialisation entre les utilisateurs. De haut en bas, la sphère des notions est de plus en plus précise. Ils plaident aussi pour que les RSN sont incluses dans les médias sociaux.

La catégorisation (Figure 11) de Coutant et Stenger (2013) est fondée, quant à elle, sur les fonctionnalités des médias sociaux et les pratiques des usagers.





**Figure 11 : Cartographie des médias sociaux par Coutant et Stenger (2013)**

À la différence du schéma de Girard et Fallery qui traite les différentes notions en montrant une hiérarchie à la manière des poupées gigognes qui rentrent les unes dans les autres, Coutant et Stenger proposent une cartographie dans laquelle des déplacements sont possibles. À partir de leur classification axée sur la participation et la visibilité, on obtient quatre repères principaux pour distinguer les variétés des médias sociaux. Le premier axe horizontal distingue participation conduite par un intérêt (à gauche) et participation conduite par l'amitié (à droite). Le second axe vertical établit la distinction entre la mise en visibilité de soi (la présentation de soi) ou la mise en visibilité de contenu tiers (c'est-à-dire le partage du contenu des autres). Dans l'espace défini par ces axes figurent des catégories (par exemple « Sites de rencontre » ou « Forums privés ») et/ou des noms de médias sociaux (par exemple *Wikipédia* ou *Tumblr*). Ces deux axes combinant à la fois les principales fonctionnalités des plateformes constatées et les usages dominants, servent, selon les auteurs, à distinguer les divers dispositifs dans une perspective sociotechnique.

Nous pouvons constater que les RSN se retrouvent au sein de la cartographie dans la partie supérieure gauche. Cet emplacement implique que les médias sociaux englobent les dispositifs de RSN qui sont caractérisés par leurs pratiques centrées sur l'amitié et la présentation de soi, par exemple *Facebook*. Les sites de réseautage se différencient des RSN parce que les pratiques sur ces dispositifs selon Coutant et Stenger sont organisées autour d'un intérêt précis et sont

placés en haut à droite de la cartographie. À titre d'exemple, *LinkedIn* est considéré comme un site de réseautage pour son intérêt lié au domaine professionnel. Nous retrouvons en bas à droite de la cartographie plusieurs types de dispositifs caractérisés par le partage de contenu et un centre d'intérêt précis, dont les forums publics et les communautés virtuelles. *YouTube* qui a pour fonction de partager des vidéos se trouve ainsi dans ce secteur. Il est intéressant de constater que les dispositifs de microblogging, notamment *Twitter* qui est souvent considéré comme un RSN est ici davantage vu comme une plateforme centrée sur un intérêt ainsi que le partage du contenu tiers.

Il est à noter que la position des plateformes dans la cartographie n'est pas fixe et que le changement de place est possible selon l'évolution des fonctions et des pratiques. Les auteurs ont ainsi souligné que les emplacements proposés reflètent les situations des plateformes observées lors de l'établissement de la cartographie en 2013. Le cas de *Twitter* en est un exemple. Selon sa situation initiale, il aurait été placé avec les RSN. Comment traiter ce microblog à l'heure actuelle ? C'est ce que nous allons aborder *infra*.

Si nous reprenons les définitions des deux termes, nous pouvons relever les distinctions suivantes entre les médias sociaux et les RSN sous divers aspects. Tout d'abord la finalité : les médias sociaux ont pour objectif de diffuser et partager des contenus produits par les utilisateurs alors que les RSN privilégient la mise en relation à travers le partage des informations, c'est-à-dire que le contenu, les informations sont au centre des médias sociaux tandis que les utilisateurs sont le noyau conceptuel du RSN. Ensuite concernant les conditions préalables, disposer d'un profil est indispensable pour utiliser les dispositifs des RSN. Le profil est une des caractéristiques clés des RSN, alors que la création de profil n'est pas nécessaire pour les médias sociaux. Enfin pour les fonctionnalités technologiques une liste de membres associés est obligatoire pour les RSN mais pas pour les médias sociaux, dans les définitions qui en sont données.

D'ailleurs, nous constatons que les caractéristiques des RSN sont bien précises dans les définitions, alors que ce n'est pas le cas pour celles des médias sociaux qui se limitent à une description assez générale. Nous supposons que l'hétérogénéité des applications incluses dans les médias sociaux ne permettent pas d'aboutir à une définition convenant à leurs multiples caractéristiques. Il nous paraît pertinent d'intégrer les RSN au sein des médias sociaux sans

pour autant considérer que les médias sociaux et les RSN sont une seule et même catégorie et que les deux termes soient interchangeables.

## **2.3. Le microblog, en tant que RSN**

### **2.3.1. Présentation générale**

Nous allons maintenant aborder une catégorie cruciale pour notre recherche qui est le microblog et la situer par rapport aux RSN. Le microblog désigne une sorte de plateforme permettant aux utilisateurs de communiquer instantanément à travers des publications courtes, textes, images fixes et vidéos. La longueur d'une publication est souvent limitée dans le but de transmettre un contenu léger, de faciliter une lecture rapide et de réaliser la diffusion la plus large possible. Le partage et la diffusion des informations sur microblog dépendent d'un mécanisme, l'abonnement. Autrement dit, il faut s'abonner à une liste des comptes qui devient la source principale des informations qui s'affichent sur la page d'accueil du microblog. On peut accéder au service de microblog depuis un ordinateur (à partir d'une page Web), depuis un smartphone ou une tablette (à partir d'une application).

Le microblog<sup>56</sup> est souvent considéré comme une version concise du blog, mais ce sont deux services bien différents malgré des fonctionnalités similaires. La limite des caractères dans une publication est une des spécialités exclusives du microblog.

Le mot anglais « blog » est une aphérèse du mot composé « we log » qui est employé pour la première fois par un blogueur américain Jorn Barger en 1997 ; « we » renvoie à « Web » et « log » peut signifier journal, en référence au journal de bord des navigateurs. Le mot raccourci « blog » est créé par Peter Merholz, un professionnel dans le domaine des médias numériques en 1999. Ce nom définitif de blog constitue en 2004 une entrée du dictionnaire américain *Merriam Webster* puis en 2006 des dictionnaires français (*Le Petit Larousse* et *Le Robert*).

---

<sup>56</sup> Nous avons eu recours aux références scientifiques suivantes : (Kaplan 2011 ; Mercanti-Guérin 2010) et aux ressources encyclopédiques ci-dessous :

<https://www.ecrirepourle.com/histoire-du-blog/> ;

<http://www.contentmarketingacademie.fr/33-ans-dhistoire-blogging-9-dates-vintages-blog/>;

<https://www.tests-et-bons-plans.fr/hors-serie-on-vous-explique-tout-sur-l-histoire-des-blogs.html>.

La première production sous forme de blog peut remonter à 1994, bien avant la création du nom, sur le site *Links.net*. La première plateforme de blog, *Open Diary*, est mise en ligne en 1997. Dans les années 90, le blog s'est développé surtout en tant que journal intime. Il faut attendre les années 2000 pour qu'il commence à prendre son essor. Le développement technologique permet de former des outils variés pour simplifier la création de journaux en ligne et rendre le service plus facile à utiliser. Le service de blog est ainsi popularisé. Diverses plateformes de blog sont créées à cette période. Par exemple, le produit américain *WordPress* ; du côté francophone, le *Skyblog*. Aujourd'hui, le service de blog suppose un service de publication des articles personnels en ligne, bien évidemment sans contrainte de longueur. Pour chaque publication, il permet aux lecteurs de réagir avec des fonctionnalités telles que commenter, transférer... Ses publications sont affichées dans un ordre chronologique inverse, de la publication la plus récente à la plus ancienne. Les blogs peuvent être des sites Web ou faire partie d'un site. Par exemple, le service de blog ne constitue qu'une section de la plateforme *WordPress* qui propose en même temps des sites de fonction variées dont le site de vente en ligne.

Donc, si on considère la fonctionnalité de publication, la possibilité d'interaction et le moyen d'affichage, microblog et blog comportent des points communs. Pourtant, nous voulons souligner une grande différence entre ces deux outils de communication. Même si le blog permet aux utilisateurs de s'abonner entre eux, c'est-à-dire de construire une relation réciproque et d'établir ainsi une liste de relation, on ne peut visualiser les contenus d'autrui qu'à travers la consultation de sa page d'accueil ou de sa liste de publications. Autrement dit, les utilisateurs doivent chercher par eux-mêmes les informations. Contrairement au blog, le microblog bénéficie d'une fonctionnalité enrichie de cette liste : les contenus venant des comptes d'abonnement s'affichent en temps réel et sur la page de consultation du compte d'abonné. Grâce à cette fonctionnalité, l'effort de la recherche des informations est économisé ; la diffusion rapide devient possible. Nous pensons que le partage d'information se basant sur une liste de relations préconstituée reste la caractéristique phare du microblog à laquelle s'ajoute le format léger et l'instantanéité d'interaction.

Pourtant, catégoriser le microblog dans les médias sociaux ou les RSN ne va pas de soi et peut donner lieu à débat. Il n'est pas faux de considérer le microblog comme un média social si on le considère comme une catégorie générale du Web 2.0 (voir *supra*). Pour autant est-il pertinent

d'inclure le microblog dans les RSN ?

Si nous reprenons la définition (2011) que Boyd et Ellison donnent des RSN, nous pouvons constater que les aspects technologiques du microblog correspondent bien aux trois caractéristiques qu'elles énoncent. Premièrement, les utilisateurs du microblog doivent se procurer un compte et constituer un profil. Deuxièmement, une fonctionnalité essentielle du microblog est la liste de relations. Troisièmement, sur un microblog, le flux d'informations basé sur la liste favorise la diffusion du contenu créé par les utilisateurs et l'interaction interpersonnelle. D'un point de vue fonctionnaliste, le microblog peut être considéré comme un RSN. Mais ce point de vue peut être discuté.

Rappelons que la première version de Boyd et Ellison (2007) est critiquée par Stenger et Coutant qui montrent que le manque de critères sur les usages cause souvent des confusions lors de la catégorisation des services. Ainsi, plutôt qu'une analyse fonctionnaliste, ils adoptent une approche sociotechnique pour classer les médias sociaux : « une analyse fondée à la fois sur les fonctionnalités offertes par les plateformes et sur les pratiques effectives des utilisateurs » (Coutant et Stenger 2013:112).

Selon la cartographie de la Figure 11 constituée de plusieurs axes, proposée par les deux chercheurs, nous avons remarqué que la participation sur les microblogs, surtout sur *Twitter* est conduite par l'intérêt et non par l'amitié ; la mise en visibilité du contenu tiers est privilégiée par rapport à la publication de soi. Ajoutons que les auteurs soulignent que les usages des RSN doivent se fonder sur la mise en relation et non sur une activité particulière motivée par des intérêts. De ce point de vue, le microblog s'il est bien une sorte de média social ne fait pas partie des RSN à proprement parler.

Nous avons mentionné ci-dessus que les médias sociaux mettent en avant le partage des informations, ce qui correspond à « la publication de contenu tiers » (Coutant et Stenger 2013), alors que les RSN visent à favoriser la communication et l'interaction à travers la mise en relation et la présentation de soi, qui selon la définition de 2011 combine le profil d'utilisateur et le contenu généré sur le flux d'information. Les techniques et les pratiques changent en continu alors que les définitions des RSN n'ont pas connu de mise à jour depuis 2011 à notre connaissance. On peut observer que, côté fonctionnalités, les microblogs ne cessent pas

d'enrichir leurs services<sup>57</sup> ; et que, côté usages, les pratiques peuvent varier d'un individu à l'autre, d'une communauté à l'autre, d'un pays à l'autre ou d'une langue à l'autre<sup>58</sup>. Nous considérons qu'aujourd'hui sur les microblogs, les types de participations et de publications ont tendance à se mixer et se multiplier ; c'est-à-dire que la participation peut être conduite par l'amitié et la recherche de contact social, par un centre d'intérêt déterminé, par un objectif commercial ou politique, par un souci d'expression de ses opinions. L'ouverture des comptes publics de compagnies, institutions, personnages politiques en témoigne. Les tweets d'un compte d'individu ne sont pas forcément tous concernés par la publication de soi. Le partage d'un contenu tiers est tout à fait possible. Ces sont les pratiques qui déterminent les usages opérés des RSN. Cela rejoint les points de vue de Boyd et Ellison qui pensent que le rôle du profil et la liste d'amis sont moins importants qu'auparavant. Ces deux caractéristiques technologiques et sociales restent des fonctionnalités basiques visant à favoriser la communication. Les RSN sont aujourd'hui une plateforme de communication se basant sur le partage d'informations. Toutes ces réalités technologiques et sociales nous font considérer le microblog comme un RSN actuel.

En effet, prenant en compte les évolutions permanentes au niveau des fonctionnalités et des pratiques, nous pensons que le microblog d'aujourd'hui tend à combiner les rôles de transmission d'informations et de mise en relation. Même si ses fonctionnalités relatives aux réseaux sociaux numériques sont atténuées par les pratiques concrètes des internautes, on peut considérer le microblog en gros comme un média social et plus précisément un réseau social numérique.

### **2.3.2. Twitter et Sina Weibo**

Dans cette étude, nous nous intéressons spécialement aux aspects technologiques et relationnels d'un RSN particulier, le microblog. Dans l'objectif de réaliser une étude comparative, nous

---

<sup>57</sup> Par exemple, les divers fonctions élaborées (y compris la centralisation automatique des messages sous un même sujet ou bien la rubrique *Tendances*) autour de « hashtags » ne sont pas originellement appliquées dans les microblogs.

<sup>58</sup> *Sina Weibo* dispose une version en chinois et une version en anglais : <http://english.sina.com/weibo/>. *Twitter* est actuellement disponible en anglais, français, allemand, italien, japonais, coréen et espagnol : [https://blog.twitter.com/fr\\_fr/a/fr/2011/traduire-twitter.html#:~:text=Twitter%20est%20actuellement%20disponible%20en,pouvoir%20am%C3%A9liorer%20les%20traductions%20pr%C3%A9existantes.](https://blog.twitter.com/fr_fr/a/fr/2011/traduire-twitter.html#:~:text=Twitter%20est%20actuellement%20disponible%20en,pouvoir%20am%C3%A9liorer%20les%20traductions%20pr%C3%A9existantes.)

avons choisi un microblog chinois *Sina Weibo* et l'un de ses homologues international<sup>59</sup>, *Twitter* (version française) comme support d'analyse. Ce choix prend en compte à la fois les caractéristiques technologiques et les aspects communicationnels et socio-culturels. Nous allons procéder dans un premier temps à une brève présentation de ces deux dispositifs et nous donnerons par la suite un aperçu des études existantes portant respectivement sur les deux microblogs.

Le nom de *Sina Weibo* (新浪微博, xīn làng wēi bó) est composé par deux termes : « Sina » (新浪, xīn làng) renvoie au nom de la compagnie de création pour distinguer les services de microblog fournis par les autres marques en Chine ; « weibo » est la traduction chinoise du mot « microblog » (微 wēi signifie micro et 博 bó est l'abréviation du terme 博客 bó kè qui signifie blog). Le service de *Sina Weibo* est lancé en 2009 suite au blocage de *Facebook* et *Twitter* en Chine<sup>60</sup>. Il s'agit de la première initiative chinoise proposant un service de microblog. Avec l'augmentation forte des utilisateurs inscrits et une quantité importante de publications journalières<sup>61</sup>, il devient très vite le microblog le plus populaire en Chine<sup>62</sup> et il est considéré comme « le *Twitter* chinois » à l'international (*Sina Weibo* dispose aussi d'une version anglaise). Essayons de montrer la pénétration importante de *Sina Weibo* dans la vie quotidienne des chinois par une expression humoristique qui circule en ligne : « Quelle est la première chose que fait un chinois au réveil le matin après avoir ouvert les yeux ? Il se connecte à *Sina Weibo* ». L'importance de *Sina Weibo* en Chine peut aussi être expliquée par les usages développés autour du dispositif sur lequel les informations peuvent être diffusées rapidement à une échelle nationale. Prenons un exemple : *Sina Weibo* a été choisi comme un canal pour publier et diffuser les avis de recherche. Ainsi, les influenceurs de domaines différents, les organisations chinoises et internationales, les institutions gouvernementales, les célébrités ou les individus ordinaires sont inscrits sur *Sina Weibo* et disposent de comptes.

---

<sup>59</sup> Le siège de l'entreprise se trouve aux Etats-Unis.

<sup>60</sup> Le blocage a été motivé par des « Émeutes au Xinjiang en juillet 2009 » où une manifestation initiale rassemblant les Ouïghours s'est transformée par la suite en attaques contre des civils. Ces attaques ont entraîné 184 morts dont 137 Hans. Source : <https://www.fmprc.gov.cn/ce/cejo/chn/dtxw/t573214.htm>.

<sup>61</sup> Selon la compagnie *Sina*, le nombre des utilisateurs avait déjà atteint 50 millions fin 2010, soit un tiers de la population chinoise. Plus de 25 millions de publications journalières ont été constatées. Voir : <https://tech.sina.com.cn/i/2010-11-16/10314870771.shtml>.

<sup>62</sup> Nous énumérons quelques microblogs chinois dont certains ne sont plus en usage actuellement : Fanfou, microblog de *Tencent*, microblog de *Sohu*, microblog de *Phoenix New Media*, etc.

Le service de *Twitter* a été lancé en 2006 par l'entreprise américaine. Pour représenter ses caractéristiques fonctionnelles, son nom se réfère aux gazouillis d'oiseaux<sup>63</sup> qui sont courts et rapides. Une fois lancé, ce service est rapidement devenu populaire. Selon les statistiques fournies par l'agence We Are Social et Hootsuite (avril 2020)<sup>64</sup>, *Twitter* est classé au 6<sup>e</sup> rang des dispositifs<sup>65</sup> le plus utilisés dans le monde et le 7<sup>e</sup> en France<sup>66</sup>. Parmi les sept dispositifs les plus utilisés en France, *Twitter* est le seul dispositif de microblog, occupant ainsi une place prédominante. Les usages et pratiques de *Twitter* en France sont aussi très divers et nous remarquons la communication politique et informative. Selon une enquête<sup>67</sup> menée sur l'utilisation de *Twitter* dans la sphère politique française, 59 des 61 personnalités politiques françaises parmi les plus importantes sont inscrites et actives sur *Twitter*. Le compte de François Hollande a été classé en 1<sup>er</sup> par le nombre d'abonnés en 2015 (980.000). Aujourd'hui, le compte du Président Emmanuel Macron dépasse 6.7 millions d'abonnés. En même temps, la population sur *Twitter* est également très politisée. D'après une étude menée sur les usages politiques de *Twitter* (Boyadjian 2016), 56% des participants au questionnaire ont répondu qu'ils s'intéressent beaucoup à la politique. Le niveau élevé de politisation des participants est aussi manifesté par une pratique importante d'activités politiques comme : signer une pétition, manifester, adhérer à un parti ou faire grève, etc. Côté informatif, *Twitter* est également choisi par des organes de médias de presse ou des journalistes pour diffuser les informations en temps réel. Leurs comptes inscrits en sont un témoin. Par exemple, *Le Monde* dispose d'un compte sur *Twitter* depuis 2009 et il est suivi maintenant par 9,1 millions d'abonnés (beaucoup plus que celui du Président de la République française). En parallèle, *Twitter* est également utilisé par les Français pour diffuser et partager les informations de première main dont ils témoignent. Par exemple, après le choc des attentats du 13 novembre 2015, les Français ont réagi rapidement sur *Twitter* pour s'informer, s'organiser, et se rassurer. Des hashtags ont été créés autour de ce sujet dans les heures qui ont suivi comme : #porteouverte, #prayforparis, #bataclan.

Les points communs constatés dans leurs fonctionnalités techniques (surtout la contrainte du

---

<sup>63</sup> Le mot anglais *twitter* signifie gazouillis et le logo du service est une silhouette bleue d'oiseau.

<sup>64</sup> Disponible sur : <https://wearesocial.com/fr/blog/2020/04/le-digital-dans-le-monde-en-avril-2020>.

<sup>65</sup> Le Top des 5 sont les suivants : *Google, YouTube, Facebook, Baidu, Wikipédia*.

<sup>66</sup> Le Top des 6 sont les suivants : *YouTube, Facebook, Messenger, WhatsApp, Snapchat*.

<sup>67</sup> Disponible sur :

<https://www.sia-partners.com/fr/actualites-et-publications/de-nos-experts/lutilisation-de-twitter-dans-la-vie-politique-et>.



message) et leurs importantes utilisations dans les deux pays (Chine et France) nous ont incitée à effectuer une observation sur ces deux microblogs. Le développement des services de microblog a aussi attiré l'attention des chercheurs venant de domaines différents. Ainsi, *Sina Weibo* et *Twitter* suscitent des questionnements qui intéressent à la fois les chercheurs chinois et français. Notons quelques axes de recherche concernant ces deux microblogs.

En Chine, un bon nombre<sup>68</sup> d'études sont menées par les chercheurs en Sciences de l'Information et de la Communication. Les propriétés technologiques du microblog et leurs influences sur les médias traditionnels constituent les premiers sujets d'étude (Huang 2011) ; le mécanisme de diffusion au sein du microblog et les pratiques des usagers sont ensuite analysés à travers des études de cas. Par exemple, dans son travail de mémoire de master, Zhang (2011) a choisi d'observer et de classer les contenus publiés par une actrice dont le nombre d'abonnés sur *Sina Weibo* est au premier rang, pour dégager à partir de là les caractéristiques de diffusion ; Chen (2012) s'intéresse spécialement aux comptes d'influenceurs sur *Sina Weibo* pour étudier les relations qu'entretiennent la publication de soi et l'environnement du microblog. En parallèle, les thématiques du microblog sont aussi abordés dans des perspectives sociologique, marketing, anthropologique et dans celle de l'analyse du discours : Hong (2015), par exemple, a réfléchi au développement du discours féministe chinois dans le contexte du microblog ; Huang et Ruan (2018) ont choisi d'étudier les pratiques publicitaires sur le microblog *Sina Weibo* ; l'étude de Wang (2017) porte sur la violence verbale ; Zhao (2012) a tenté d'analyser les interactions sur *Sina Weibo* d'un point de vue anthropologique.

En France, les études portant sur le microblog *Twitter* sont menées en Sciences de l'Information et de la Communication, en linguistique, en sociologie, en marketing. Notons quelques exemples en SIC et SDL. Bigey (2018) analyse les modalités d'usages et de pratiques de *Twitter* ; Alloing et Vanderbiest (2018) se sont interrogés sur l'importance des fonctionnalités

---

<sup>68</sup> Selon les recherches dans la base de donnée chinoise CNKI, *China National Knowledge Infrastructure* (Infrastructure nationale de connaissances de Chine), créée par l'Université Tsinghua et supervisée par le ministère de l'Education de la République populaire de Chine, est l'un des premiers éditeurs autorisés par l'Etat à l'édition et la publication en ligne de ressources numériques. Il est le portail de connaissances le plus complet de la Chine, proposant des contenus de collection bien distincts : périodiques, journaux, annuaires, ouvrages de référence, thèses, mémoires de maîtrise.

de *Twitter* dans la circulation de la fausse information. Le produsage<sup>69</sup> (Paveau 2015) des hashstags sur *Twitter* est étudié par Thiault (2015). Julliard (2015) a tenté d'analyser des controverses sur *Twitter* d'un point de vue sémiotique. En Sciences du langage, Gautier (2017) et Longhi (2013) s'intéressent tous deux aux tweets à teneur politique et à leurs formes linguistiques. Les recherches de Paveau (2012 ; 2013) sont centrées sur les caractéristiques des productions discursives dans le contexte numérique et des méthodologies d'analyse. Elle a pris ainsi l'étude de *Twitter* comme un exemple. Reboul-Touré (2016) s'est interrogée sur les influences des nouvelles modalités d'expression de *Twitter* sur la créativité lexicale.

Dans notre étude, nous voudrions observer la circulation des nouveaux phénomènes langagiers sur les deux microblogs. Le premier apport de notre étude est cette perspective comparative qui est rarement constatée dans les travaux dont nous avons fait état. En plus, nous nous inscrivons également dans une perspective interdisciplinaire alliant la linguistique au communicationnel et à l'analyse discursive ce qui amène à considérer à la fois le dispositif, les productions et les usagers. Les études relevées *supra* s'inscrivent souvent dans un ancrage unique. Enfin, à la différence de certaines études quantitatives (par exemple, celle de Thiault et Longhi ou Bigey), nous allons nous appuyer sur une méthodologie qualitative pour acquérir une compréhension en profondeur, et nous allons opter pour une analyse contextualisée<sup>70</sup> (voir *infra*).

#### **2.4. Rôles des usagers et des chercheurs dans le contexte numérique**

L'influence de l'Internet et ses services consiste non seulement à modifier les moyens de communication mais aussi les pratiques des usagers. En retour les pratiques des usagers modifient les dispositifs qui s'adaptent à ces évolutions et en engendrent de nouvelles dans une suite de réactions. Les internautes interviennent eux-mêmes dans le processus de création et de partage du contenu au lieu d'être des récepteurs passifs. Nous allons nous appuyer principalement sur deux notions que nous définirons celles de culture participative et

---

<sup>69</sup> Initialement proposé par Axel Bruns, le produsage décrit le fait que « dans une communauté collaborative en ligne, la frontière entre producteur et usager de contenu s'efface [...] ». Source : *Dictionnaire d'Analyse du Discours Numérique* (DADN) en ligne : <https://technodiscours.hypotheses.org/682>, consulté le 1er mars 2017.

<sup>70</sup> « L'analyse qualitative par contextualisation consiste à faire surgir le sens d'un élément (objet, événement, récit, etc.), et notamment d'un texte ou d'un élément d'un corpus, en lui posant des questions après l'avoir mis en rapport avec différents contextes constitutifs de différents environnements dans lesquels il s'insère[...] » (Paillé et Mucchielli 2012).

d'observation participante pour identifier certaines caractéristiques engendrées par ces pratiques et pour choisir une démarche qui nous aide à mieux cerner les nouveaux phénomènes langagiers et leurs diffusion dans le contexte de l'Internet .

#### **2.4.1. La culture participative**

La notion de « culture participative » (*participatory culture* en anglais) est employée pour la première fois par le chercheur américain Henry Jenkins, spécialisé dans le domaine des nouveaux médias, dans son ouvrage intitulé *Textual Poachers* en 1992. Remarquons que cette publication est bien antérieure à l'avènement du Web 2.0. Ce travail portant sur une communauté spéciale, les fans, a mis en lumière une culture dans laquelle les individus n'agissent pas uniquement comme des consommateurs de contenu mais aussi comme créateurs ou contributeurs. Cela rejoint la notion de « produsage », mot-valise construit à partir de production et usage, proposé plus tard en 2008 par le spécialiste des médias Axel Bruns. Cette notion décrit le fait suivant : « dans une communauté collaborative en ligne, la frontière entre producteur et usager de contenus s'efface : l'internaute cumule les deux fonctions et devient un agent hybride, un produser. » (Paveau 2015<sup>71</sup>).

Dans les publications ultérieures de Jenkins (2013 ; 2014) sa réflexion à propos de la culture participative s'est articulée en rapport avec le contexte d'Internet. Il est à noter que cette dimension participative de Jenkins est fortement axée sur les fans et le *fandom*<sup>72</sup> dans ses études mais que la notion est désormais répandue hors de ce domaine, dans de multiples disciplines des Sciences humaines et sociales, dont les Sciences de l'éducation, les Sciences de l'Information et de la Communication et, notamment, les études de médias sociaux.

C'est la réflexion sur les mutations des rôles et des pratiques dans le contexte de l'Internet qui nous incite à nous intéresser à la culture participative. La notion met bien évidence que les nouvelles technologies de l'information et de la communication ne se réduisent pas à fournir des nouveaux outils. Elles produisent de nouvelles façons de comprendre, de réagir, de penser qui sont manifestées par l'engagement des individus dans différentes activités créatives. Par

---

<sup>71</sup> Dictionnaire d'Analyse du Discours Numérique (DADN) en ligne : <https://technodiscours.hypotheses.org/682>, consulté le 1<sup>er</sup> mars 2017.

<sup>72</sup> La culture propre à un ensemble de fans.

exemple à travers la pratique du *fansubbing*<sup>73</sup> que nous avons étudiée dans notre recherche de master, les individus élaborent librement leurs propres interprétations, – exactes ou non –, d’une production culturelle étrangère et font partager leurs connaissances linguistiques et culturelles par une traduction (Gao 2013), ouvrant ainsi une autre possibilité d’appropriation des œuvres étrangères ; le *Danmaku*<sup>74</sup> (appelé également en chinois 弹幕, dān mù) permettant d’afficher les commentaires sur une vidéo en temps réel a non seulement modifié l’emplacement et la manifestation des commentaires, mais aussi changé profondément le rapport aux productions audiovisuelles, favorisant l’investissement personnel. Pendant qu’on regarde une vidéo, on peut réagir immédiatement au contenu et ainsi exprimer un point de vue ou expliciter un scénario, etc... Cette réaction est lue par d’autres internautes au cours du visionnage de la vidéo qui peuvent choisir de répondre aux commentaires, d’ajouter les leurs ou de ne rien faire. Les commentaires ne constituent plus un élément séparé et sont intégrés dans le contenu. Une vidéo est ainsi enrichie voire recrée dans un processus continu. L’appropriation d’une production audiovisuelle devient collective, interactive par rapport au dispositif et interactionnelle entre internautes.

La culture participative implique de penser les éventuelles influences sur les productions langagières de la participation des individus dans le processus de création et/ou de diffusion. Dans le cadre de notre recherche sur les phénomènes langagiers en contexte numérique, notamment sur les microblogs, nous supposons que les individus interviennent sur ces phénomènes. Nous nous demandons : quelles sont les caractéristiques linguistiques engendrées par cette participation ? Quelles sont les situations d’énonciation qui favorisent la création ? Quelles sont les spécificités de la diffusion quand les individus y participent ? Dans un deuxième temps, ce cadre de réflexion nous conduit également à prendre en considération les identités des comptes et les objectifs des usagers de RSN concernant l’interprétation et la diffusion des contenus.

#### **2.4.2. L’observation participante**

Le contexte d’Internet nous invite à nous interroger sur les questions touchant à la fois aux

---

<sup>73</sup> Celle-ci consiste pour des amateurs à sous-titrer une production audiovisuelle dans une langue donnée.

<sup>74</sup> Fonction d’origine japonaise, permettant d’afficher les commentaires d’une vidéo en temps réel sur l’écran. Cette fonction est facultative sur les sites proposant ce service.

usagers et aux observateurs. Cela nous amène à nous appuyer sur une notion qui concerne la posture de recherche en Sciences sociales, l'« observation participante ». L'observation participante est une méthode de recherche couramment utilisée dans les travaux ethnographiques en anthropologie, en sociologie. Par exemple, en sociologie, Bogdan et Taylor donnent la définition suivante :

une période d'interactions sociales intenses entre le chercheur et les sujets, dans le milieu de ces derniers. Au cours de cette période, des données sont systématiquement collectées [...]. Les observateurs s'immergent personnellement dans la vie des gens. Ils partagent leurs expériences (Lapassade 2002).

Une telle méthode exige l'engagement des chercheurs sur le terrain et dans les activités d'étude pour atteindre la compréhension des sujets. La question de la prise de distance ou non du chercheur sur son sujet d'étude est particulièrement discutée en Sciences humaines et sociales bien au-delà du cas de l'observation participante. Sans entrer dans le débat de la neutralité du positionnement, nous voudrions souligner que la participation des chercheurs est nécessaire mais spécifique dans le cas de l'observation des créations langagières et de leur circulation au sein des RSN, en tenant compte du fait que les interactions se font en ligne. Côté technique, la maîtrise des fonctionnalités des RSN est obligatoire pour les chercheurs. Pour le cas du microblog, il faut disposer tout d'abord un compte pour pouvoir effectuer les observations et le chercheur doit donc être un usager de microblog. Les opérations de recherche sont basées directement sur une pratique indispensable à l'observation et à la collecte de données. Dans ce cas-là, la combinaison des rôles d'usager et de chercheur est nécessaire. Côté linguistique et socioculturel, une étude comparative sur le plan linguistique et discursif suppose au-delà des connaissances une sorte de double appartenance. Pour notre étude, notre expérience de locutrice bilingue du chinois et du français ainsi que notre pratique d'usager des microblogs et des nouvelles créations langagières sont déterminantes mais peuvent aussi engendrer des biais.

Les productions comme les pratiques, ainsi que la circulation des contenus au sein des RSN et vers d'autres espaces sont dynamiques et relationnelles. Ces spécificités ont-elles des influences sur les données discursives ? Quelles approches convient-il de mobiliser pour la collecte et l'analyse d'un corpus issu de ce contexte, notamment celui du microblog ? Ces questions méthodologiques seront abordées lors de la constitution du corpus dans le Chapitre IV.

## Conclusion

L'essor de l'Internet a bouleversé la communication interpersonnelle et son développement continue à influencer, former et transformer les modes de communications d'aujourd'hui. Web 2.0, Médias sociaux et Réseaux Sociaux Numériques, ces termes sont apparus en accompagnant le développement de cette infrastructure informatique. Les difficultés à trouver une définition précise et figée de ces termes et les confusions à leur sujet peuvent s'expliquer tout d'abord par les particularités des contenus de l'Internet et du Web qui sont à la fois dynamiques et éphémères, ensuite par les usages de plus en plus diversifiés, enfin par les interprétations qui se font sous des angles variés : marketing, Information et Communication, connaissances encyclopédiques, etc.

En dépit des difficultés à les définir, nous pouvons poser des distinctions claires de ces termes. Le service Web est une application concrète de l'Internet offrant une interface homme-machine. Le Web 2.0 est une étape de développement de l'Internet permettant un accès très simplifié. Les utilisateurs commencent à avoir un rôle actif et à participer à la production des contenus en ligne. L'interaction interpersonnelle devient pratique et instantanée. Tous ces changements se réalisent concrètement par les services des médias sociaux et les RSN. Ils constituent des outils de communication différents tant sur le plan des fonctionnalités que des pratiques. En outre, ils se distinguent de la notion de réseau social qui désigne un ensemble de relations entre des personnes.

Outre les aspects technologiques, nous pouvons constater que ces différents termes renvoient en réalité aux changements des manières de produire, de se procurer et de partager des informations ainsi qu'aux modes de mise en relation. L'Internet n'est pas invariable, et ses services ne cessent de se renouveler. Cela nous conduit à nous intéresser non seulement au contexte technologique mais aussi aux usagers et à leurs pratiques. Nous avons également constaté que l'engagement des chercheurs dans les activités d'observation active et participante est nécessaire. Notre hypothèse consistant à dire que contexte numérique, notamment le dispositif de microblog provoque des influences sur les productions langagières, il nous faut maintenant en venir à celles-ci que nous aborderons dans le chapitre suivant sous l'angle de la nouveauté et dans une perspective linguistique.



## CHAPITRE III

### Des néologismes aux nouveautés langagières en contexte numérique : définition de l'objet de recherche

#### Introduction

Une langue vivante évolue, change et se modifie progressivement au fil du temps. Ce sont les changements lexicaux qui sont particulièrement perceptibles : alors que les transformations des autres niveaux linguistiques échappent souvent aux locuteurs, l'apparition de mots nouveaux ou de sens nouveaux sont très visibles dans l'usage courant comme dans la littérature au sein de certains groupes professionnels ou générationnels par exemple. Ces changements trouvent aisément des explications. Pour désigner de nouvelles réalités et communiquer de nouvelles idées, de nouveaux mots et expressions sont créés spontanément. L'apparition du « verlan »<sup>75</sup> et de « l'argot d'Internet »<sup>76</sup> en sont des exemples dans la société d'aujourd'hui. Les langues différentes s'influencent mutuellement et les mots voyagent dans l'espace et dans le temps sans s'arrêter aux frontières des pays. Le phénomène de l'emprunt est bien connu et favorisé par les échanges, autrefois comme de nos jours. Notons qu'il devient très fréquent de rencontrer des anglicismes dans les langues française et chinoise (et parfois les mots anglais viennent lointainement du français...).

Avec le développement de technologies qui favorisent une communication internationale et numérique, on peut supposer que l'évolution des langues est plus accélérée que jamais et que

---

<sup>75</sup> Le verlan est un procédé d'inversion des syllabes. Il est associé à une pratique de dissimulation du sens pour ne pas être compris en dehors d'un groupe de connivence ou encore à une forme de langage jeune et éventuellement transgressif. Selon *l'Encyclopedia universalis*, il s'agit d'une « *Figure de diction* (Du Marsais) qui opère un changement phonétique ou graphique dans le mot par suppression, adjonction ou permutation. ». Dans l'acception élargie que fait le groupe Mu du terme *métaplasme*, c'est un métaplasme qui concerne la morphologie des mots. L'aphèrese (*las* pour *hélas*) est un métaplasme par suppression, la prosthèse par adjonction (*esquelette*) et le verlan par permutation (*béton* pour *tomber*). Le verlan est également rapproché des procédés de l'argot. Voir : <https://www.universalis.fr/encyclopedie/metaplasme-rhetorique/>; <https://www.universalis.fr/encyclopedie/registres-de-langue/>.

<sup>76</sup> Les listes de mots et d'expression qu'on peut trouver (par exemple dans Wikipédia) sous cette appellation regroupent un ensemble assez hétéroclite d'abréviations et d'acronymes ou encore de mots spécialisés et d'emprunts utilisés dans les communications en ligne et/ou dans les SMS.



les créations s'accompagnent d'une diffusion très large. Au niveau de la quantité comme du contenu, ces créations langagières s'enrichissent de manière inouïe. D'autre part, les changements les plus visibles ne se limitent plus au plan lexical, l'emploi des émoticônes et des signes typographiques inédits ainsi que de nouvelles structures apparaissent. Cette évolution s'accompagne en même temps d'une diffusion plus rapide que dans le passé où il fallait attendre des années pour que les nouveautés entrent dans l'usage quotidien. Aujourd'hui, elles peuvent être transmises très vite et employées alors par le grand public.

Notre problématique porte sur l'influence du contexte numérique sur la créativité lexicale et le présent chapitre est destiné à faire le point sur les notions linguistiques les plus adaptées pour définir et nommer notre objet de recherche à partir des métadiscours savants spécialisés de différentes traditions. Nous allons tout d'abord nous intéresser à la néologie et aux néologismes en retraçant l'évolution des termes et en évoquant les typologies de ces productions lexicales. Ensuite nous tenterons de faire un état de l'art comparé entre chercheurs chinois d'un côté et francophones et anglophones de l'autre dans un champ de recherche spécifique à la créativité sur le Web mais abordé par différentes disciplines avec des angles variés. Nous terminerons en passant en revue plusieurs possibilités pour dénommer notre objet de recherche et en justifiant une dénomination personnelle, celle de nouveauté langagière en contexte numérique.

### **3.1. Les métadiscours savants en question**

Dans ce chapitre nous cherchons tout d'abord à résoudre la question de la désignation des phénomènes qui passe par une réflexion sur les métatermes en usage en linguistique, dans la tradition française, mais aussi chez les chercheurs chinois, c'est-à-dire que notre comparaison porte ici sur les approches théoriques et les concepts descriptifs auxquels nous pouvons avoir accès dans les deux langues. Nous allons examiner tout d'abord quelques définitions de dictionnaires français avant d'en venir à des cadres théoriques qui expliquent les phénomènes néologiques dans le droit fil de la linguistique structurale.

#### **3.1.1. Les notions de *néologie* et *néologisme* dans les dictionnaires français**

Nous commençons par nous intéresser aux définitions lexicographiques fournies par des dictionnaires de langue usuels contemporains avant de nous appuyer sur des sources

spécialisées. Le dictionnaire *Le petit Robert* (édition 2012) définit respectivement les termes « néologisme » et « néologie » comme suit:

**Néologisme**, n.m. - 1735 de néo-et-logisme **1.** VIEILLI ET PEJ. Affectation de nouveauté dans la manière de s'exprimer. **2.**(1800) MOD. Emploi d'un mot nouveau (soit créé, soit obtenu par dérivation, composition, troncation, siglaison, emprunt, etc. : néologisme de forme) ou emploi d'un mot, d'une expression préexistants dans un sens nouveau (néologisme de sens). **3.** Mot nouveau ; sens nouveau d'un mot. **MED.** Mot forgé par un malade, incompréhensible pour l'entourage.

**Néologie** n.m – 1759 de néo-et-logie **1.** Création de mots nouveaux dans une langue, afin de l'enrichir. **2** (milieu XXe) LING. Processus par lesquels le lexique d'une langue s'enrichit, soit par la dérivation et la composition, soit par emprunts, calques, ou par tout autre moyen (sigle, acronymes...).

Ces définitions nous permettent de faire plusieurs constats. Tout d'abord, même si les deux termes sont associés aux phénomènes d'accroissement lexical dans la langue, le sens des termes n'est pas stable dans le temps et c'est à travers cette évolution historique que des convergences ou divergences se manifestent dans leur définition. Lorsqu'ils apparaissent en français ces mots désignent deux phénomènes langagiers différents ou plutôt jugés différemment. « Néologisme » qui est décrit comme une affectation de nouveauté dans la langue comporte une connotation négative. Alors que « néologie » est considéré comme une création qui enrichit la langue, bénéficiant d'une appréciation favorable. Par rapport à « néologie » dont l'évolution de sens reste moins saillante, le terme « néologisme » a acquis une valeur plus méliorative qui l'associe à l'emploi des différentes manifestations de nouveautés lexicales sur le plan morphologique et sémantique.

Le dictionnaire du CNRTL<sup>77</sup> fait la remarque suivante en précisant les repères historiques : « au 18<sup>e</sup> siècle, néologie et néologisme étaient tout d'abord employés de façon péjorative. Au début du 19<sup>e</sup> siècle, il s'est opéré une différenciation sémantique entre les deux paronymes ». Ce dictionnaire cite une définition du 19<sup>e</sup> siècle donnée par Mercier (1801: 28)<sup>78</sup> à l'appui de ce constat : « La néologie est l'art de former des mots nouveaux pour des idées ou nouvelles ou mal rendues. Le néologisme est la manie d'employer des mots nouveaux sans besoin ou sans goût. La néologie a ses règles ; le néologisme n'a pour guide qu'un vain

---

<sup>77</sup> Disponible sur : <https://www.cnrtl.fr/etymologie/neologisme> (consulté le 28 février 2016).

<sup>78</sup> Il s'agit d'un dictionnaire élaboré par Mercier, intitulé : *Néologie ou Vocabulaire de mots nouveaux à renouveler, ou pris dans des acceptions nouvelles*. Dans ce dictionnaire, Mercier a établi une opposition entre néologie et néologisme pour justifier son choix du titre de son dictionnaire.

caprice. »

Cette opposition entre néologie et néologisme formulée par Mercier est soulignée aussi bien par le linguiste français Louis Guilbert et par le chercheur chinois Zhenhua Xu, avec la citation suivante : « Néologie se prend toujours en bonne part, et néologisme en mauvaise, il y a entre ces deux mots la même différence qu'entre religion et fanatisme, philosophie et philosophisme » (Guilbert 1973:9) ; (Xu 2001:13).

Alors que les deux termes sont péjoratifs au 18<sup>e</sup> siècle, une distinction intervient entre les deux pour rendre compte d'un même phénomène (la nouveauté) : la notion de *néologisme* est marquée négativement (manque d'utilité, manque de règles, manque de goût) alors que celle de *néologie* est positive (nouveauté répondant à un besoin). L'évolution de sens et la relation de ces deux termes connaît une nouvelle étape au 20<sup>e</sup> siècle, allant vers une distinction entre le processus et les résultats de ce processus. En effet selon *Le petit Robert*, à partir du milieu du 20<sup>e</sup> siècle *néologie* désigne un processus d'enrichissement du lexique d'une langue. Quant au *néologisme*, il est abordé sous l'angle de deux procédés principaux : mot nouveau et sens nouveau d'un mot.

Notre consultation des définitions dans quelques dictionnaires contemporains nous conduit à faire les remarques suivantes :

1. Une évolution de sens est constatée qui permet de distinguer trois périodes décisives (18<sup>e</sup>, 19<sup>e</sup> et 20<sup>e</sup> siècles) où les nouvelles acceptions de ces deux mots se sont formées. Un article de Steuckardt (2016) retrace cette évolution historique très précisément en renvoyant à une abondante bibliographie. L'auteur évoque les prises de position dominantes envers les nouveaux phénomènes langagiers de l'Antiquité à nos jours, et nous pouvons ainsi comprendre que c'est à travers les changements d'opinion sur ces phénomènes, – et sur la conception de la langue et de ses normes –, que les deux mots eux-mêmes et leurs sens se sont construits. Concernant l'origine de la valeur péjorative, nous relevons la citation suivante :

C'est sur le mode de la plaisanterie qu'est imaginé, par le poète Jean-Baptiste Rousseau, le substantif *néologue*, puis, sous la plume acerbe de l'abbé Desfontaines, l'adjectif *néologique*. Il s'agit, en ce début du XVIII<sup>e</sup> siècle marqué par l'hostilité classique à l'innovation lexicale, de se moquer des trop audacieux inventeurs de mots nouveaux (Steuckardt 2016:34).

2. Les dictionnaires mentionnés ci-dessus n'ont pas présenté les acceptions d'un point de vue synchronique et n'ont pas non plus retenu les mêmes moments de l'évolution. Cette hétérogénéité dans les définitions est soulignée par Sablayrolles (2000). À partir d'une étude comparative entre vingt dictionnaires et travaux encyclopédiques dont la date de publication s'échelonne sur environ deux siècles (de 1751 à 1995), Sablayrolles, un des grands spécialistes français de la néologie a récapitulé cinq notions<sup>79</sup> qui fournissent des nouvelles acceptions dans les définitions de « néologie » et « néologisme », montrant ainsi la complexité de l'histoire de ces deux mots ainsi que la polysémie qui les caractérise. Cette analyse peut également nous aider à comprendre la relation entre les deux termes tout au long de ces évolutions historiques. Nous reprenons ci-dessous (Tableau 12) un tableau synthétique (Sablayrolles 2000:67) :

1/ opposition au 18 <sup>e</sup> siècle	
Néologisme « abus »	Néologie « usage normal »
Puis, par disparition de la valeur péjorative	
2/ synonymie	
Néologisme « emploi normal » (sens donné comme vieilli)	Néologie « emploi normal » (encore et déjà sens actuel)
Puis, par différenciation de deux synonymes	
3/ nouvelle différence	
Néologisme « mot nouveau » (sens actuel)	Néologie « emploi normal » (encore, et déjà sens actuel)
Puis, polysémie par nouveaux sens ajoutés à cause de besoins terminologiques	
4/ état actuel	
Néologisme « mot nouveau » + « terme formé par un malade » (terme récent, de divulgation restreinte)	Néologie « création/emploi normal de mots nouveaux » + « créativité lexicale » (terme récent et spécialisé)

**Tableau 12 : Evolution des relations sémantiques entre néologie et néologisme selon Sablayrolles (2000)**

Ce tableau fait bien apparaître que la définition de la néologie reste relativement stable, alors que celle du néologisme connaît des fluctuations. En fait, « néologie » et « néologisme » sont

<sup>79</sup> « 1. Abus de l'emploi de mots nouveaux, 2. Emploi de mots nouveaux, 3. Mot nouveau, 4. Processus d'enrichissement du lexique, créativité lexicale, 5. Création d'un malade en état délirant » (Sablayrolles 2000:57)

eux-mêmes des mots nouveaux (avec création formelle), construits sur des racines grecques savantes pour dénommer un phénomène qui existait bien entendu auparavant, mais n'était pas conceptualisé comme tel. Le processus d'évolution de leurs définitions illustre le néologisme de sens c'est-à-dire le fait que les acceptions d'un mot changent. Autrement dit l'apparition des concepts à travers des mots nouveaux et ensuite les sens nouveaux en font de bons exemples de ce qu'ils dénomment.

Malgré les diverses acceptions, la relation actuelle évoquée dans le tableau entre « néologie » et « néologisme » peut nous aider à éclaircir la principale distinction entre ces deux notions : la néologie concerne plutôt le processus et la manière dont les nouveaux mots et expressions sont formés, alors que le néologisme met l'accent sur les nouveaux mots eux-mêmes, et peut être considéré comme la concrétisation de la néologie.

Nous avons poursuivi notre enquête du côté des dictionnaires chinois pour voir si nous pouvions remonter à partir de là à une conceptualisation originale ou à des parentés.

### 3.1.2. Ressources chinoises

L'étymologie du terme « néologie » est simple : il est composé de deux mots d'origine grecque, *néos* qui signifie « nouveau » et *logos* qui signifie « parole, discours, savoir » et qui est devenu en français un suffixe courant (*logie* qui se rapporte à une « étude », une « doctrine »). Le suffixe *-isme*, d'origine grecque également, est très productif en français<sup>80</sup> et souvent porteur d'un sens d'appartenance idéologique, de référence à un système mais aussi d'exercice d'une activité, d'une « manière de faire ». Nous avons essayé de trouver l'expression équivalente dans la langue chinoise courante en nous appuyant sur un dictionnaire bilingue français-chinois. Selon le dictionnaire Larousse bilingue en ligne<sup>81</sup>, le *néologisme* est traduit comme 新/词<sup>82</sup> ; 新/词义 (xīn cí ; xīn cí yì). 新 (xīn) signifie « nouveau, ce qui vient d'apparaître » et 词 (cí) est « la plus petite unité porteuse de sens dans la langue chinoise ». 词义 (cí yì) signifiant « le sens de l'unité ».

---

<sup>80</sup> Voir : <http://www.academie-francaise.fr/construction-en-isme>; <https://www.cnrtl.fr/definition/-isme>.

<sup>81</sup> Disponible sur : [Dictionnaire Français-Chinois en ligne - Larousse](#).

<sup>82</sup> Nous employons la barre oblique dans cette étude pour marquer les unités de sens dans les expressions chinoises.

Il est intéressant de constater premièrement que « néologisme » est associé à la notion de 词 (cí) et non 字 (zì, le caractère) dans la traduction. Nous avons déjà mentionné que l'écriture chinoise traditionnelle est réalisée par des caractères monosyllabiques et morphologiquement interchangeables. 字 (zì, le caractère) renvoie donc à la réalisation graphique des « monèmes » chinois, unités minimales associant une forme et un sens. Mais il est souvent considéré comment l'équivalent de « mot » dans l'usage quotidien. Nous avons déjà indiqué *supra* que 词 est considéré comme une unité minimale associée à un sens qui peut être utilisée de manière autonome en chinois. La différence principale entre 字 et 词 repose sur l'autonomie d'utilisation. 字 est en effet un élément constitutif du 词. Selon le nombre de caractères, 词 comporte deux catégories : une forme simple qui se compose d'un seul caractère (ex. : 猫 māo, « chat ») ; une forme composée qui regroupe au minimum deux caractères (ex. : 学生 xué shēng, « étudiant » ; 电影院 diàn yǐng yuàn, « cinéma »). Plusieurs 词 peuvent se combiner pour construire des groupes et former des expressions et locutions. Dans ce cas-là, dans la langue chinoise, 新词 (qu'on traduira littéralement « nouvelle unité ») consiste la plupart du temps en une nouvelle combinaison structurelle de caractères s'il s'agit de formes composées. En outre, il ne faut pas oublier le phénomène de création de nouveaux caractères, ou plutôt de « 词 » simples, comme par exemple 婊 (biǎo, « ne le faites pas ») nouvellement créé à partir des caractères 不 et 婊 ; ou encore 蹿 (cuān, « en un clin d'œil ») qui est la combinaison des caractères 成 et 龙.

Après avoir précisé la notion de 词 (cí) en chinois, nous pouvons ainsi constater que la traduction chinoise de « néologisme » que nous avons citée transpose deux dimensions de nouveauté tout en prenant compte les caractéristiques linguistiques du chinois : la nouveauté formelle et la nouveauté sémantique. Ces constats nous amènent ensuite à nous demander si le terme 新词 (xīn cí) est bien une expression qui existe en chinois et non une traduction descriptive du phénomène. Pour cela, nous nous sommes appuyée sur plusieurs ressources englobant dictionnaire chinois, manuel de linguistique et certains travaux scientifiques chinois.

Nous avons tout d'abord consulté le dictionnaire *Xinhua* (version en ligne)<sup>83</sup>, ouvrage de référence important en Chine. La présence du terme 新词 (xīn cí) dans le dictionnaire est vérifiée et nous avons repéré ses définitions<sup>84</sup> que nous traduisons comme suivant :

1. une poésie chinoise nouvellement composée ;
2. des mots ou expressions nouvellement créés ;
3. des termes qui ne sont pas encore appris.

Dans le premier cas, 词 désigne un genre littéraire chinois très ancien qui date de la Dynastie du Sud et du Nord (420-589 apr. J.-C.). Il s'agit d'un poème chanté, écrit sur une mélodie préexistante. Dans le deuxième cas, 新词 désigne bien les nouvelles créations verbales. Dans le troisième cas, un aspect de nouveauté apparaît bien, mais associé à un processus d'apprentissage et non à un résultat de création. Donc le recours au dictionnaire montre que le terme 新词 est enregistré dans l'usage partagé pour décrire les nouveaux phénomènes langagiers.

Dans le manuel universitaire 现代汉语 (Yang *et al.* 2010:242) qui sert de référence pour les cours de linguistique chinoise, nous avons repéré l'expression 新词/新语 (xīn cí / xīn yǔ, *litt.* « nouvelle unité / nouveau langage ») qui est employée pour désigner les nouveautés langagières apparues dans le chinois contemporain depuis 1978<sup>85</sup>. Toujours selon ce manuel, ces nouveautés englobent des nouvelles créations lexicales, des sens nouveaux, l'utilisation élargie de certains dialectes, les emprunts, les formes mixtes (combinaison de caractères chinois et/ou de lettres étrangères et/ou de chiffres, par exemple : T 恤, 985 高校<sup>86</sup>).

---

<sup>83</sup> «新华字典» (xīn huá zì diǎn), disponible sur: <http://xh.5156edu.com/html5/72270.html>. Il fut le premier dictionnaire du chinois contemporain publié après la fondation de la République populaire de Chine. Il occupe une place importante en Chine avec une grande diffusion. En avril 2016, le Guinness World Records avait reconnu ce dictionnaire comme le « dictionnaire le plus populaire » et un « best-seller régulièrement mis à jour ».

<sup>84</sup> « 1.新作的诗词, 2.新产生的词语, 3.没有学过的词语».

<sup>85</sup> 1978 est l'année où a commencé le programme 改革开放 (gǎi gé kāi fàng, « réforme et d'ouverture ») en Chine. Cela a amené des changements sociaux, économiques, culturels, etc. Nous supposons que ces changements ont beaucoup favorisé le contact des langues aux niveaux national et international. C'est peut-être pour cette raison que cette année est marquée dans le manuel.

<sup>86</sup> T 恤 (xù) : T-shirt ; 985 高校(gāo xiào) : les universités qui participent au programme 985 destiné à construire des universités de premier rang et de niveau international.

Nous avons aussi vérifié que le terme 新词 (xīn cí) est adopté par Shuxin Liu, un des grands linguistes contemporains en Chine dans son livre « 汉语描写词汇学 » consacré à la lexicologie chinoise. Il donne la définition suivante<sup>87</sup> : 新词 est constitué par certaine (s) combinaison (s) phonétiques associée (s) à un nouveau concept et un nouveau sens, ou une nouvelle forme phonétique combinée avec un sens préexistant.<sup>88</sup> Cela rejoint à première vue la distinction des dictionnaires français entre néologisme formel et néologisme sémantique. Mais il insiste sur le fait que la forme est à comprendre comme phonétique et non comme affectant les caractères.

Nous avons complété ces informations par une recherche par mots clés<sup>89</sup> dans la base de données scientifique chinoise CNKI et nous avons constaté que les termes comme 新词/新语 (xīn cí xīn yǔ, voir *supra*), 新/词语 (xīn cí yǔ, *litt.* « nouveaux / mots et expressions ») sont utilisés parallèlement au terme 新词 (xīn cí) par les autres chercheurs chinois. En plus, ces chercheurs ont également tenté de caractériser les nouveaux phénomènes langagiers constatés. Citons une de leurs définitions<sup>90</sup> : les nouveautés phonétiques, sémantiques ou grammaticales composent les trois critères de 新词 (xīn cí).

Ces ressources nous montrent qu'ils existent des expressions en chinois<sup>91</sup> qui servent à caractériser les nouveaux phénomènes langagiers sans qu'il soit possible de trouver un accord unanime sur l'appellation ou la catégorisation de ces phénomènes. Nous l'avons constaté tant dans l'usage communément partagé que dans le contexte scientifique.

En nous intéressant aux métatermes existant en français et en chinois nous avons découvert que les notions de néologie/néologisme étaient plus difficiles à cerner que nous le pensions et qu'il fallait tenir compte de la variété des appellations selon les époques et les chercheurs, de la divergence des points de vue ainsi que de la complexité des phénomènes en relation avec les

---

<sup>87</sup> « 新词是以一定的语音形式与一个全新概念或意义相结合的或以一个新的语音形式与一个旧概念或意义相结合的在语言运用中具有稳定性和复呈性的语言单位 » (Liu 2005:33).

<sup>88</sup> Notre traduction.

<sup>89</sup> Nous avons entré le terme « 新词 » (xīn cí) dans le moteur de recherche de la base CNKI pour chercher les études dont les titres et mots clés associés contiennent ce terme.

<sup>90</sup> « 如果我们综合上述新词定义中的“意义和用法是新的”和“一个新的语音形式”来看, 一个词语只要在语音、语义或语法特点中有一个方面是新出现的, 我们就判定为新词 » (Sun 2015:17). Notre traduction.

<sup>91</sup> Faute de références qui retracent l'étymologie, nous ne pouvons toutefois pas établir s'il s'agit d'expressions d'origine chinoise ou de termes ayant subi l'influences des notions occidentales.



systemes linguistiques consideres. Toutefois, nous n'avons pas constate de disparite majeure d'acceptions en passant des ressources en francais aux ressources chinoises. Nous cherchons maintenant a preciser le cadre conceptuel sur la base de la tradition de la linguistique structurale.

### 3.1.3. Cadre conceptuel pour l'interpretation des neologismes et de leur formation

#### *La linguistique structurale pour comprendre la creation lexicale (Saussure, Martinet, Guilbert)*

Pour eclairer le phenomene de la neologie, nous reprenons la celebre distinction langue-parole due a l'un des fondateurs de la linguistique contemporaine. Selon Saussure :

L'etude du langage comporte deux parties : l'une, essentielle, a pour objet la langue, qui est sociale dans son essence et independante de l'individu ; cette etude est uniquement psychique ; l'autre, secondaire, a pour objet la partie individuelle du langage, c'est-a-dire la parole y compris la phonation : elle est psycho-physique (Saussure 1972:37).

Il ressort de cette opposition langue/parole que la langue, en tant que phenomene social, est collective, et que la parole est l'acte d'expression d'une pensee individuelle dans une situation concrete. De ce point de vue, la creation des nouveaux mots et des sens nouveaux est tout d'abord un acte individuel. Comme le dit Louis Guilbert<sup>92</sup>, un eminent specialiste de lexicologie et de terminologie scientifique et technique : « c'est donc le sujet parlant qui cree le neologisme » (1972:13). Mais il ajoute ensuite que la formation du neologisme n'est pas seulement un acte de parole, et qu'elle est destinee a etre aussi un phenomene de langue dans la mesure ou les sujets parlants, en tant que membres d'une communaute creent les neologismes dans le but d'enrichir la communication. On peut donc conclure que la creation neologique individuelle se realise essentiellement dans la communication et presente necessairement un aspect social.

Un autre apport de Saussure est sa conception originale du signe linguistique qui comporte deux

---

<sup>92</sup> Louis Guilbert (1912-1977) a ete le directeur de publication du Grand Larousse de la langue francaise. Ses etudes sont centrees sur la mise en rapport la creation lexicale et l'evolution du vocabulaire avec les inventions techniques. Il animait une equipe de recherche dans les annees 70 avec pour objectif d'integrer dans la description linguistique le phenomene neologique. C'est dire qu'il est particulierement bien place pour cerner l'ampleur du phenomene neologique dans le francais d'aujourd'hui. Voir :

<https://maitron.fr/spip.php?article75506>;

[https://www.lemonde.fr/archives/article/1977/06/20/louis-guilbert\\_2868220\\_1819218.html](https://www.lemonde.fr/archives/article/1977/06/20/louis-guilbert_2868220_1819218.html).

faces indissociables, le signifiant et le signifié. Tout néologisme comprend inévitablement ces deux composants inséparables. A cet égard, l'opposition entre le signifiant et signifié constitue une base fondamentale de l'analyse du néologisme. Le néologisme de sens associerait un nouveau signifié à un signifiant existant, alors que le néologisme de forme serait un signifiant nouveau avec un signifié nouveau.

S'inspirant du structuralisme issu des travaux de Saussure et de la définition de la langue comme système de signes, André Martinet développe une théorie originale et désormais bien connue, celle de la double articulation du langage. Selon lui :

Une langue est un instrument de communication selon lequel l'expérience humaine s'analyse, différemment dans chaque communauté, en unités douées d'un contenu sémantique et d'une expression phonique, les monèmes ; cette expression phonique s'articule à son tour en unités distinctives et successives, les phonèmes, en nombre déterminé dans chaque langue, dont la nature et les rapports mutuels diffèrent eux aussi d'une langue à une autre (Martinet 2008:44).

Ainsi, la théorie de Martinet permet d'expliquer l'apparition des néologismes en les situant sur une échelle de liberté croissante dans la langue selon les unités : très limitée pour les phonèmes, infinie pour les phrases. D'après lui : « le nombre des énoncés possibles dans chaque langue est théoriquement infini, car il n'est pas de limite au nombre de monèmes successifs qu'un énoncé peut comporter » (*Ibid*:43). Cette liberté de combinaison des monèmes<sup>93</sup> au sein des phrases en respectant les structures syntaxiques de la langue concernée régit aussi la production des monèmes lexicaux. Il considère que ces derniers appartiennent à des séries ouvertes parce qu'ils doivent pouvoir être fabriqués au fur et à mesure des besoins qui sont constants, illimités, c'est-à-dire que, dans toute communauté, de nouveaux besoins se manifestent sans cesse et ces besoins font naître de nouvelles désignations. La création néologique concerne donc essentiellement les monèmes lexicaux tout en mobilisant des monèmes qui se grammaticalisent (comme les suffixes).

Il est bien connu que le lexique est la composante la plus sensible et active dans les changements

---

<sup>93</sup> D'après Martinet, qui souligne l'ambiguïté de la notion de mot, usuelle mais peu opérationnelle en linguistique, le monème est une unité minimale significative (associant une forme et un sens) dite de première articulation alors que le phonème est une unité minimale non significative et distinctive dite de deuxième articulation qui permet de constituer et de différencier les monèmes. Les monèmes grammaticaux en nombre limité, caractérisés par une amplitude sémantique réduite et une forte stabilité (par exemple les désinences verbales) sont distingués des monèmes lexicaux.

linguistiques mais pour autant il n'y a pas de consensus chez les linguistes sur l'unité de la création lexicale. Pour Guilbert, l'unité de la création lexicale n'est pas la plus petite unité de la signification de la langue (monème) mais c'est le mot qui sert de point de départ à l'analyse linguistique. Alors que Sablayrolles (2020) reprenant Pottier considère que l'unité pertinente du néologisme n'est ni le monème, ni le mot mais la lexie, c'est-à-dire une unité lexicale qui peut être simple (un mot) ou composée (plusieurs mots associés) jusqu'à la locution complexe mais qui se comporte fonctionnellement comme une unité simple comme le montre la possibilité de permutation avec une telle unité simple ; la lexie a la propriété d'être mémorisée en langue et ménage un continuum entre mot et énoncé<sup>94</sup>.

Nous avons rencontré cette question de délimitation de l'unité du néologisme en chinois où, on emploie plutôt 新词 (la plus petite unité porteuse du sens dans la langue chinoise) au lieu de 新字 (caractère ou nouveau mot).

### *Une comparaison des postulats de recherche (Guilbert et Xu)*

Nous avons voulu comparer les postulats de recherche d'un linguiste français héritier du structuralisme Louis Guilbert et d'un chercheur chinois en Sciences du langage Zhenhua Xu qui associe à la linguistique la question des dimensions sociales particulièrement pour aborder les néologismes<sup>95</sup>.

Nous allons citer assez longuement Guilbert qui dresse une sorte de bilan théorique en situant les néologismes comme des créations des locuteurs en synchronie dans son ouvrage *Théorie du néologisme* (1972:18) :

1. Une langue fonctionne selon son propre code en vertu duquel sont produits des actes de discours et des formations lexicales. Tout ce qui provient d'une langue autre doit être considéré comme relevant d'un autre code.
2. Le néologisme est un signe linguistique comportant une face « signifiant » et une face « signifié ». Ces deux composantes sont modifiées conjointement dans la création néologique, même si la mutation semble porter sur la seule morphologie du terme ou sur sa seule signification.

---

<sup>94</sup> <https://www.cnrtl.fr/definition/lexie>.

<sup>95</sup> Il est notamment l'auteur d'un ouvrage en français *Le néologisme et ses implications sociales* (2001) auquel nous renvoyons.

3. La formation néologique, le plus souvent, n'est pas une unité de signification minimale. Elle résulte de la combinaison d'éléments plus simples existant dans la langue. La création réside alors dans le mode de relation établie entre les éléments.
4. La création du néologisme ne peut être dissociée du discours tenu par le créateur-individu intégré à une communauté, s'exprimant dans une situation donnée.
5. Le néologisme présente un aspect oral et un aspect écrit. Les modifications graphiques doivent donc être considérées comme relevant de la néologie.

Ces postulats de Guilbert permettent de baliser différentes questions que nous allons trouver quand il s'agira de classer les néologismes : celle de l'emprunt (1) ; celle de la distinction à faire ou non (Guilbert la refuse) entre néologisme formel et néologisme sémantique (2) ; celle des conditions d'énonciation et du rôle du contexte communicationnel (4) ; celle de la réalisation écrite pertinente pour les discours diffusés en ligne qui exploitent la dimension graphique et iconique (5).

Nous citons également les postulats principaux de Xu (2001:59-60) :

1. Les créations et renouvellements lexicaux sont apparus d'abord comme une manifestation d'énonciation, le lexique étant une partie dépendante de la phrase, c'est cette dernière qui assure la fonction lexicale d'un acte de création[...]
2. Le système des signes est productif, et ces signes s'ouvrent à toutes les virtualités d'enrichissement et de changement, répondant aux besoins de la société ou à l'inspiration créatrice de l'individu[...]
3. Comme la langue n'est pas une structure abstraite homogène, on doit prendre en considération à la fois sa pluralité fonctionnelle et des effets sociaux du langage lui-même[...]
4. L'acte de parole contribue au dynamisme de l'interaction sociale. L'étude de la nouvelle formation lexicale...il faudrait qu'elle s'oriente également sur l'analyse du langage en situation[...]
5. Dans la plupart des cas, la création lexicale consiste dans l'application d'un certain nombre de règles linguistiques ou extralinguistiques[...]
6. L'apparition d'un terme nouveau n'efface pas nécessairement l'usage ancien[...]

Ces points font une part importante à l'énonciation et aux situations d'énonciation (1, 4), à la part de l'extralinguistique dans la création linguistique elle-même (2, 3, 4, 5) et plus généralement aux relations entre langue, langage et société.

Les postulats adoptés par les deux chercheurs concernent le fonctionnement de la langue mais sous les angles différents. Ils ne nous paraissent pas incompatibles mais plutôt complémentaires. Certains postulats se croisent alors que d'autres se complètent les uns les autres. D'après eux, la création lexicale fait le lien entre l'activité individuelle des sujets parlants d'une communauté donnée et la dimension collective de la langue. Il nous semble que Xu apporte un élément important qui correspond aux objectifs de notre recherche en suggérant que l'analyse de la

formation des néologismes doit être traitée non seulement sur le plan linguistique, mais aussi sur le plan social. La formation des néologismes n'est pas un acte purement linguistique. Elle est à la fois un phénomène langagier et social.

#### **3.1.4. Les typologies des néologismes**

Les dictionnaires usuels comme *Le petit Robert* distinguent comme nous l'avons vu *supra* le néologisme de forme et le néologisme de sens. Un classement très souvent mentionné (y compris dans les études chinoises) comporte, lui, trois catégories : les néologismes formels – créés à partir des ressources morphologiques de la langue –, les néologismes sémantiques – résultant de nouveaux sens attribués à des mots existants –, les emprunts – à un dialecte, une langue étrangère ou ancienne.

Pour reprendre les deux chercheurs dont nous avons comparé les postulats, Guilbert classe autrement les néologismes en distinguant les sous-systèmes linguistiques. Il distingue la néologie phonologique, la néologie syntaxique, la néologie sémantique, l'emprunt et la néologie graphique. Xu préfère au classement de Guilbert celui de Pierre Guiraud<sup>96</sup> en quatre types : 1. Type onomatopéique 2. Type morphologique 3. Type sémantique 4. Type allogénique (ce dernier correspondant aux emprunts) (cité dans Xu 2001:61). Nous voyons à partir de là que la complexité de la notion de « néologisme » se retrouve quand il s'agit de dégager des critères de classement.

Ce sous-chapitre tente de synthétiser les typologies des néologismes existantes issues des travaux de certains chercheurs et n'a pas l'ambition de construire une nouvelle classification. Nous cherchons une typologie qui nous serve de référence pour nos analyses ultérieures.

#### ***Une grande diversité de classements***

Les définitions des dictionnaires que nous avons mentionnées au début du chapitre nous ont permis d'apercevoir dans un premier temps deux grands types de nouveautés manifestées, celles de forme et celles de sens. L'ajout ou non de l'emprunt comme constituant un troisième type

---

<sup>96</sup> Ce linguiste français (1912-1983) de grand renom a notamment travaillé dans le domaine de la stylistique, du lexique (sur l'argot par exemple) et est considéré comme un précurseur de la statistique textuelle.

fait souvent débat. L'emprunt est parfois exclu des procédés de néologie ou considéré comme une sous-catégorie de néologismes morphologiques et/ou sémantiques et non comme une catégorie à part entière.

Si nous admettons que l'emprunt relève d'un troisième type, nous pouvons déjà constater de premières divergences sur les critères de classement. Par exemple, les néologismes formels et sémantiques sont distingués sur la base de leurs procédés de création alors que les emprunts sont abordés à partir de leurs origines.

En fait, les typologies de néologismes sont très nombreuses. Sablayrolles (1996:26-28) a recensé et discuté les principales et les répartit en deux ensembles.

1) Dans le premier ensemble la trichotomie (néologisme formel / néologisme sémantique / emprunt) que nous avons évoquée *supra* est enrichie par d'autres distinctions. Sablayrolles en énumère 12 différentes et représentatives, parmi lesquelles nous pouvons retrouver celles de Guilbert ou celle de Tournier qui fait des néologismes purement morphologiques une quatrième catégorie sur le même plan que celle de la trichotomie, nous en reparlerons plus loin.

2) Dans l'autre ensemble, la trichotomie n'est pas forcément reprise et les propositions de typologie sont bien différentes les unes des autres. Par exemple, menant une analyse morphologique du français parlé des jeunes Parisiens dans les années 1980, Walter (1984, cité dans Sablayrolles 1996:45) intègre le verlan dans sa typologie avec sur le même plan « nouveaux signifiants » et « nouveaux signifiés ». Verdelhan-Bourgade (1990, cité dans Sablayrolles 1996:28) propose la typologie suivante : changement de sens, créativité lexicale et phénomènes syntaxiques dans son étude du français « branché ».

Dans ses travaux récents (Sablayrolles 2019:90-95), à travers une grille de comparaison des diverses typologies, Sablayrolles évoque des critères de classement plus précis et développés, au nombre de 8 : les domaines ; les créateurs et les modes de création ; l'origine des formants ; la structure apparente ; les causes de la néologie ; la sémantique ; les grands principes linguistiques<sup>97</sup> ; les procédés de formation. Il est à noter que les procédés de formation

---

<sup>97</sup> Sablayrolles cite des principes linguistiques comme le transfert ou la coalescence de Henri Bonnard ; l'élargissement ou la fixation d'une base de Robert-Léon Wagner.

constituent un des critères de classement le plus choisi par les chercheurs selon les études menées par Sablayrolles.

En comparant ces typologies, Sablayrolles formule des remarques auxquelles nous souscrivons : les points de vue adoptés dans l'étude conduisent à l'adoption de critères différents ; les corpus d'analyse constitués sont également un élément phare qui détermine les critères. Il exprime une critique à l'égard de la plupart des typologies : elles relèvent selon lui du « catalogue », c'est-à-dire un regroupement de cas étudiés sans hiérarchie et sans analyse approfondie de leurs relations. Quant à lui, il propose à partir du modèle anglais de Jean Tournier, un tableau raisonné et hiérarchisé des néologismes. Il a choisi le modèle de Tournier comme base de travail pour les raisons suivantes : il s'agit d'un modèle assez récent parmi ceux qu'il a étudiés et ses préoccupations de recherche sur la productivité lexicale et l'unité lexicale du néologisme sont proches de celles de Tournier. Même s'il a été établi à l'origine pour l'anglais, ce classement permet de traiter un plus grand nombre de corpus collectés dans les études de Sablayrolles par rapport aux autres classements selon une étude comparative (Sablayrolles 1996 ; 2000). Ainsi, il se base sur les procédés de création en les faisant reposer sur deux sortes de matrices : les matrices<sup>98</sup> internes et les matrices externes.

### *Une typologie évolutive (Sablayrolles et Pruvost)*

Sablayrolles et Pruvost ont fait évoluer leurs propositions de typologie au fil de leurs travaux (Pruvost et Sablayrolles 2003 ; Sablayrolles 2019). Nous allons comparer trois tableaux différents correspondant aux phases successives de leur réflexion car cette présentation synthétique nous permet d'embrasser du regard l'organisation typologique afin d'examiner ensuite plus en détail telle ou telle catégorie, tel ou tel changement et aussi de déterminer la typologie à retenir pour nos analyses.

Le premier tableau présenté ci-dessous (Tableau 13) se trouve dans l'ouvrage de la collection « Que sais-je ? » portant sur les néologismes publié en 2012 (Pruvost et Sablayrolles 2012)<sup>99</sup>.

---

<sup>98</sup> Tournier (1985, cité dans Sablayrolles 1996:13) nomme la grille des procédés de formation : matrice lexicogénique.

<sup>99</sup> Il s'agit d'une 2<sup>e</sup> édition d'un ouvrage paru initialement en 2003 et qui est régulièrement mis à jour. Nous avons choisi de prendre en considération cette 2<sup>e</sup> édition et de la comparer à des tableaux plus récents

## LES PROCÉDÉS DE CRÉATION

Matrices internes	morpho-sémantiques	construction	affixation	préfixation ----- suffixation ----- dérivation inverse ----- parasynthétique
			flexion	flexion
			composition	composition synapsie quasi-morphème ----- mot-valise
		imitation et déformation		onomatopée fausse coupe jeu graphique paronymie
		syntactico-sémantiques	changement de fonction	conversion ----- combinatoire syntaxique/lexicale
	changement de sens		métaphore ----- métonymie ----- autres figures	
	morphologiques	réduction de la forme	troncation ----- siglaison	
	sémantico-pragmatique		détournement	
	matrice externe			emprunt

**Tableau 13 : Les procédés de création (Pruvost et Sablayrolles 2012:117)**

Un classement pouvant aller jusqu'à cinq niveaux principaux est présenté avec des sous-types échelonnés, en allant du général au particulier. Les matrices internes réparties en quatre rubriques correspondant chacune à des niveaux linguistiques mixtes (morpho-sémantique, syntactico-sémantique, sémantico-pragmatique) ou non (morphologique) s'opposent à une unique matrice externe horizontale associée à un procédé unique, l'emprunt. Les rubriques des matrices internes sont à leur tour subdivisées en opérations formelles, ou sémantiques et formelles, qui, elles-mêmes s'actualisent en différents procédés au nom plus ou moins familier. Par exemple, l'opération de changement de sens dans la rubrique du niveau syntactico-sémantique peut déboucher sur une métaphore, une métonymie ou d'autres figures ; l'opération de construction dans la rubrique du niveau morphosémantique se décompose en trois catégories (affixation, flexion, composition), et, par exemple, parmi les procédés de la troisième catégorie (composition) on trouve le mot-valise.



Les apports d'un classement précis et hiérarchisé permettent non seulement de prendre en considération la diversité des données rencontrées, mais aussi de distinguer les données d'une manière homogène. Dans l'ensemble, les cinq niveaux distingués en font la typologie la plus hiérarchisée par rapport à une centaine de classements relevés dans les études comparatives (voir *supra*). A cet égard, à la différence des typologies qui regroupent sur le même plan de grands types de procédés, comme par exemple la trichotomie traditionnelle, le classement pourrait permettre d'entreprendre une analyse dans les détails et de bien situer les néologismes dans chaque regroupement. Mais il est à noter que le nombre des niveaux hiérarchisés n'implique pas la quantité la plus élevée des procédés établis. A titre d'exemple, la typologie de Goose proposée en 1975 (cité dans Sablayrolles 1996:37-38) en compte une quarantaine sur trois niveaux. La disparité dans la quantité peut s'expliquer par des corpus différents.

Regardons maintenant de plus près les niveaux hiérarchisés. Tout d'abord, nous avons constaté au niveau supérieur une opposition fondamentale entre matrices internes et externe. Cette opposition, qui dégage un procédé de formation sous l'influence des éléments étrangers, peut apporter éventuellement une solution à la place délicate de l'emprunt qui fait souvent débat (voir *supra*), tout en gardant un critère homogène de classement sur le premier niveau. Les changements formels et sémantiques constatés dans les emprunts ne pourraient plus provoquer de confusions lors des analyses. Ensuite, au deuxième niveau les hiérarchisations essentielles sont manifestées par les procédés de création interne. Ces hiérarchisations sont basées dans un premier temps sur les principaux aspects linguistiques. Des procédés linguistiques mixtes sont ainsi proposés. Cela pourrait permettre d'analyser les néologismes d'un point de vue général, contrairement à certaines typologies restreintes pour des études spécialisées, comme celles de Walter et Verdelhan-Bourgade qui s'intéressent respectivement au verlan et au français branché. Les procédés mixtes impliquent également la prise en considération de la complexité dans le processus de formation. En soulignant les changements simultanés sur le plan formel et sémantique, les matrices morphosémantiques se distinguent des autres procédés qui ne marquent prioritairement qu'une face du signe<sup>100</sup>, comme le classement binaire néologisme formel ou néologisme sémantique. D'ailleurs, les précisions sur des faces affectés d'un signe,

---

<sup>100</sup> En référence aux faces signifié / signifiant de Saussure. Cela est aussi compatible avec ce qu'écrit Guilbert (1972) dans ses postulats, cité *supra* : « Ces deux composantes sont modifiées conjointement dans la création néologique, même si la mutation semble porter sur la seule morphologie du terme ou sur sa seule signification ».

permettent de mieux comprendre les différences entre les matrices internes : les matrices syntactico-sémantiques permettent de mieux cerner les cas de changement sur la face *signifié* ; les matrices purement morphologiquement modifient le *signifiant*. Plus les niveaux sont hiérarchisés, plus les différences sont visibles.

En résumé, l'appui sur un tableau hiérarchisé qui est basé sur les procédés de formation pourrait éventuellement aider à dissiper des confusions de classement non résolues dans d'autres typologies et mettre en évidence les caractéristiques différentes des néologismes.

Le tableau suivant (Tableau 14) datant de 2016 est issu de la nouvelle édition du même ouvrage, la troisième (Pruvost et Sablayrolles 2016).

LES PROCÉDÉS DE CRÉATION				
Matrices internes	morpho-sémantiques	construction	affixation	préfixation
				suffixation
				dérivation inverse
				parasynthétique ?
		flexion	flexion	
		composition	composition synapsie fractocomposition compoaction	
	mot-valise factorisation			
	imitation et déformation		onomatopée fausse coupe jeu graphique paronymie	
	syntactico-sémantiques	changement de fonction	conversion	
			combinatoire syntaxique/ lexicale	
		changement de sens	métaphore	
	métonymie			
autres figures				
morpho-logiques	réduction de la forme	troncation		
		siglaison		
phraséologique		création		
		détournement		
matrice externe		emprunt		

Tableau 14 : Les procédés de création (Pruvost et Sablayrolles 2016:115)

Les principaux changements résident dans l'extension d'une des matrices internes, celle qui concerne les procédés morphologiques. La matrice sémantico-pragmatique de 2012 disparaît en tant que rubrique et le procédé qui lui était associé (le détournement) se retrouve avec un nouveau procédé (la création) dans une nouvelle catégorie, la matrice phraséologique. Cette dernière, intégrée comme sous-ensemble de la matrice morphologique, concerne davantage « une lexie longue et complexe » ; deux procédés sont distingués : la création directe de longues expressions et le détournement créé à partir des expressions existantes. Les autres changements concernent les sous-types de « composition » inclus dans les matrices morpho-sémantiques. Ces changements apportent des précisions aux procédés de composition comme par exemple la distinction entre *mot-valise* et *factorisation*.

Le 3<sup>ème</sup> tableau (Tableau 15) que nous allons voir à présente est issu d'un ouvrage récent (Sablayrolles 2019). Il ne s'agit pas d'une réédition du livre de la collection « Que sais-je ? » mais d'un ouvrage visant à synthétiser les travaux, les questions et les solutions disponibles sur le sujet de néologie, à faire le point sur les avancées en ce domaine.

m a t r i c e s  i n t e r n e s	morpho- sémantiques	construction	affixation	préfixation	<i>détatouer, rediaboliser</i>	
				suffixation	<i>statuesque, tontoniser</i>	
				dérivation inverse	<i>dissider, prester</i>	
				parasynthétique	<i>désidéologisé</i>	
				flexion		<i>ils closirent, lesbien une représaille</i>
			compo- sition	composition « rég. »	<i>voiture-bélier</i>	
				synapsie	<i>lanceur d'alerte</i>	
				composition savante	<i>ptochophobie</i>	
				hybride	<i>nounoursothérapie</i>	
				compocation	<i>glamping, héliport</i>	
				fractocomposition	<i>télespectateur, xgate</i>	
				mot valise	<i>peopolitique, quinquado</i>	
				factorisation	<i>optimessimiste, hépathétique</i>	
		substitution	<i>monokini, trilogue</i>			
	imitation et déformation	onomatopée	<i>dzoing, whishsh</i>			
		fausse coupe ou paronymie	<i>la nesthésie, pestacle</i>			
		redoublement	<i>tuture (voiture)</i>			
		déf. syst. : verlan...	<i>reuch, loufoque</i>			
s y n t a c t i c o- s é m a n t i q u e s	syntactico- sémantiques	changement de fonction	conversion	<i>la gagne, bande-dessiner</i>		
			conversion verticale	<i>son ex, les chic ouf</i>		
			déflexivation	<i>le manger, le pleurer</i>		
			combinatoire syntax° / lexicale	<i>ça craint, encourir la liberté</i>		
		changement de sens	métaphore	<i>souris (inform.), bounty</i>		
			métonymie	<i>passoire thermique</i>		
			autres figures	<i>déficient auditif (sourd)</i>		
purement morpho- logiques	réduction de la forme	troncation	<i>ricain, blème, déj'...</i>			
		siglaison	<i>LMD</i>			
		acronyme	<i>ecue, pacs, perdre</i>			
phraséologiques		création d'expression	<i>faire du huit mégabits</i>			
		détournement	<i>faire marcher la planche à promesses</i>			
matrice externe		emprunt	<i>quantified self, fact checking, fioul, redingote</i>			

Tableau 15 : Tableau raisonné et hiérarchisé des matrices (Sablayrolles 2019:127)

Si nous comparons ce tableau aux deux précédents, nous pouvons observer dans un premier temps que l'unique matrice externe (l'emprunt) reste toujours inchangée et que d'autre part une colonne supplémentaire à droite est ajoutée, fournissant un ou plusieurs exemples en italiques par procédé pour pouvoir bien illustrer ces procédés de création. Les autres changements peuvent être considérés comme des mises à jour concernant les matrices internes. Nos constatations sont les suivantes :

- la matrice phraséologique devient indépendante et constitue une rubrique de même niveau que les autres matrices internes (les trois autres étant morpho-sémantiques, syntactico-sémantiques, purement morphologiques) ;
- l'ancienne matrice morphologique est réduite et renommée « purement morphologique » ;
- des ajouts sont effectués dans la liste des procédés qui se trouvent enrichis, surtout ceux relevant de l'opération de *composition* ;
- à l'inverse certains procédés sont exclus du tableau comme « *jeu graphique* » présent dans les deux précédents.

Ces tableaux présentent un récapitulatif des principaux types de procédés en les rattachant à des matrices caractérisées. Pour connaître plus d'information précises et mener à partir de là des études approfondies, la consultation des explications est nécessaire. Notre objectif n'est pas de discuter la pertinence de cette typologie mais au sein de différentes possibilités de classement, de présenter celle qui nous paraît particulièrement élaborée en raison de ses paliers successifs, et la plus dynamique puisqu'elle se modifie tout en gardant sa logique. Il faut souligner que l'élaboration des typologies hiérarchisées est basée sur la collecte continue des néologismes rencontrés, ce qui, selon Sablayrolles, est un élément d'enrichissement du classement. Les rectifications sont inévitables après un certain temps. Outre l'intégration des exemples qui amène une réflexion, le modèle peut être repensé par ses auteurs en vue de l'améliorer. Mais cela nous montre aussi que les frontières entre les procédés ne sont pas faciles à établir et ne le sont pas forcément une fois pour toutes.

### ***Deux cas en discussion***

Nous avons ainsi relevé dans un ouvrage très récent (Sablayrolles 2020), de nouveaux changements apportés par Sablayrolles dans les sous-ensembles de la typologie. Notre attention est plus particulièrement attirée par deux modifications qui nous semblent relativement liées aux nouveaux phénomènes langagiers qui nous intéressent<sup>101</sup>.

Il s'agit de deux cas d'exclusion motivés dans cette publication de 2020. Nous avons remarqué que *extension et restriction de sens* ainsi que *lexies phrastiques et textuelles* sont éliminés des néologismes. Ces deux cas ne sont pas affichés dans les tableaux mais présents dans les explications des tableaux. Le premier cas était originalement intégré dans les matrices morpho-

---

<sup>101</sup> Nous prendrons éventuellement en compte les autres modifications au fil de nos analyses concrètes si besoin.

sémantiques. Les explications indiquent que « le changement de sens néologique est un saut de signification brutale » (Sablayrolles 2020:24) alors que les évolutions de sens de l'ordre de l'extension et de la restriction sont progressives. La prise en compte du changement de sens évolutif n'est pas immédiate. Un autre argument souligne que la différence entre les cas exclus et les néologismes est que « aucun effet stylistique est présent » (*Id.*) dans le processus évolutif du sens. Nous nous demandons toutefois si quand le sens néologique en rupture (« saut de signification brutale ») est utilisé en parallèle du sens original, il ne peut pas être considéré comme une extension, notamment dans le cas de sens spécifiques appliqués à des nouvelles réalités. Si le sens évolutif est déterminé par un décalage entre sens émis par le locuteur et sens construit par le récepteur interprétant, comment peut-on classer les nouveaux sens qui entrent à une vitesse accélérée dans les usages quotidiens ? Il est également à noter que le sentiment néologique<sup>102</sup> est différent d'un individu à l'autre. La prise en compte non immédiate d'un néologisme est tout à fait possible.

La deuxième exclusion concerne les lexies phrastiques et textuelles. Tout d'abord, il est opportun de comprendre la notion de « lexie ». En se basant sur les travaux de Pottier et Tournier, Sablayrolles opte pour cette notion comme unité lexicale pertinente du néologisme. En effet un néologisme est classiquement défini comme « mot nouveau » et/ou « sens nouveau ». C'est pour questionner la pertinence de la notion de « mot » comme unité de signification minimale du néologisme que la notion de « lexie » (celle-ci pouvant être constituée d'un mot ou de plusieurs mots associés) s'est imposée dans la mesure où la notion de morphème ou monème pour reprendre le terme de Martinet ne convenait pas vraiment non plus.

Sans vouloir nous engager dans une discussion sur ces concepts qui ont donné lieu à de multiples travaux, il faut admettre que nous avons effectivement constaté des innovations lexicales hybrides ou longues qui dépassaient le cadre du « mot » ou du « morphème » comme nous aurons l'occasion de le voir plus en détail dans les chapitres d'analyse.

Rappelons maintenant les trois propriétés des « lexies » selon Sablayrolles (2000) : sur le plan sémantique, elles constituent des unités qui ont une stabilité référentielle ; ce sont des unités qui

---

<sup>102</sup> Il s'agit de l'intuition qu'ont les locuteurs du caractère nouveau de certains mots ou expressions. (Gardin *et al.* 1974).

ont une fonction syntaxique ; enfin les lexies sont des éléments mémorisés ou mémorisable. C'est alors à partir de ces trois aspects que l'exclusion est prononcée pour les innovations au niveau phrastique et textuel. Selon l'auteur, les phrases et textes, même s'ils sont mémorisables, ne sont pas des constituants de syntagmes ou de phrases mais sont sur le même niveau. Donc, ils ne sont pas considérés comme « lexies » et sont éliminés du cadre des néologismes.

Nous nous questionnons toutefois sur cette exclusion car de telles phrases peuvent connaître un figement qui leur confère bien un nouveau sens, voire un nouveau statut syntaxique et être intégrées ainsi dans les énoncés. Comme le souligne lui-même Sablayrolles « les locuteurs ne donnent pas l'impression de construire leurs énoncés à partir de la combinaison des plus petites unités porteuses de sens mais à partir d'unités sémantiques déjà constituées, plus vastes » (2000 : 147). Dès lors les rôles syntaxiques des innovations de niveau phrastique et textuel ne sont-ils pas influencés par les pratiques des locuteurs? Il peut arriver par exemple que ces innovations phrastiques se transforment au fil de leurs emplois en une structure raccourcie par exemple : la longue expression 我和我的小伙伴们都惊呆了 (wǒ hé wǒ de xiǎo huǒ bàn men dōu jīng dāi le, *litt.* « mes copains et moi nous sommes tous étonnés ») est en fait une phrase. La forme raccourcie 我伙呆 (wǒ huǒ dāi, *litt.* « moi copain(s) étonné») est ensuite créée et utilisée en parallèle de sa longue version comme un adjectif pour exprimer un sentiment de surprise. Dans ce cas-là, convient-il toujours de ne pas les traiter comme néologismes ?

Cette discussion amorce les développements que nous consacrerons dans nos corpus d'étude aux phénomènes langagiers en contexte numérique qui font l'objet de notre attention. C'est dans cette optique que nous avons cherché à adopter une typologie parmi la variété des classements et des critères possibles et que nous avons retenu les propositions de Pruvost et Sablayrolles et plus particulièrement celle de Sablayrolles 2019. Nos raisons sont les suivantes : les matrices « lexicogéniques » sont clairement définies et distinguées ; il s'agit d'un classement rigoureux, récent et dynamique.

Certes cette typologie est élaborée pour la langue française alors que nous aurons à traiter un corpus français et un corpus chinois. Précisons pourquoi les typologies chinoises ne font pas dans ce chapitre l'objet d'un propos détaillé. Nous avons mentionné *supra*, au début du chapitre, que les chercheurs chinois se situent, sur la question des néologismes sur la même ligne que les travaux français (malgré les différences majeures entre les deux langues). Comme une des

hypothèses de cette recherche est que les mêmes tendances de création peuvent être observées dans les corpus issus de microblogs en chinois et en français tout en tenant compte des spécificités des langues, nous allons nous appuyer sur la même typologie (Sablayrolles 2019) pour vérifier notre hypothèse en menant à bien la comparaison. Si une grande proportion de procédés communs est attestée dans les deux corpus chinois et français, notre hypothèse sera considérée comme confirmée, sinon il nous faudra la réviser. Mais en tous les cas l'appui sur une typologie commune permet une véritable comparaison.

Toutefois dans un domaine spécifique qui ne concerne pas les typologies, les chercheurs chinois ont poussé loin leur réflexion : il s'agit des phénomènes langagiers dans le contexte de l'Internet que nous allons nous attacher à explorer dans la perspective des créations néologiques. Nous allons donc nous consacrer maintenant à un état de l'art sur ces questions dans le sous-chapitre suivant en prenant en compte les travaux français et chinois.

## **3.2. Créations lexicales en contexte numérique**

### **3.2.1. Caractéristiques communicationnelles générales des créations lexicales**

Les caractéristiques des productions langagières que nous nommons provisoirement « créations lexicales » avant de voir si la dénomination de néologisme peut être retenue ou non se situent à plusieurs niveaux si nous les envisageons sous l'angle de la communication.

*1. Les causes de création.* Quelles sont les causes de l'apparition de nouveautés lexicales ? Les créations sont attribuables selon nous et sur la base de différentes lectures à deux causes principales :

- *L'acte de désignation* répond au besoin de nommer les nouvelles réalités. Cela est particulièrement pertinent pour celles qui accompagnent le développement de l'Internet et englobent non seulement les termes techniques, mais aussi les termes spécifiques de divers domaines élaborés pour l'adaptation au contexte numérique. C'est surtout sur la précision de ces termes que les divergences de catégorisation interviennent.
- *L'acte de communication* correspond aux exigences de l'interaction. Sablayrolles



mentionne dans ses récents travaux (2019) qu'une distinction ancienne oppose des néologismes de chose et des néologismes d'expression. En réfléchissant à ces deux catégories, nous estimons que si les nouvelles productions associées à l'acte de désignation peuvent être considérées comme des néologismes de chose – qui ont une fonction dénomminative –, les productions issues de l'acte de communication ne peuvent pas être incluses dans les néologismes d'expression qui ont une fonction esthétique. Les néologismes d'expression consistent plutôt en créations stylistiques dans les œuvres littéraires alors que les productions ici concernées sont étroitement liées à la communication interpersonnelle en ligne. Cette dimension communicationnelle ajoute une caractéristique spontanée et relativement imprévisible aux productions même s'il ne s'agit pas de néologismes d'expression dans l'acception ancienne.

2. *Les créateurs.* Dans leurs travaux de 2003, Pruvost et Sablayrolles indiquent la présence des néologismes dans des lieux privilégiés : presse écrite et audiovisuelle, dictionnaires, littérature, instances officielles. C'est le seul domaine littéraire qui permettrait de trouver des créations individuelles. Nous observons cependant que d'un point de vue synchronique, le contexte numérique permet aux internautes de participer au processus de la création et de transmission de nouveaux mots et expressions. Cette dimension individuelle est donc présente. Il est à noter que derrière le statut d'internaute, les identités de chaque individu constituent un élément important dans leur création verbale. Ces identités ne se limitent pas à leurs places sociales comme par exemple celle d'un homme politique ou d'un influenceur dans un réseau social numérique donné. Les identités numériques (Cardon 2008 ; Georges 2009) comptent également, y compris celles qui relèvent de l'appartenance à une communauté comme par exemple les usagers d'une communauté de jeux de vidéo.

3. *Nouveauté relative des néologismes.* Les jugements concernant la nouveauté d'un néologisme sont aussi un sujet de débat et il existe une sorte de paradoxe dans le repérage des néologismes dont le statut est lié à la perception qu'on en a et une durée variable. Sablayrolles (2019) propose de considérer le nombre de réemplois, le nombre des locuteurs différents et le nombre de cercles par lesquels transite la lexie comme critères. Un temps de circulation d'environ 10 ans ou plus est considéré par Guilbert (1973) comme la durée pendant laquelle persiste le statut de néologisme. La dictionnarisation, c'est-à-dire l'intégration des

éléments lexicaux dans des ouvrages lexicographiques<sup>103</sup> comme processus conduit à une perte du statut néologique pour certains chercheurs. Pourtant, vu la diversité des néologismes, ces critères même si on les superpose, ne peuvent pas être appliqués universellement.

En prenant en compte toutes ces réflexions appliquées au contexte numérique, nous avons découvert certaines caractéristiques significatives des créations lexicales dans ce contexte : à propos du moment de création, il est très difficile de remonter aux premières occurrences et de s'interroger sur les raisons de la création lors des innovations individuelles et spontanées ; il est presque impossible de savoir si la création a eu lieu en ligne ou hors ligne, volontairement selon les objectifs de communication, ou involontairement par suite par exemple d'une erreur de saisie ensuite entérinée. En ce qui concerne le réemploi, il est souvent instantané et viral. Les observations empiriques montrent que les néologismes pendant la période pré-numérique, par rapport au lexique en usage, sont caractérisés non seulement par leur aspect de nouveautés mais aussi par un usage moins fréquent. Or, avec les nouveaux dispositifs de communication, il est possible que les nouveautés soient reconnues par le grand public en très peu de temps et la quantité de leurs usages est importante durant cette période. En plus, un usage inter-plateformes et multi-supports (en ligne et hors ligne) se produit souvent. Cela correspond aux critères du nombre de locuteurs et de cercles différents invoqués par Sablayrolles et mentionnés ci-dessus. Cette haute fréquence d'utilisation peut conduire à deux résultats : un usage stabilisé dans la langue ou une existence éphémère. Il faut souligner en tous les cas qu'il n'est plus nécessaire d'attendre 10 ans. Ainsi, pour le premier cas (usage stabilisé), les nouvelles entrées annuelles des dictionnaires en sont un exemple. Le rythme de dictionnarisation semble également susceptible d'être beaucoup plus accéléré. Une innovation peut être sélectionnée immédiatement dans l'édition de l'année suivante. Néanmoins, la dictionnarisation n'est pas un processus obligatoire pour la stabilité. D'une part, les dictionnaires n'ont pas de critères uniformes pour la sélection. Il peut arriver qu'une innovation lexicale soit sélectionnée par un dictionnaire mais pas par un autre. D'autre part, ce sont des mots isolés qui sont privilégiés par les dictionnaires. En tous les cas, les nouveautés ne sont pas toutes pris en compte mais ce qui les caractérise à l'ère de l'Internet c'est la densité et l'ampleur de diffusion dans une durée

---

<sup>103</sup> C'est sur ce processus que nous nous appuierons pour constituer le corpus en français comme nous l'expliquerons dans le chapitre suivant.

relativement courte.

### 3.2.2. Nouveaux phénomènes langagiers en contexte numérique : le point de vue des chercheurs chinois

Nous souhaitons dans ce sous-chapitre faire une place aux recherches développées en Chine. Si elles croisent la question des néologismes en l'englobant, ce n'est pas en ces termes qu'elles la posent et les présenter nous amènera à revenir sur la dénomination de notre objet. Ce à quoi ces recherches réfléchissent c'est ce que l'Internet et ses usages font au langage à travers la circulation des messages des internautes.

Dès la première période de développement de l'Internet, son influence sur les productions langagières a attiré l'attention des chercheurs chinois. Nous allons ici nous intéresser dans un premier temps à des travaux pionniers, ceux de Zhou. Par la suite, nous allons effectuer une recherche dans une base de données scientifiques de Chine pour faire un état de l'art des études chinoises relatives à ces phénomènes dans les 20 dernières années.

#### *Weblinguistics et Weblanguage*<sup>104</sup>

Le concept de 网络语言学 *wǎng luò yǔ yán xué*<sup>105</sup> (en anglais : *Weblinguistics*<sup>106</sup>) a été en premier proposé comme une discipline en Chine par le chercheur Haizhong Zhou (2000) il y a une vingtaine d'années, en raison de la popularisation de la communication numérique et de l'émergence de la culture Web. Il suggère de s'intéresser à l'étude du *Weblanguage* (网络语言 *wǎng luò yǔ yán*, « langage de l'Internet »<sup>107</sup>) avec cette nouvelle discipline. Le

---

<sup>104</sup> En prenant compte la date de publication des articles cités dans cette recherche, nous avons ajouté des exemples récents ou usuels pour bien illustrer les différentes pratiques langagières en plus de certains exemples introduits par les auteurs.

<sup>105</sup> Traduit littéralement en français comme : 网络(*wǎng luò*) réseau informatique + 语言学(*yǔ yán xué*) linguistique. Le terme 网络 peut correspondre à plusieurs notions selon le contexte : par exemple, il est souvent utilisé pour désigner ou l'Internet ou le Web en chinois dans l'usage courant. Nous adoptons dans la suite la traduction « langage de l'Internet ».

<sup>106</sup> Il est à noter qu'en Chine du fait de la bonne diffusion des travaux anglophones et des exigences de la publication internationale (résumé, mots-clés à proposer en version anglaise par exemple), l'anglais a une place importante dans les productions scientifiques. En outre c'est la première langue étrangère enseignée.

<sup>107</sup> Notre traduction. Il s'agit d'une possibilité de traduction. Il peut aussi être considéré comme par exemple l'argot de l'Internet.

*Weblanguage* est défini comme une variété de langage créée par les internautes dans le but de répondre à leurs besoins de communication en ligne. Selon lui, ce langage est composé de mots et expressions du Web, de chiffres et de caractères informatiques. Il a divisé les mots et expressions du Web en trois catégories :

1. les termes techniques liés aux produits et usages du Web : 登录 dēng lù (connexion), 浏览器 liú lǎn qì (navigateur) ;
2. les mots et expressions pour désigner des phénomènes sociaux spécifiques du Web : 网虫 wǎng chóng (cybernaute), 网恋 wǎng liàn (amour en ligne) ;
3. les mots et les expressions usuels dans les échanges en ligne.

Pour cette troisième catégorie, il existe encore des subdivisions :

- le lexique « traditionnel » possédant un nouveau sens ; par exemple 楼主 lóu zhǔ, signifie « le propriétaire d'un immeuble » au sens originel, alors qu'en ligne, il désigne l'auteur d'une publication, que ce soit un billet de blog ou un tweet ;
- le lexique lié au contact des langues, par exemple *E-mail* est transcrit en chinois comme ceci 伊妹儿 yī mèi er<sup>108</sup> ;
- les nouveaux mots et expressions liés à la vie quotidienne des internautes : pour dire « chose », on utilise 东东 dōng dōng<sup>109</sup> à la place de 东西 dōng xī ;
- les paronymes<sup>110</sup> dus à une erreur de saisie informatique : 版主 bǎn zhǔ (Webmaster) devient 斑竹 bān zhú (avec le même sens), 美国 měi guó (les Etats-Unis) est remplacé par 米国 mǐ guó.

Les chiffres et caractères informatiques sont constitués respectivement de deux catégories :

1. Les chiffres dont la combinaison donne un sens connotatif : à titre exemple, 996 signifie un emploi du temps qui débute à 9 a.m. et se termine à 9 p.m. pour 6 jours ouvrés par semaine. Cette expression est utilisée pour décrire les conditions de travail dans les grandes entreprises en Chine en soulignant le concept de rallonger le temps de travail pour gagner la possibilité de réussir.

---

<sup>108</sup> C'est une expression employée au moment de l'introduction des E-mails en Chine. Elle n'est plus utilisée actuellement. 伊 yī est un nom de famille en Chine. 妹儿 mèi er désigne une jeune fille dans un registre familier. L'ensemble de l'expression est une translittération phonétique donnant lieu à une nouvelle structure combinatoire des caractères chinois.

<sup>109</sup> La création se fait par redoublement.

<sup>110</sup> Dans la langue chinoise le ton sert à différencier les prononciations, les caractères associés et le sens (voir chapitre I). Dans le cas des paronymes on met en jeu les différents tons ou une prononciation approchée. Dans ces exemples il n'y a pas de changement de sens.

2. Les chiffres dont la prononciation est rapprochée de certains caractères chinois : 88 prononcé comme bā bā est utilisé à la place de 拜拜 bái bái<sup>111</sup> (au revoir). Un autre exemple assez romantique : 520 wǔ'èr líng est utilisé pour exprimer 我爱你 wǒ ài nǐ (je t'aime) sur la base d'une prononciation rapprochée.

Quant aux caractères informatiques, sont distingués les sigles et les émoticônes. Voici quelques exemples de sigles en Pinyin<sup>112</sup> : le sigle *XSWL*, composé par les initiales de xiào sǐ wǒ le (mort de rire) est l'équivalent de l'expression 笑死我了 ; le sigle anglais, *GF* (*girl friend*) ou *BF* (*boy friend*). Pour les émoticônes nous ne donnerons qu'un exemple : ORZ, d'origine japonaise, ressemble à quelqu'un qui s'agenouille et baisse la tête (O figure la tête, R le corps et Z la position à genoux mains au sol). Les internautes s'en servent pour exprimer leur déception ou dire « on n'y peut rien ».

Zhou mentionne également que les changements sont aussi apparus au niveau grammatical et que ce faisant des règles de grammaire spécifiques se sont formées. Par exemple : 这/很/有/中国特色(zhè/hě'n yǒu/zhōng guó tè sè, *litt.* « Ceci / beaucoup / avoir / les caractéristiques chinoises ») devient 这/很/中国 (zhè/hě'n/zhōng guó, *litt.* « Ceci / très / Chine »). L'adverbe est employé directement pour modifier un nom. L'expression signifie : « c'est très chinois ».

Les premières connexions internationales à l'Internet en Chine débutent en 1994. Dès lors, les chercheurs chinois se sont penchés sur les phénomènes langagiers natifs de l'Internet. Nous pouvons constater que ces phénomènes langagiers se sont développés à une vitesse inédite et ont engendré des changements lexicaux, orthographiques et grammaticaux. L'étude de Zhou est parmi les travaux pionniers consacrés à ce sujet. Cette première étude marquante met en évidence de nombreuses caractéristiques spécifiques des pratiques langagières en ligne pour lesquelles il est toujours possible de trouver des productions correspondantes qui sont advenues les années suivantes. C'est le cas de certains exemples que nous avons introduits *supra*.

Certes, les catégories de Zhou ne peuvent pas englober toutes les productions, vu que de

---

<sup>111</sup> Il s'agit d'un emprunt à l'anglais « bye-bye ».

<sup>112</sup> Nous avons présenté dans le chapitre I le pinyin (système d'écriture pour la transcription phonétique du chinois mandarin) et signalé qu'aujourd'hui la plupart des internautes utilisent cette méthode pinyin pour la saisie numérique.

nouveaux dispositifs et pratiques surgissent continuellement. Nous citerons la forme composée d'un chiffre et d'une lettre : 3Q (merci) dans laquelle 3 s'ān remplace *thank* et Q est utilisé à la place de *you*. Au lieu de taper en chinois 谢谢 (xiè xiè) ou en anglais *thank you*, on a créé une forme concise en cherchant un rapprochement phonétique. Cette combinaison mixte constitue une nouvelle catégorie à part entière que nous n'avons pas trouvée chez Zhou. D'autre part, les frontières entre les catégories semblent floues, surtout pour les mots et les expressions. Comment peut-on distinguer les termes liés aux usages du Web de ceux spécifiques du Web ? Les produits du Web ne sont-ils pas spécifiques pour le contexte numérique ou ne sont-ils pas liés à la vie quotidienne des internautes ? Les divisions de catégorie sont peut-être adéquates mais les précisions manquent.

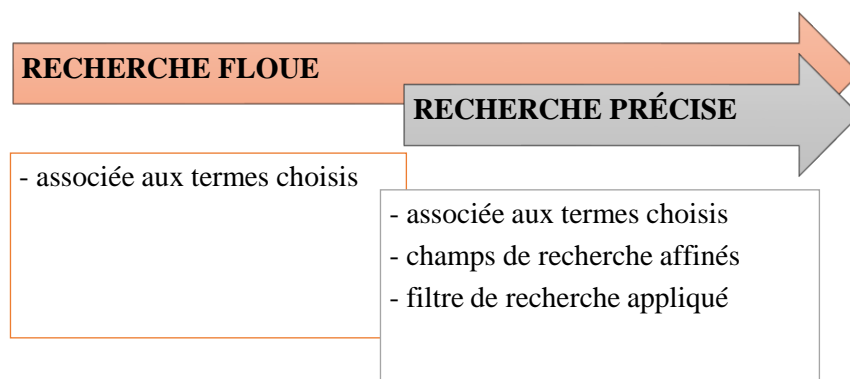
Pourtant, les aspects technique et pragmatique sont bien pris en compte pour expliquer plusieurs propriétés linguistiques dans une démarche qui associe langage et environnement. Zhou souligne que certaines productions langagières sont associées étroitement avec la saisie numérique par pinyin et les intentions des usagers. Il suppose que pour réaliser une communication rapide, les internautes font moins attention lorsqu'ils choisissent les mots ou expressions énumérés par le logiciel de saisie. Cela faisant, les homonymies et paronymes sont employées à la place du lexique correspondant. Suite à la diffusion rapide, ces « erreurs » typographiques sont reconnues par les autres internautes et leurs utilisations sont stabilisées. Les logiciels de saisie en Chine sont actuellement dotés d'une fonction qui a pour but de mémoriser les saisies fréquentes et ainsi proposer des choix approchés le plus rapidement possible. En plus, ils sont aussi connectés à leurs propres bases de données en nuage qui détectent les nouveaux usages des mots et expressions, et mettent à jour les choix possibles en temps réel. Ces fonctions techniques en ligne et hors-ligne contribuent fortement à la naissance, la diffusion et l'intégration du *weblanguage*.

### ***Un état de l'art***

L'étude de Zhou a introduit scientifiquement en Chine la notion de 网络语言 (langage de l'Internet). Cette désignation a été largement reprise ultérieurement par les autres chercheurs chinois. Nous avons mené une recherche dans la base de données chinoise CNKI, le portail de ressources bibliographiques le plus complet de Chine, dans le but de connaître les études

académiques associées à ce terme<sup>113</sup>. La base de données dispose d'un moteur de recherche principal et offre des filtres différents permettant d'obtenir un résultat plus précis. Donc, une recherche associée aux termes choisis constitue le moyen principal d'investigation.

Dans un premier temps, nous avons entrepris une recherche floue en entrant simplement le terme 网络语言 (langage de l'Internet) sans appliquer de filtres pour obtenir une observation générale. Nous avons relevé environ 8420 résultats publiés de 1988 à 2018. Etant donné que toutes les références dans la base de données sont regroupées en 10 catégories, c'est-à-dire, 10 grands champs disciplinaires, il nous a été nécessaire d'appliquer un filtre en limitant les résultats aux domaines suivants : Philosophie, Sciences humaines et sociales. Il faut souligner que la recherche (associée aux termes), par défaut du moteur, consiste en une recherche approchée. Il se peut que les caractères chinois composant un terme se dispersent dans les titres des études. Cela faisant, nous avons réalisé en même temps une recherche avancée (recherche précise du terme apparu dans les titres et les mots clés) en appliquant le filtre du champ d'étude pour obtenir les résultats qui comportent les termes exacts (exactement les mêmes mots dans le même ordre). Nous nous sommes aussi appuyée sur des travaux synthétiques portant sur les mêmes sujets pour nous aider à analyser les résultats obtenus.

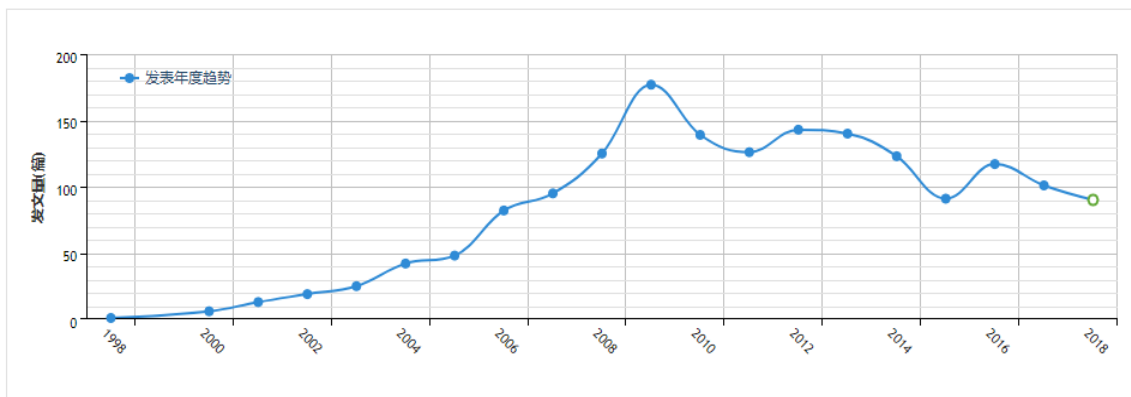


**Figure 12 : Méthodologie de recherche documentaire dans CNKI**

Une fois les conditions appliquées, nous avons obtenu un nombre moins important de résultats soit 1663 études affichées, dont 190 mémoires de master, 3 thèses et 1516 articles. La

<sup>113</sup> Cette recherche a été effectuée le 21 juin 2018. Des recherches complémentaires ont été réalisées le 18 septembre 2018.

progression dans le temps des publications est la suivante :



**Figure 13 : Nombre de publications annuelles des études portant sur 网络语言 (langage d’Internet) (1998-2008)**

Selon ces données, d’un point de vue diachronique, l’année 2000 peut être considérée comme une année de départ pour les études portant sur 网络语言 (langage de l’Internet). Le nombre de publications est en augmentation au fil des ans et atteint un pic en 2009. Il faut souligner le fait que le microblog *Sina Weibo* a été officiellement lancé la même année. À ce moment-là, il est la seule plateforme de microblog en Chine. En plus, il est ouvert pour tout public, alors que *Renren.com* (aussi reconnu comme le « Facebook chinois ») est un RSN destiné à rechercher et joindre des camarades de classe, du primaire jusqu’à l’université. Nous supposons qu’une plateforme unifiée, peut favoriser dans une certaine mesure de nouvelles pratiques langagières et, à leur suite, des études. Après 2009, les données montrent que 网络语言 (langage de l’Internet) est toujours un sujet qui retient l’attention des chercheurs chinois.

Ainsi, selon les synthèses, les perspectives de recherche varient pour des périodes différentes. Chen et Bai (2012) relèvent que :

- de 2001 à 2005, les études sont inscrites dans les perspectives linguistiques, focalisées sur le vocabulaire, la nature, la catégorisation, la norme, etc. ;
- de 2005 à 2010, les champs d’analyse sont élargis aux sciences cognitives, à la psychologie, aux aspects socioculturels et communicationnels ; l’étude de cas constitue une méthode principale durant cette période.



Néanmoins, Chu (2016) démontre que ce sujet ne gagne que la faveur des jeunes chercheurs. La question des normes reste un grand obstacle au choix de ce sujet par les chercheurs car le caractère informel du 网络语言 (langage de l'Internet) est souvent mal vu dans les domaines scientifiques. De ce fait les travaux publiés sont relativement peu nombreux. Chu ajoute que la plupart des études se sont penchées sur l'aspect lexical et que les analyses manquent d'ampleur. Selon lui, alors que ce langage spécifique s'élabore dans un contexte particulier, les outils théoriques et méthodologiques pour l'étudier sont empruntés à la linguistique « traditionnelle ». Ajoutons toutefois que les travaux émanant de différents champs disciplinaires sont attestés et que les angles de recherches ont tendance à se diversifier.

À travers l'outil d'analyse intégré à la base de données que nous avons consultée, parmi les 1663 résultats, nous avons obtenu les mots clés ci-dessous qui sont associés au terme 网络语言 (langage de l'Internet).



visibles les traces de l'anglais (Rong 2012 ; Sun et Wang 2017) et du japonais (Huang 2010 ; Liu 2011).

Nous pouvons dire que les recherches inscrites dans des perspectives interdisciplinaires n'occupent qu'un faible pourcentage même si certains travaux commencent à s'intéresser aux usagers, aux médiums, etc. Dans son article, Li (2011) s'interroge sur la relation entre les variables sociales des internautes et le développement du langage d'Internet. Les approches discursives ont été appliquées à l'analyse (Liu et Li 2017). Les différences entre sexe et genre au niveau de l'emploi des productions langagières numériques sur le réseau social chinois *Wechat* ont été investiguées (Dai 2016). Les phénomènes langagiers sur le médium *TikTok* ont attiré l'attention des chercheurs (Yan 2018). Dans le domaine de l'éducation, on se soucie de l'impact du 网络语言 (langage de l'Internet) sur l'enseignement du chinois (Hu 2016) et sur l'apprentissage des jeunes (Liang 2014).

Nous avons aussi effectué des recherches visant à collecter les travaux relatifs à la langue française dans ce champ de recherches. Il n'existe que trois travaux chinois qui datent respectivement de 2003, 2013 et 2014 (2 articles et 1 mémoire de master).

Pour nous résumer, les études associées au terme 网络语言 (langage de l'Internet) en Chine sont toujours dans une phase florissante. Plusieurs dictionnaires spécifiques consacrés à recueillir les productions langagières ont été publiés durant les deux dernières décennies. Nous pouvons constater que de plus en plus de nouveaux enjeux et perspectives. Mais jusqu'à aujourd'hui, il s'agit encore des travaux épars et il y a à ce jour très peu d'études chinoises qui visent à relier ces phénomènes langagiers en Chine et en France. De ce point de vue, notre recherche peut dans une certaine mesure y ajouter des nouveaux éléments.

### ***Questions liées à la délimitation des notions***

A travers la revue que nous venons de faire, nous pouvons arriver à la conclusion que le terme chinois 网络语言 (langage de l'Internet) est largement reconnu par les chercheurs chinois. Son utilisation est stabilisée en presque 20 ans d'évolution. Cependant, il reste toujours une notion assez floue. Cela se manifeste sous des aspects divers.

Dans un premier temps, la définition du terme et la catégorisation des productions langagières sont peu claires. Après la proposition de Zhou en 2000, une grande majorité des chercheurs se mettent d'accord sur le fait que le terme 网络语言 (langage de l'Internet) renvoie à un sens large et à un sens strict mais le contenu associé à chacun des deux est différent et à partir de là, les divergences se multiplient.

Certains chercheurs (Liu et Shi 2015) approuvent la trichotomie proposée par Zhou et citée *supra* (les termes techniques, le lexique lié aux phénomènes sociaux et les mots et expressions usuels dans les échanges en ligne) et incluent la troisième catégorie, c'est-à-dire, les mots et expressions usuels dans les échanges en ligne dans le sens strict alors que les deux autres relèvent selon eux du sens large.

Hu, Liu et Yang (2014) proposent une division en trois catégories comme suit :

- termes techniques liés à l'Internet, y compris les termes informatiques (*modem*<sup>114</sup>) et la terminologie spécifique (*souris, navigateur*) ;
- langage de communication issu des domaines professionnels (*média Internet, e-commerce, Web-littérature*), destiné à l'emploi dans le contexte d'Internet ;
- langage créé et/ou utilisé quotidiennement par les internautes dans la communication en ligne.

Liu (2015) pense que, au sens large, 网络语言 (langage de l'Internet) désigne toutes les productions langagières associées à l'Internet alors que le sens strict ne peut recevoir une description universelle mais dépend des études en situation. Ainsi, il montre que dans le champ linguistique, les travaux sont majoritairement consacrés à étudier le *Weblanguage* au sens strict, étant admis que les productions destinées à la communication en ligne et leurs formes diffèrent de celles qu'on utilise *offline*.

Dans un deuxième temps, nous avons aussi observé que quelques termes sont employés en parallèle du terme 网络语言 pour décrire ces phénomènes langagiers numériques :

- n°1 : 网络/词语 wǎng luò cí yǔ (*litt.* « l'Internet / mot et expression ») ;
- n°2 : 网络/流行/语 wǎng luò liú xíng yǔ (*litt.* « l'Internet / populaire / langue ou langage ») ;
- n°3 : 网络/热/词 wǎng luò rè cí (*litt.* « l'Internet / recherché / mot et expression ») ;

---

<sup>114</sup> Nous donnons ici des exemples en français pour illustrer leurs catégories.

- n°4 : 网路/新/词 wǎng lù xīn cí (*litt.* « l'Internet / nouveau / mot et expression »).

Le terme n°1 peut être considéré comme synonyme de 网络语言 (langage de l'Internet). Les termes n°2 et n°3, ils sont tout d'abord, synonymes de l'un et l'autre ; le n°2 correspond au français « langage en vogue, à la mode » ; le n°3 désigne *les* termes ou expressions « branchée », ce qu'on entend aussi par *buzzword*<sup>115</sup>. Nous avons procédé à une recherche précise des termes n°2 et 3 dans la base de données, et obtenu 1024 travaux dont le titre inclut le terme n°2 et 165 résultats le terme n°3. Si nous nous référons aux définitions de ces deux termes, nous pouvons récapituler certaines caractéristiques : il s'agit des productions langagières en ligne marquées par la popularité. Leur présence est récurrente pour une période donnée. Mais ils ne sont pas forcément « nouveaux » (que ce soit au niveau de la forme ou du sens). Leur utilisation peut impressionner les utilisateurs hors ligne qui vont ensuite les pratiquer en dehors de l'Internet et contribuer davantage à la propagation. Pour résumer ces deux termes qui concernent le langage dans la communication entre les internautes sont inscrits dans des perspectives socioculturelle et communicationnelle.

Le terme n°4 lui-même emploie l'expression 新词 xīn cí (que nous pouvons associer à la notion de *néologisme*) pour catégoriser ces phénomènes langagiers. Il oriente vers une perspective micro alors que 网络语言 s'inscrit dans un point de vue macro. Par rapport aux autres termes, la quantité des études est moins importante<sup>116</sup>. Cela montre que les chercheurs chinois semblent plus prudents avec ce terme ou qu'il y a encore peu de travaux sur cette perspective.

Revenons sur ces appellations, pour récapituler selon les articles de presses et les travaux scientifiques quelques résultats en commun et ainsi mieux situer notre propre étude : 网络语言 wǎng luò yǔ yán, 网络词语 wǎng luò cí yǔ, 网络热词 wǎng luò rè cí, 网络流行词 wǎng luò liú xíng cí, 网络新词 wǎng luò xīn cí. À première vue, ces appellations ont toutes

---

<sup>115</sup> Voici des définitions du mot anglais *buzzword* d'après deux dictionnaires : 1.) a word or phrase used by members of some in-group, having little or imprecise meaning but sounding impressive to outsiders» (Webster's New World Dictionary (Third College Edition) ; 2.) An stylish or trendy word or phrase, especially when occurring in a specialized field (The American Heritage® Dictionary of the English Language, 5th edition). Notre traduction : 1) un mot ou une phrase utilisé par les membres d'un groupe interne, dont le sens est imprécis mais qui semble impressionnant pour les personnes extérieures 2.) un mot ou une expression à la mode, surtout dans un domaine spécialisé. Traduit avec [www.DeepL.com/Translator](http://www.DeepL.com/Translator) (version gratuite).

<sup>116</sup> La recherche précise nous a permis d'obtenir 253 travaux.

adopté 网络 (l'Internet) comme déterminant. Donc, la relation entre les productions langagière et l'Internet est reconnue. Pour ce qui est du déterminé nous avons remarqué que : 网络语言 (langage de l'Internet) est un terme assez large et polysémique qui peut englober tous les phénomènes langagiers en ligne, y compris les termes techniques et informatiques, ou le lexique du jeu vidéo. Même si le nombre de catégories et leurs définitions sont variés selon les chercheurs, il apparaît que le sujet de notre recherche porte sur une sous-catégorie du langage de l'Internet. Cela dit, nous pensons qu'il est nécessaire de trouver un terme plus précis pour nommer le sujet auquel nous nous intéressons. Quant au terme n°4 网络新词, il est le seul terme qui souligne l'aspect néologique des productions langagières. Néanmoins, nous nous demandons si ces productions peuvent être intégrées dans le champ des néologismes, autrement dit, si les fondements théoriques de la notion de « néologisme » que nous avons rappelés dans ce chapitre (3.1.) sont toujours applicables dans le contexte spécifique de l'Internet. Pour les deux termes n°2 et 3 qui restent, ils désignent les mots et expressions populaires sur l'Internet mais un terme populaire n'est pas forcément un terme néologique.

Admettons qu'il n'est pas facile de définir un phénomène en cours de développement possédant d'ailleurs des potentialités inconnues. Nous espérons partir d'un paysage scientifique qui nous est familier mais ces appellations variées<sup>117</sup> si elles ont aiguisé notre réflexion ont constitué un obstacle pour nous en nous empêchant de distinguer dans les sources chinoises, les choix les plus adéquats pour choisir les mots clés équivalents en français, soit pour entreprendre des recherches sur les champs correspondants, soit pour établir des listes d'échantillon comme nous le verrons au chapitre suivant. Nous allons nous intéresser plus brièvement aux différentes notions proposées par quelques chercheurs en dehors de la Chine, francophones et anglophones, sans prétendre en faire un panorama et essayons y récapituler la situation d'étude.

### **3.2.3. Nouveaux phénomènes langagiers en contexte numérique : le point de vue des chercheurs francophones et anglophones**

Sous l'influence du réseau mondial, les chercheurs de pays différents s'intéressent aux

---

<sup>117</sup>Cette ambiguïté est encore plus évidente dans les termes anglais employés par les chercheurs chinois. *Cyberlanguage, Weblanguage, Online language, Netspeak, Internet language, Internet slangs, Network buzzword, Cyber néologisme*, sont tous concernés.

nouveaux phénomènes langagiers en ligne. Le linguiste américain Baron (2003:2) indique quatre sens associés à « Language of Internet »<sup>118</sup> que nous traduisons comme suit :

- (1) langage naturel utilisé sur l'Internet (par exemple, dans le courrier électronique, les listes de diffusion, les pages Web) ;
- (2) systèmes de codage spéciaux pour la construction de communication sur l'Internet (par exemple, les langages de balisage, les langages de programmation, les programmes d'application) ;
- (3) langage naturel, systèmes de codage ou interfaces de traduction utilisés pour recueillir des informations sur le Web (par exemple, les moteurs de recherche, les agents intelligents, les programmes de traduction automatique) ;
- (4) terminologie spéciale utilisée pour désigner les services ou productions de l'Internet (par exemple, réseau, serveurs, navigateurs, HTTP).

Par rapport à la proposition de Zhou, c'est l'opposition entre langage naturel – qui désigne le langage parlé par les humains –, et langage informatique – qui est un langage formel ou code de signaux entre machines –, qui structure les différentes acceptions de l'interprétation large de « Language of Internet » par Baron. Ses distinctions sont centrées sur la nature des langages présents ou liés au contexte de l'Internet tandis que celles de Zhou précisent les différents usages du langage naturel dans ce contexte.

Le linguiste anglais, Crystal a proposé une notion alternative à *Internet language*, le *Netspeak* en 2001 (Crystal 2001) pour nommer les variations linguistiques en ligne et ensuite *Internet linguistics* en 2011 (Crystal 2011) pour l'étude scientifique de toutes les manifestations langagières en ligne. Selon lui, les notions de 1) *electronic discourse* et 2) CMC (*computer-mediated communication* en anglais et en français CMO (communications médiées par ordinateur) se focalisent respectivement sur l'interaction et le médium de communication, des perspectives d'études qui ne sont pas spécifiquement linguistiques. En outre la CMC peut englober toutes les formes de communication, y compris des photos, de la musique ou des vidéos. Suite à l'apparition de nouveaux supports comme le smartphone et la tablette, les appellations de *electronically mediated communication* (EMC) or *digitally mediated communication* (DMC) ont été utilisées à la place de CMC. Crystal pense que *Internet*

---

<sup>118</sup> « (1) natural language usage carried via the Internet (e.g., in email, listservs, Web pages). (2) special coding systems for constructing communication carried via the Internet (e.g., markup languages, programming languages, applications programs) (3) natural language, coding system, or translation interfaces used for gathering information from the World Wide Web (e.g., search engines, intelligent agents, machine translation programs) (4) special terminology used in talking about the Internet (e.g., networking, servers, browsers, HTTP) » (Baron, 2003).

*linguistics* est un terme neutre et global qui peut inclure toutes les activités langagières en ligne.

Le rapport entre l'écrit et les nouvelles techniques se développe dès la fin des années 1990 en France mais surtout en Sciences de l'Information et de la Communication. Les linguistes semblent alors moins mobilisés par les productions écrites des nouveaux dispositifs. Toutefois les travaux pionniers de Jacques Anis dès 1990 (2006) doivent être signalés. Ce spécialiste de l'écrit étudie précisément l'incidence sur la langue de l'écriture informatisée ; il fait une typologie des dispositifs de ce qu'il appelle la communication électronique scripturale et une description systématique des formes des SMS et des chats à tous les niveaux linguistiques : du point de vue graphique, phonétique, morpho-lexical, syntaxique. Il publie notamment en 2001 un ouvrage destiné au grand public *Parlez-vous texto: guide des nouveaux langages du réseau*<sup>119</sup>. Par la suite des analyses linguistiques et socio-linguistiques à grande échelle sont menées à partir de corpus authentiques par des collectifs de chercheurs comme dans le projet international Sud4science<sup>120</sup> qui invite le public à « léguer ses SMS » à la sciences afin de collecter et d'analyser des SMS dans le monde entier et en grand nombre.

Les Sciences de l'Information et de la Communication en France ont examiné les nouvelles formes de l'écrit issues de l'usage d'appareils informatiques et d'Internet et fait des propositions théoriques importantes dans la décennie 90-2000 qui restent très porteuses dans les travaux actuels. Citons ici la notion d'*écrit d'écran* proposée par Emmanuël Souchier (1996) et d'*écriture numérique* travaillée par Bachimont (2000), et plus tard par Bouchardon (2014).

La notion de « l'écrit d'écran » postule « l'inter-détermination des modalités d'écriture, du support et des outils, des dispositifs techniques et des pratiques sociales d'écriture »<sup>121</sup> et invite à réfléchir aux différentes mutations portant sur « la matérialité et les supports, l'acte et les pratiques d'écriture, sur les partenaires de l'écrit, sur la division du travail, sur le temps, l'espace et la diffusion de l'écrit » (Souchier 1996:106). Selon Souchier, l'écran est le lieu où l'écriture informatique s'effectue en introduisant un nouvel espace visuel et de nouvelles pratiques.

---

<sup>119</sup> Editions du Cherche Midi.

<sup>120</sup> [www.sud4science.org](http://www.sud4science.org).

<sup>121</sup> Disponible sur : <http://www.noscreen.com/ecriture.html> (consulté le 7 mars 2020).



En ce qui concerne « l'écriture numérique », Bachimont met en relief le fait que l'informatique introduit une nouvelle dimension de l'écriture qui n'est plus seulement graphique mais passe par les procédures de calcul des ordinateurs aussi bien pour la production que pour l'interprétation. Dans une certaine mesure, elle s'éloigne de la langue dont elle n'est plus la transcription directe et passe par des codes plus ou moins visibles. Dans le même ordre d'idées, Bouchardon (2014) s'intéresse à l'« écriture numérique » sous deux angles : un angle anthropologique et un angle artistique avec la littérature numérique. Il considère que l'avènement de machines qui écrivent et lisent modifie le rapport général à l'écriture, par une rupture entre traces et supports et une omniprésence du calcul informatique. Il développe des perspectives d'enseignement et des propositions pédagogiques.

Ces considérations sont assez éloignées des préoccupations linguistiques mais elles vont en quelque sorte se rejoindre dans les travaux de Marie-Anne Paveau qui pense que le langage verbal ne peut pas être dissocié de ses dimensions sociales mais aussi techniques et corporelles et que la linguistique doit connaître un tournant et dépasser les méthodes qu'elle appelle logocentrées (Paveau 2017). Elle appelle son approche *linguistique symétrique* (Paveau 2009) ou approche *écologique* du discours (Paveau 2013). Plutôt que de linguistique proprement dite, on peut dire qu'il s'agit d'une nouvelle approche en analyse de discours en faisant du discours numérique un nouveau champ de recherches avec des outils spécifiques. Nous y reviendrons car nous nous appuyons principalement sur ses travaux pour nos analyses de corpus.

Cette approche de Paveau qui fait le lien entre linguistique communication et discours se distingue de travaux d'autres analystes de discours comme Maingueneau (2013) par exemple qui étudient les productions en ligne, notamment sur les blogs et réseaux sociaux en adoptant une approche pré-numérique, c'est-à-dire en utilisant des concepts non spécifiques comme le genre, l'énonciation ou l'argumentation.

Nous avons constaté en tous les cas que les études sur le langage dans le contexte numérique prennent en compte l'environnement / le support à des degrés divers. Ces travaux anglophones et francophones nous ont fourni des perspectives alternatives ou complémentaires aux approches chinoises sans que nous puissions faire coïncider les terminologies et sans que nous puissions vraiment isoler ce qui correspond à notre intérêt précis de recherche.

Au terme de ce sous-chapitre, nous faisons le constat que l'étude des phénomènes langagiers en contexte numérique dans l'espace de recherche chinois d'une part, anglophone et francophone d'autre part, était abordée sous des perspectives différentes, par des disciplines différentes et toujours en évolution. De ce fait, les notions et les appellations sont variées.

Notre objet de recherche se situe à l'intersection de ce champ d'études du langage et contexte numérique et du champ d'études de la néologie puisque nous nous intéressons aux productions verbales dont la création et/ou diffusion sont favorisée par le contexte de l'Internet. Notre réflexion s'est élaborée à ce croisement en ayant pour projet de mobiliser les connaissances des différents champs, de nous orienter parmi les travaux en mettant au point notre propre questionnement. Nous avons peu à peu été persuadée qu'il importait de trouver une appellation adaptée. Nous faisons état maintenant de nos réflexions et de notre cheminement.

### **3.3. Nommer notre objet d'étude**

Comment nommer notre objet d'étude ? Cette question a été présente dès l'origine de la thèse et a aussi été liée à nos choix de démarche méthodologique que nous exposerons dans le chapitre IV. Notre centre d'intérêt n'a pas changé mais nous avons pris conscience qu'il présentait des facettes différentes et qu'il était nécessaire de clarifier pour nous ce qui était au premier plan. Nous allons évoquer successivement les différentes appellations que nous avons envisagées.

Les aspects morphologiques et sémantiques particulièrement manifestés des productions verbales que nous observions nous ont conduit à nous référer dans un premier temps à la notion de « néologisme », présentée dans la première partie du présent chapitre. Nous avons supposé que les nouveaux phénomènes langagiers liés à l'usage de l'Internet répondaient aux fondements de la néologie et constituaient ainsi une catégorie dynamique de néologismes.

Nos lectures sur la notion de 网络新词 wǎng luò xīn cí (langage de l'Internet) dans les travaux scientifiques chinois nous ont inspirée. Nos connaissances empiriques en tant que internaute ont aussi eu un rôle. Par analogie, avec l'expression de *Internet language* proposé par Crystal et Baron, nous nous sommes proposé de nommer notre objet d'étude « **néologisme de l'Internet** ».

Mais nous avons remarqué une utilisation non uniformisée des notions « Internet » et « Web » dans les diverses appellations existantes. Nous avons constaté cela en comparant les notions chinoises et les notions anglophones, mais aussi les notions chinoises de périodes différentes. Par exemple, le linguiste chinois Zhou a proposé la notion « Weblanguage » comme pionnière dans le domaine. Mais dans les autres travaux chinois, il est difficile de distinguer l'usage de « Internet » et « Web » dans les noms chinois proposés puisque le même terme 网络 (wǎng luò) est utilisé comme équivalent des deux en chinois. De ce fait, c'est le résumé en anglais des travaux qui permet de les distinguer. Les confusions sur ces deux concepts et la traduction chinoise non précise nous ont amenée à préciser la relation entre « Internet » et « Web » comme nous l'avons présentée au début du Chapitre II.

Il est devenu clair pour nous que l'Internet joue le rôle d'infrastructure et que Web constitue un de ses services. En même temps, nous avons pu situer les nouveaux phénomènes langagiers qui nous intéressent dans le Web 2.0, c'est-à-dire une étape d'évolution du Web. Depuis les années 2000, le Web 2.0 offre des services qui favorisent les initiatives créatives des internautes et un échange interpersonnel instantané. Les réseaux sociaux numériques constituent ainsi les dispositifs socio-techniques représentatifs de cette étape. Nous pensons que les nouveaux phénomènes langagiers récents sont étroitement liés à ces services. Pour bien prendre en considération les éléments environnementaux et délimiter notre sujet d'étude, nous avons abandonné l'appellation « Néologisme de l'Internet » comme trop générale et large. Nous nous sommes arrêtée à l'appellation « **Néologisme 2.0** ».

Les évolutions informatiques et leurs influences ne restent pas inaperçues dans les domaines de recherche en France. Pourtant, les essais de délimiter et définir les nouveaux phénomènes langagiers nous paraissent moins développés que les études chinoises. Ou bien, les termes que nous avons rencontrés au cours de la recherche n'ont pas un objectif définitoire. Par exemple, nous avons relevé l'utilisation du terme « Cybernéologisme » dans une étude sur l'identification et le traitement des néologismes en ligne (Issac 2011). Mais ce terme vise surtout à mettre en relief le nouvel environnement de diffusion des néologismes qui sont interprétés d'un point de vue informatique<sup>122</sup>. L'expression « les néologismes de l'Internet » est utilisée par Reboul-

---

<sup>122</sup> Il s'agit d'élaborer un outil de détection et de traitement d'unités en les comparant à des ressources linguistiques répertoriés par un système de traitement automatique des langues et de repérer les « inconnus », les formes non modélisées dites « candidats néologismes » afin de les ventiler dans différentes catégories.

Touré (2010) mais cela concerne seulement les termes terminologiques en lien avec l'Internet, qui y réfèrent.

En travaillant sur les notions francophones que nous avons présentées plus haut pour les comparer avec les anglophones et les chinoises, notre attention a été attirée par la notion de « culture numérique » qui souligne à la fois le façonnement que les changements informatiques opèrent sur les productions culturelles et sur les pratiques dont les initiatives participatives ne sont pas à négliger. Pour marquer la relation entre les nouveaux phénomènes langagiers et ces deux aspects, nous avons pensé à l'expression « **Néologisme numérique** » pour nommer notre objet d'étude.

Mais un autre problème s'est posé : nous avons intuitivement considéré que les créations verbales étaient produites dans le contexte numérique, – sans pouvoir pour autant en apporter la preuve précise –, et cela nous a fait longuement réfléchir aux lieux les plus opportuns de recueil de notre corpus.

Qu'il s'agisse d' Internet ou du Web, les associations du concept 新词 xīn cí (néologisme) aux divers phénomènes langagiers par les chercheurs chinois donnent également lieu à des confusions. Cela tient principalement à la tendance à inclure sous la désignation de « néologisme » tous les phénomènes langagiers en ligne. Plus particulièrement, la distinction n'est pas faite entre des néologismes et des mots ou expressions fréquemment utilisés en ligne pendant une période donnée (ce qu'on nomme parfois 网络流行语 « langage d'Internet en vogue »). Ou bien encore le néologisme est pris comme synonyme de « jargon ». Cela nous a conduit à nous recentrer sur l'étude linguistique de la néologie et des néologismes et à nous intéresser précisément aux typologies.

Vu la complexité et le dynamisme constatés dans les études sur les néologismes et leurs typologies, nous avons trouvé qu'il était ambitieux et hasardeux de proposer une nouvelle catégorie de néologismes. En plus, ce n'est pas l'objectif principal de notre étude.

En nous interrogeant toujours plus précisément sur notre objet d'étude à partir de la collecte du corpus, nous avons découvert que ce n'était pas que les procédés linguistiques qui étaient concernés mais aussi les éléments qui participent à la communication (interactions, dispositifs)

et peuvent ne pas être strictement verbaux (repères techniques, chiffres, icônes). C'est pourquoi nous préférons *langagier* plus englobant, à *linguistique*. Mais ce ne sont pas tous les phénomènes langagiers que nous voulons observer : plus spécifiquement les nouveautés. Nous avons donc décidé de nommer notre objet de recherche « **nouveautés langagières en contexte numérique** ». Il s'agit de nouvelles productions langagières tant individuelles que institutionnelles dont la création et/ou la diffusion sont inséparables des services d'Internet (nous avons renoncé au sentiment intuitif que la création émanait strictement du contexte numérique) et dont les emplois sont massifs pour une durée déterminée au moins.

## **Conclusion**

Dans un premier temps nous nous sommes centrée sur les travaux de lexicologie et la question des néologismes pour aborder la problématique des créations lexicales sous l'angle linguistique. Dans un deuxième temps, nous avons pu mettre en relief certaines caractéristiques communicationnelles en nous interrogeant sur les causes et les créateurs de nouveautés. Nous avons ensuite constaté que les phénomènes langagiers en contexte numérique intéressent des chercheurs venant de pays et de disciplines différents qui les abordent sous des perspectives variées ; l'hétérogénéité des travaux offrant une abondance de termes ne nous a pas permis de trouver une appellation unique et consensuelle pour désigner notre objet de recherche : nous avons trouvé en somme la même difficulté que pour les études consacrées à la néologie et à la création lexicale. Nous sommes parvenue toutefois à préciser notre objet de recherche, les nouveaux mots et expressions utilisées dans les communications interpersonnelles, notamment en ligne, bénéficiant une densité et une ampleur de diffusion dans une durée relativement courte. En reprenant les notions de dispositif abordées dans le Chapitre II, les diverses appellations exposées dans ce chapitre et les caractéristiques observées, nous avons fait le choix de dénommer notre objet « nouveautés langagières en contexte numérique ». La richesse des points de vue dans des champs de recherche qui se croisent sans se confondre a pu causer notre perplexité, mais nous a aussi aidée à nous orienter non seulement vers les aspects linguistiques mais aussi communicationnels et énonciatifs pour l'observation de notre objet que nous allons dans les chapitres qui suivent mettre à l'épreuve des corpus.

## **PARTIE II : MÉTHODOLOGIE ET ANALYSES**



## **CHAPITRE IV**

### **Présentation des corpus et de la méthodologie de constitution des corpus**

#### **Introduction**

Le sujet d'étude que nous avons choisi l'a d'abord été intuitivement sur la base de notre expérience et nous l'avons peu à peu affiné en nous posant de nombreuses questions, en entreprenant des recherches théoriques dont la première partie de la thèse fait le bilan, en nous heurtant aussi à des difficultés. Ces difficultés sont liées soit à l'objet lui-même (les nouveautés langagières dans le contexte numérique des RSN et spécifiquement des plateformes de microblogging), soit à l'objectif de comparaison (Chine / France), soit à la méthodologie à mettre en œuvre pour mener à bien cette comparaison dans des contextes sociaux et linguistiques différents. Les difficultés ne sont pas des éléments négatifs car elles ont été aussi pour nous des opportunités de réflexion et les solutions que nous avons trouvées restent soumises à la discussion. C'est pourquoi nous voulons retracer notre cheminement dans ce chapitre en commençant par évoquer ces difficultés caractéristiques de la complexité de notre objet et les principales options qui sont les nôtres.

Ensuite nous précisons les étapes de nos démarches préparatoires à la collecte des corpus avant de décrire les grandes lignes qui ont présidé à nos choix. Ceux-ci seront détaillés dans les chapitres d'analyse suivants.

#### **4.1. Questionnements et principales options relatives au corpus**

Premièrement, comme nous l'avons signalé dans l'introduction générale, nous avons constaté que trois contextes étaient impliqués et devaient être pris en compte dans la recherche des nouveautés langagières : le contexte numérique d'origine que nous aurions voulu pouvoir isoler, celui de l'apparition supposée des créations ; le contexte numérique et non numérique de diffusion des nouveautés langagières ; le contexte institutionnel et scientifique d'observation, de caractérisation ou de régulation de ces nouveautés. Cela va de pair avec la richesse et



l'importance de notre objet mais n'est pas sans entraîner des chevauchements et une pluralité de perspectives avec des risques de confusion.

Deuxièmement, en relation avec ce dernier contexte scientifique en particulier, la variété des appellations chez les chercheurs chinois et français comme dans l'usage courant, mais aussi les frontières entre les catégories de phénomènes sont un problème qui n'est pas vraiment résolu. Nous avons adopté une solution provisoire et personnelle en parlant quant à nous de « nouveautés langagières en contexte numérique » comme nous l'avons exposé à la fin du chapitre III. Nous devons cependant tenir compte dans notre protocole d'établissement du corpus des différentes appellations autres que celle sur laquelle nous nous sommes arrêtée et de l'équivalence relative d'une langue à une autre afin de recueillir des éléments pertinents mais éventuellement nommés différemment.

Troisièmement, un des principaux fils conducteurs de notre travail, qui prend sa source dans notre expérience personnelle et universitaire, est la mise en relation entre Chine et France aussi bien pour les langues et particulièrement le lexique, que pour les discours (messages des RSN), et pour les pratiques institutionnelles et communicationnelles. Comme ces dernières sont différentes, nous n'avons pas eu affaire à deux contextes analogues et n'avons pas pu procéder exactement de la même façon.

Enfin, dans la perspective de recueillir et analyser un corpus en relation avec le changement linguistique et avec des plateformes à très haute productivité discursive, nous avons affaire à un corpus en constante expansion, considérable dans son volume. Or notre démarche est qualitative et ne peut donc envisager un travail que sur des données jugées significatives et pas forcément représentatives. Donc nous sommes amenée à travailler sur des échantillons et à faire en sorte de les déterminer, de les choisir et de justifier ce choix. C'est ce que nous allons faire dans ce court chapitre en préalable à l'analyse comparée de deux sortes de corpus.

En effet, notre corpus d'étude comporte deux catégories : un corpus de nouveautés langagières et un corpus de tweets contenant ces nouveautés langagières. L'analyse comparative porte sur les langues chinoise et française puisque nous constituons notre corpus dans les deux langues inscrites dans les deux sociétés. Elle porte aussi sur les discours issus des deux plateformes de microblog que nous avons présentées au chapitre II, respectivement *Sina Weibo* et *Twitter*.

La période d'observation pour la collecte du corpus de nouveautés langagières est délimitée de la manière suivante : du mois d'août 2009 au mois de décembre 2017. Cette période définie en amont du début de notre thèse et sur une durée suffisamment longue pour pouvoir constater des évolutions répond aussi à des critères de fonctionnement des deux plateformes. Pour la borne initiale, 2009 est l'année de fondation du microblog *Sina Weibo*, plus tardivement que *Twitter* qui a été créé en 2006. Pour la borne finale, 2017 est l'année où un peu plus tard que *Sina*, *Twitter* a annoncé le passage de 140 à 280 signes pour un tweet<sup>123</sup>. Cette contrainte technique constitue la spécificité du microblog qui le distingue des autres RSN. C'est aussi pour observer l'influence de cette modalité d'expression sur les lexiques et les discours que nous choisissons d'observer les nouveautés langagières sur les plateformes de microblog. Donc, notre choix de période prend en considération l'environnement discursif et technologique du nouveau dispositif. Pour les corpus de tweets, nous n'avons pas délimité une période d'observation identique à celle des corpus de liste en raison des différences entre les deux plateformes et des situations d'utilisation variées des nouveautés langagières. Nous allons exposer cela en détail ultérieurement dans le chapitre VI.

Nous avons pour objectif d'analyser un corpus de mots ou d'expressions mis en liste que nous appellerons « corpus de listes » et un corpus de discours que nous appellerons « corpus de tweets<sup>124</sup> ». La nature de ces corpus ainsi que nos orientations nous incitent à ne pas dissocier le linguistique, le discursif et le communicationnel et nous nous efforcerons d'articuler ces trois dimensions en nous appuyant notamment sur les réflexions et informations des trois premiers chapitres, ainsi que sur certains compléments que nous apporterons au fur et à mesure.

Nous nous intéresserons dans un premier temps aux caractéristiques linguistiques d'un point de vue comparatif en nous centrant sur les procédés néologiques (chapitre V) et dans un deuxième temps aux caractéristiques discursives des tweets (chapitre VII) après avoir décrit en profondeur la constitution de leur corpus (chapitre VI). Comme nous l'avons indiqué *supra* les différents

---

<sup>123</sup> *Sina Weibo* a annulé officiellement la limite de 140 caractères en novembre 2016. Selon les informations disponibles sur : <http://www.topedmeds.net/news/4943.html> (consulté le 13 mars 2018).

En novembre 2017, *Twitter* a annoncé le passage de 140 à 280 signes. Source d'information disponible sur : [https://www.lemonde.fr/idees/article/2017/11/17/twitter-un-changement-de-taille\\_5216117\\_3232.html](https://www.lemonde.fr/idees/article/2017/11/17/twitter-un-changement-de-taille_5216117_3232.html) (consulté le 14 mars 2018). Les présentations de *Sina Weibo* et *Twitter* ont été faites dans le chapitre II.

<sup>124</sup> Nous utilisons le terme de tweet aussi bien pour les messages en chinois qu'en français car le terme est assez usuel.

défis qu'induit la complexité des phénomènes langagiers que nous voulons étudier nous amènent dans ce chapitre à présenter nos démarches qui nous apparaissent comme une tentative progressive pour dépasser les difficultés rencontrées, en faisant des choix que nous justifierons.

## **4.2. Démarches préparatoires en vue de la collecte des corpus**

Notre objectif était d'opérer une sélection de termes déjà cautionnés par un métadiscours autorisé savant ou institutionnel afin d'échapper au risque de subjectivité de notre part. Nous n'avons pas procédé de la même manière pour le corpus en chinois et pour le corpus en français mais nous avons essayé de respecter les mêmes principes. Ces corpus se distinguent par la langue, certes, mais aussi par leurs contextes technologique et social : deux plateformes différentes quoique voisines, deux pays dont l'organisation relative à la langue, à son enseignement et sa diffusion ne sont pas analogues. Nous avons essayé de tracer notre voie d'étude de manière personnelle en réfléchissant au fur et à mesure à ce que nous entreprenions.

### **4.2.1. Observations et démarche de collecte pour le corpus (en) chinois**

Nos observations pour le chinois portent sur le microblog *Sina Weibo* dont nous sommes à la fois utilisatrice et observatrice et nous avons eu recours à certains outils qu'il propose<sup>125</sup>. Nous avons constaté la présence de multiples nouveautés langagières dans le contexte d'expansion en ligne et hors ligne et le contexte scientifique. En commençant par l'expansion en ligne, nous citerons les moteurs de recherche comme Google ou Baidu (le moteur prédominant en Chine)<sup>126</sup> pour obtenir des résultats en saisissant les nouveautés, les articles encyclopédiques qui présentent et commentent ces phénomènes langagiers ou les articles qui les emploient, notamment la presse, d'autres productions linguistiques ou culturelles des internautes (ex. : le *fansub*<sup>127</sup>, les ressources encyclopédiques partagées, etc.), les communications des entreprises commerciales ou les textes de publicité. Dans le contexte hors ligne, la pénétration des nouvelles productions langagières dans la société chinoise est omniprésente et profonde. Si la

---

<sup>125</sup> Nous nous appuyons sur le moteur de recherche interne du microblog. Les observations sont basées sur les résultats de recherche, c'est-à-dire, les tweets contenant les nouveaux mots et expressions que nous avons saisis dans le moteur. Ces nouveautés langagières sont choisies à partir de nos propres expériences d'utilisation du microblog.

<sup>126</sup> Google n'est plus disponible en Chine depuis 2010 pour des raisons de restriction politique.

<sup>127</sup> Le sous-titre d'une production audiovisuelle étrangère créé par des amateurs dans une langue d'arrivée et diffusée en ligne.

diffusion en ligne de ces productions touche plutôt la jeune génération, la transmission hors ligne grâce aux médias peut atteindre tous les âges. Dans les journaux, les œuvres littéraires, cinématographiques ou les feuillets, à la télévision ou à la radio (en fonction des types d'émissions), leur présence est sensible. Il y a quelques années, les spectacles du Gala de Nouvel An de CCTV<sup>128</sup> ont commencé à employer les nouvelles productions langagières. Précisons que ce Gala constitue une émission télévisée très populaire en Chine (Figure 14). Regarder cette émission la veille du Nouvel An chinois est traditionnel pour tous les Chinois. Sans exagérer l'importance de cette émission, nous pouvons dire qu'une fois diffusées via le Gala et la chaîne CCTV, ces nouveautés seront reconnues au niveau national.



Figure 14 : Une scène du sketch 不差钱 (bú chā qián)<sup>129</sup> présenté au Gala de Nouvel An 2009

Nous avons aussi remarqué plusieurs activités organisées en Chine par des magazines de grande notoriété, certains organes de presse de référence ou encore des institutions qui veillent sur la langue chinoise. Ces activités visent à établir une liste regroupant les nouvelles productions langagières les plus populaires de l'année. En nous inspirant de ces activités, nous avons estimé que les critères de popularité et de fréquence d'utilisation étaient convenables pour constituer un corpus chinois au sein de la masse langagière à considérer. Le critère de popularité peut

<sup>128</sup> La Télévision centrale de Chine (中国中央电视台, zhōng guó zhōng yāng diàn shì tái), généralement abrégée en CCTV, de l'anglais *China Central Television*, est le réseau principal de télévision publique de la République populaire de Chine. Le lien de son site officiel est : <https://www.cctv.com/>.

<sup>129</sup> Un des mots sélectionné de l'année 2009 par le magazine 《咬文嚼字》(yǎo wén jiáo zì), qui signifie « ne manque pas d'argent », voir *infra*.

avant tout nous aider à limiter le nombre des choix. Il indique que les productions concernées ont été employées à un moment donné et ont connu une diffusion intense. La fréquence d'utilisation permet d'observer ultérieurement des itinéraires de circulation multiples. Mais en même temps, nous avons pris conscience qu'en tant que chercheur indépendant, nous ne disposons ni des moyens d'accès individuels à la masse de données ni des outils capables d'effectuer une veille sur des données à grande échelle et d'établir une liste d'autorité selon les deux critères comme le font les organes ou les institutions. Donc après avoir pris connaissance des processus de sélection adoptés par les institutions, nous avons décidé de constituer le corpus chinois de nouveautés langagières sous forme de liste, directement à partir des résultats issus de diverses activités de sélection.

Nous allons maintenant présenter ces activités :

1) La première se nomme 《咬文嚼字》十大流行语 (« yǎo wén jiáo zì » shí dà liú xíng yǔ) et consiste à identifier les 10 mots et expressions les plus populaires de l'année choisis par le magazine 《咬文嚼字》(yǎo wén jiáo zì)<sup>130</sup> (Figure 15).

---

<sup>130</sup> Le magazine mensuel dont le nom traduit directement signifie « peser les paroles et les mots » existe depuis 1995. Sa mission principale est de révéler et rectifier les erreurs linguistiques de la langue chinoise, surtout les caractères fautifs utilisés dans les médias, les œuvres des célébrités, et par les personnes connues, les artistes, etc. Tout comme décrit dans son objectif: 宣传语文规范, 传播语文知识, 引导语文生活, 推动语文学习 (Notre traduction : Diffuser la norme et la connaissance linguistiques chinoises, guider la culture chinoise et promouvoir l'apprentissage du chinois). Il est considéré aujourd'hui comme le « défenseur » d'une culture chinoise qualifiée. Ses lecteurs ciblés sont rédacteurs, journalistes, relecteurs, professionnels de la publicité, secrétaires, présentateurs, enseignants et étudiants. En 2007, il a déjà un tirage annuel de 200 000 exemplaires.

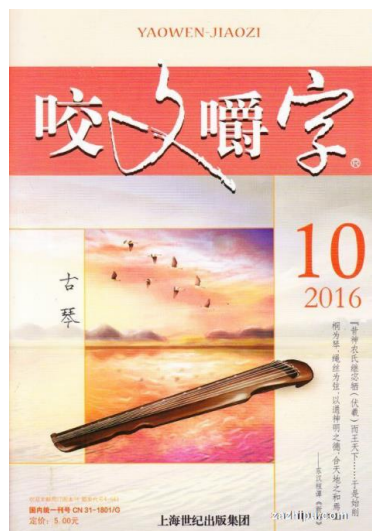


Figure 15 : La couverture du magazine 《咬文嚼字》 (octobre 2016)

À partir de l'année 2008, cette institution a commencé à publier dans ses archives annuelles les résultats de sélection chaque année. La sélection commence à partir du mois d'août et le résultat est publié vers fin décembre. Mais nous ne pouvons obtenir d'informations directes (celles qui sont disponibles en ligne) de cette activité qu'à travers le compte *Sina Weibo* et le blog du magazine. Il ne dispose pas d'un site propre, et est souvent hébergé par des sites de périodiques avec seulement une rubrique de présentation simple. Il a publié son premier billet sur le blog en 2007, mais son compte *Sina Weibo* a été créé bien tardivement, en 2012. Comme le RSN limite les messages à 140 caractères, la publication de la sélection sur son compte microblog n'en affiche qu'une partie, et le reste du contenu se présente sous la forme d'un lien hypertexte (Figure 16). En cliquant sur le lien, on est renvoyé directement vers la page de son blog<sup>131</sup>. Donc, le lieu principal pour constituer le premier corpus, est le blog de ce magazine<sup>132</sup>.

---

<sup>131</sup> L'adresse du blog est la suivante : <http://blog.sina.com.cn/u/1275434742>.

<sup>132</sup> Les billets du blog collectés sont disponibles dans l'Annexe n°6a.

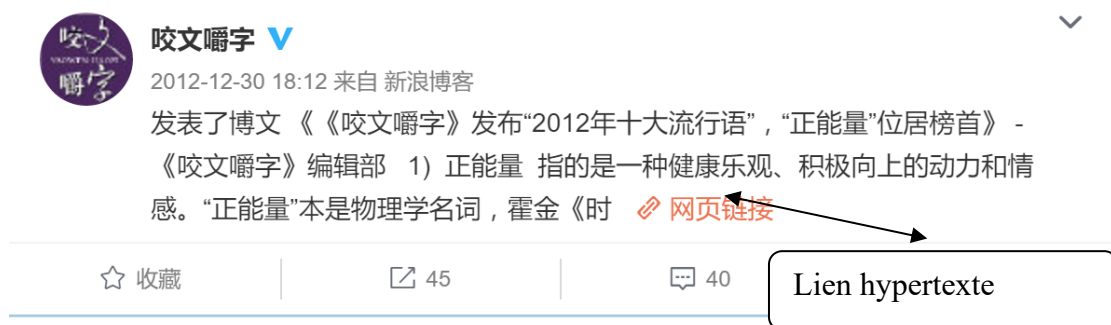


Figure 16 : Billet du compte *Sina Weibo* de magazine annonçant le résultat de la sélection 2012

2) La deuxième activité s'appelle 《汉语盘点》 (hàn yǔ pán diǎn, « Inventaire de la langue chinoise »), et publie un palmarès intitulé 十大网络用语 (shí dà wǎng luò yòng yǔ) que nous pouvons traduire comme « le top des 10 termes les plus utilisés en ligne ». C'est une des activités culturelles les plus influentes en Chine. Elle est lancée par le *National Language Resources Monitoring and Research Center*<sup>133</sup> en coopération avec différentes institutions de grande importance (la maison d'édition *The Commercial Press*, le site web *Sina*, l'organe de presse *Le Quotidien du Peuple*, le site de la chaîne de télévision *CCTV.com*, etc.) en Chine (Figure 17). Cette activité annuelle a débuté en 2006. Son but initial est de faire appel aux internautes pour décrire l'année courante par un mot ou une expression. À partir de l'année 2010, la date de début de l'activité a été changée et placée en fin d'année au lieu du début d'année. Depuis 2011, le démarrage de l'activité est fixé au 20 novembre, et le résultat est publié le 20 décembre. De 2006 à 2010, cette activité s'est concentrée sur la sélection de mot ou expression de l'année, à savoir 年度字词 (nián dù zì cí). À partir de 2011, son contenu commence à s'enrichir, et comme peut en témoigner la publication des catégories 中国媒体十大流行语 (zhōng guó méi tǐ shí dà liú xíng yǔ, « le top des 10 mots et expressions les plus populaires dans les médias »), 中国媒体十大新词语 (zhōng guó méi tǐ shí dà xīn cí yǔ, « le top des 10 nouveaux mots et expressions les plus utilisés par la presse »). Depuis 2012, les résultats sont encore plus précisés, et la catégorie de 十大网络用语 (shí dà wǎng luò yòng yǔ,

<sup>133</sup> 中国国家语言资源监测与研究中心 : <http://cnlr.blcu.edu.cn/>. Ce centre promoteur est établi par le Ministère de l'Éducation chinois pour veiller sur la langue chinoise et favoriser les études scientifiques.

« le top des 10 termes les plus utilisés en ligne ») a vu le jour. Nous nous intéressons à cette dernière catégorie qui est relativement liée à un usage individuel et un usage dans le contexte de l'Internet. Les autres, comme indiqué dans leurs noms, sont des préférés des médias.



Figure 17 : Capture d'écran d'une émission d'information de la Chaîne CCTV (le 19 décembre 2015) qui rapporte l'activité 《汉语盘点》 2014

3) La troisième émane de 人民网舆情监测室<sup>134</sup> et propose un palmarès<sup>135</sup> des mots populaires de l'Internet (网络流行语, wǎng luò liú xíng yǔ) de l'année 2014 lors du 20<sup>e</sup> anniversaire de l'accès à l'Internet en Chine (Tableau 17). La période de sélection était limitée du 1<sup>er</sup> janvier 2014 au 31 décembre 2014. En s'appuyant sur le nombre de citations sur *Sina Weibo* (c'est-à-dire la fréquence d'utilisation des mots ou expressions sur la plateforme) et les résultats de

<sup>134</sup> Le nom est devenu 人民网舆情数据中心 (Centre de données de l'opinion publique) depuis 2017. Il est hébergé par le site web du Journal « 人民日报 », *Le Quotidien du Peuple* (l'organe de presse officiel du Comité central du Parti communiste chinois). Cette branche au sein du journal a été établie en 2008. Elle est pionnière dans l'analyse et la recherche de l'opinion publique sur l'Internet en Chine, grâce à un système complet de monitoring et des centaines de chercheurs venus de domaines disciplinaires différents (y compris SIC, Sociologie, Economie, Gestion publique, Statistique mathématique, etc.), ce centre peut surveiller 24h/24 les supports principaux (version web des médias traditionnels, suivis d'actualité, communautés en ligne / forums / BBS, microblogs, blogs personnels de leaders d'opinion) de l'opinion publique en ligne et entreprendre une analyse professionnelle.

<sup>135</sup> Le résultat de la sélection a été publié dans 《2014 中国网络语象报告》 (nom que nous traduisons ainsi « Le rapport sur le langage de l'Internet en Chine de l'année 2014 »). Ce rapport est disponible à l'adresse suivante : <http://yuqing.people.com.cn/NMediaFile/2014/1219/MAIN201412190912000136898523380.pdf> (consulté le 10 mai 2016).



recherche dans leur système de journalisation, le centre a obtenu 25 résultats, les plus populaires dans le contexte d'Internet.

表 2014 网络流行语排行。

排名	网络热词	中文报刊检索结果 (标题中使用)	微博提及量
1.	萌萌哒。	1410。	214,720,184。
2.	逗比。	276。	150,043,492。
3.	我也是醉了。	298。	79,679,804。
4.	且行且珍惜。	878。	73,477,308。
5.	涨姿势。	222。	70,124,496。
6.	心塞。	132。	50,601,480。
7.	暖男。	1455。	50,564,364。
8.	买买买。	247。	41,685,392。
9.	也是蛮拼的。	643。	36,984,032。
10.	逼格。	65。	30,109,324。
11.	画面太美我不敢看。	133。	29,820,644。
12.	小鲜肉。	808。	20,879,812。
13.	那么问题来了。	309。	16,912,524。
14.	no zuo no die。	99。	16,792,928。
15.	脑洞大开。	162。	15,572,224。
16.	有钱就是任性。	855。	14,904,136。
17.	时间都去哪了。	365。	10,701,780。
18.	保证不打死你。	7。	8,124,280。
19.	我读书少你别骗我。	25。	7,282,984。
20.	习大大。	325。	6,161,256。
21.	上天台。	58。	4,379,688。
22.	什么仇什么怨。	78。	4,272,464。
23.	你家里人知道吗？	39。	3,175,480。
24.	挖掘机技术哪家强。	38。	2,033,132。
25.	APEC 蓝。	730。	1,831,056。

(数据来源: 新浪微博、人民网舆情监测室中文报刊系统。  
数据截取日期: 2014 年 1 月 1 日-12 月 31 日)。

**Tableau 17 : Liste du palmarès des 25 termes les plus populaires présentée dans le rapport 《2014 中国网络语象报告》**

Toutes ces instances médiatiques et/ou institutionnelles disposent de bases de données permettant une veille sur la situation langagière en Chine. Le magazine 《咬文嚼字》 (yǎo wén jiáo zì, ce qui signifie « peser les paroles et les mots ») met l'accent sur la popularité et se fonde sur les conseils d'experts parmi lesquels des linguistes. Il faut souligner que la sélection du magazine est faite en deux étapes. Premièrement, on fait appel aux lecteurs pour faire selon leurs constats individuels des propositions de termes populaires ou utilisées à haute

fréquence<sup>136</sup>. Deuxièmement, des linguistes et experts se réunissent pour choisir les 10 propositions finales à partir des normes linguistiques<sup>137</sup> et des critères internes du magazine qui défend le bon usage de la langue chinoise. Il est à noter que ces normes linguistiques constituent des critères importants qui déterminent directement le choix des résultats finaux. Par conséquent, certains mots et expressions élus sur la liste ne sont pas ceux qui sont attendus par le public. Cela suscite souvent la défiance des individus usagers des nouveautés langagières dès que le résultat est publié. À titre d'exemple : le magazine indique dans son billet de blog que les mots ou expressions qui ne sont pas approuvés par la presse écrite mais largement utilisés en ligne sont exclus de la sélection de 2009. Par exemple l'expression 杯具 (bēi jù, *litt.* « les verres ») beaucoup utilisée à la place de l'homophone 悲剧 (bēi jù, qui signifie « la tragédie ») est ainsi exclue malgré sa popularité en ligne.

Pour éviter d'avoir dans notre recherche un corpus qui manquerait des caractéristiques technosocio-communicationnelles, nous ajoutons aux sélections du magazine 《咬文嚼字》 (yǎo wén jiáo zì) les listes des deux autres activités qui mobilisent, elles, des statistiques qui montrent objectivement la fréquence d'utilisation sans d'adopter préalablement un jugement linguistique. Grâce à une coopération des institutions différentes, l'activité 《汉语盘点》 (hàn yǔ pán diǎn) que nous avons décrite *supra* suppose l'accès à une base de données d'environ 6 milliards de caractères chinois collectés sur les divers supports de communication en ligne : les forums célèbres, les blogs et microblogs, la presse en ligne, etc. L'obtention du résultat de chaque année dépend surtout de la technologie de traitement de l'information. Ces sources abondantes peuvent nous aider à obtenir des observations globales sur le développement néologique de la langue chinoise. La spécificité de la troisième activité qui est lancée par 人民网舆情中心 (le centre de données de l'organe de presse le *Quotidien du Peuple*) est qu'elle opère son classement de la fréquence d'utilisation à partir du microblog *Sina Weibo*, la plateforme que nous avons choisi d'étudier mais seule l'année 2014 est concernée. Après avoir

---

<sup>136</sup> De 2008 à 2017, les démarches de ce magazine pour collecter les propositions (de la 1<sup>ère</sup> étape) deviennent multiples. Non seulement les lecteurs, mais aussi les internautes sans connaître le magazine peuvent donner leur avis via des canaux variés : *Sina Weibo*, *Wechat*, *Blog* (officiel du magazine). En plus, l'année 2017, l'éditorial a établie des coopérations avec certaines universités, afin de collecter les propositions à l'aide des ressources éducatives. Cela montre que la base donnée de sélection dispose des sources abondantes.

<sup>137</sup> Nous avons essayé de contacter l'équipe éditoriale du magazine pour savoir quelles sont les normes linguistiques privilégiées mais nous n'avons pas eu de retour.

comparé les résultats de cette sélection avec ceux des deux autres activités, nous avons constaté que les termes élus coïncident avec les résultats issus des deux autres moyens de sélection. Cela vérifie indirectement une utilisation importante et étendue de certaines nouveautés langagières puisqu'elles sont repérées conjointement par des activités différentes. Cela prouve également la légitimité de nos critères de sélection du corpus de liste et nous conduit également à nous appuyer sur les mots en communs des sous-corpus issus des activités pour constituer notre corpus de liste des nouveautés langagières à observer ensuite dans des tweets. Avec ces trois sous-corpus, nous envisageons d'élaborer une liste qui, prenant en compte une comparaison inter-corpus, répondra à la fois aux deux critères de fréquence et de popularité. Cependant nous avons conscience cet échantillon de liste peut constituer une éventuelle faiblesse de notre corpus car les termes élus ne peuvent prétendre être représentatifs de la langue entière, de toutes les pratiques langagières en ligne. Nous considérons qu'ils sont significatifs.

#### **4.2.2. Observations et démarche de collecte pour le corpus (en) français**

Pour constituer le corpus français en relation avec le réseau *Twitter*, nous nous sommes posée les mêmes questions: comment nommer / retrouver les dénominations des productions langagières néologiques en français, comment définir les lieux de collecte de leurs occurrences, selon quels critères de sélection ?

Dans la chronologie de notre travail doctoral, nous avons d'abord fixé notre démarche pour le corpus chinois, constitué des mots et expressions créés et/ou utilisés par les internautes et ensuite sélectionnés par des instances institutionnelles et/ou médiatiques via plusieurs activités spécifiques et selon les critères de nouveauté, de popularité et de fréquence. Dans l'hypothèse qu'en France, les même phénomènes (nouveautés langagières et activités de sélection) étaient présents, nous avons tenté d'établir le corpus français de la même manière, autrement dit en trouvant des activités équivalentes de sélection en France en vue de collecter un corpus de façon à nous appuyer sur les mêmes critères. Cette phase de recherche nous a longuement mobilisée mais nous avons finalement adopté d'autres voies permettant toutefois de collecter des nouveautés langagières authentifiées et reconnues comme telles dans le contexte numérique. Nous allons retracer les étapes de cette phase.

La première étape a consisté à identifier les nouveautés langagières en optant pour une gamme

d'appellations possibles, proches (par traduction notamment) des appellations chinoises que nous avons repérées et utilisées. Même si ces appellations ont engendré des ambiguïtés ou des impasses, elles nous ont fourni une série de mots clés pour effectuer une première recherche générale<sup>138</sup> en France à partir de métatermes désignant les phénomènes qui nous intéressent. Après cette première recherche, nous avons remarqué que ces phénomènes étaient à l'origine de toutes sortes de publications en France parmi lesquelles des publications scientifiques qui ont nourri notre réflexion théorique, des ressources encyclopédiques<sup>139</sup> qui proposent des listes de mots et expressions nouveaux ainsi que leurs définitions, des articles de presses qui rapportent et commentent les phénomènes et certaines activités relatives aux créations lexicales, activités que nous souhaitons pouvoir exploiter en parallèle avec les activités chinoises.

Notre deuxième étape a été la quête d'activités de sélection en France équivalentes à celles que nous connaissions en Chine. Nous avons effectivement découvert certaines activités susceptibles de nous intéresser que nous allons rappeler ici et présenter plus en détail bien qu'elles ne nous aient pas permis de constituer de corpus. Il s'agit du « Festival XYZ », du « Festival du mot », de l'initiative « Dis-moi dix mots sur la toile » (édition 2016-2017). Nous allons les évoquer successivement.

1) Le « Festival XYZ du mot et du son nouveau », qui s'est tenu à partir de 2002 au Havre, avait pour vocation la promotion du néologisme et défendait une conception vivante de la

---

<sup>138</sup> La méthodologie d'ici consiste à faire des recherches générales en utilisant le moteur de *Google*. Nous avons effectué ces recherches en utilisant des mots clés traduits des appellations chinoises présentées dans les Chapitre III. Ces mots clés sont : langage, argot, parler, lexique, néologisme, nouveaux mots, web, numérique, digital, Internet, internautes, etc. Voici des exemples de combinaisons : langage d'Internet, lexique web, parler des internautes...

<sup>139</sup> Les sites partagé comme :

*Wikipédia* : [https://fr.wikipedia.org/wiki/Argot\\_Internet](https://fr.wikipedia.org/wiki/Argot_Internet);

*Wiktionnaire* : [https://fr.wiktionary.org/wiki/Annexe:Liste\\_de\\_termes\\_d%E2%80%99argot\\_Internet](https://fr.wiktionary.org/wiki/Annexe:Liste_de_termes_d%E2%80%99argot_Internet);

Les billets de blog ou des articles encyclopédiques comme par exemple :

« Top 16 des meilleurs mots absents du dico, ceux qu'on veut voir rentrer dans le Larousse » : <http://www.topito.com/top-meilleurs-mots-absents-dico>;

« Top 30 des mots chelous utilisés par les ados pour qu'on ne les comprenne pas » : <https://www.terrafemina.com/vie-privee/famille/articles/50810-top-30-des-mots-chelous-utilises-par-les-ados-pour-quon-ne-les-comprenne-pas.html>;

« Le langage des jeunes et des ados - Edition 2017-2018 ! » : <https://www.lecurionaute.fr/langage-jeunes-ados-2017-lexique-mots-termes-dictionnaire/>;

« 31 mots d'argot qu'utilisent les jeunes et que vous ne comprenez pas. Les plus de 30 ans, vous allez nous remercier ! » :

<https://www.demotivateur.fr/article/argot-nouveau-vocabulaire-expression-jeune-guide-dictionnaire-6236>;

Les dictionnaires partagés, voir note n°153 *infra* etc.

langue française. Le fondateur du Festival est Eric Donfu, ancien maire adjoint socialiste du Havre<sup>140</sup>. La dernière édition date de 2014. Chaque année, selon les propositions des participants envoyées par mail, le jury était chargé de composer une liste des mots les plus créatifs soumis ensuite au vote du public présent lors des deux soirées du mois de novembre, la première à Paris et la seconde au Havre pour sélectionner le mot de l'année<sup>141</sup> (Figure 18). C'est à travers les informations publiées sur son compte *Facebook*<sup>142</sup> que nous avons pris connaissance du Festival et des résultats de sélection.



Figure 18 : Affiche du Festival XYZ (édition 2011)

Pour cette activité, la période, le nombre et les critères de sélection nous ont posé des problèmes. Nous avons délimité notre période d'étude aux années 2009-2017. Or l'activité ne fournissait de résultats que de 2009 à 2014. Ensuite, l'objectif final de cette activité était de choisir la proposition la plus créative de l'année et la communication publique ne fait état que du mot élu

---

<sup>140</sup> Sociologue et écrivain, il est décédé en 2018.

<sup>141</sup> Nous avons constaté que les éditions récentes (probablement depuis 2011 selon les affiches du Festival disponibles) se sont simultanément déroulées au Havre et à Paris.

<sup>142</sup> Le compte Facebook du Festival XYZ :

<https://www.facebook.com/Festival-XYZ-du-Mot-et-du-Son-nouveau-142906562429000/>.

de l'année<sup>143</sup> sans que les listes au sein desquelles il était choisi soient connues. Un problème de nombre se posait. En outre nous n'avons pas eu accès aux critères de proposition et de sélection et n'avions pas de moyen d'identifier les sources des propositions ; autrement dit nous ne pouvions savoir s'il s'agissait d'un mot ou une expression néologique influente créé et connaissant un certain niveau de diffusion dans la société ou d'une proposition fabriquée spécialement par les participants pour montrer les possibilités créatives de la langue française. Nous n'avons obtenu aucune réponse à nos questions ni aucun retour de la part des responsables du Festival.

2) Le « Festival du Mot » est aussi une activité destinée à élire le mot de l'année, créée par Marc Lecarpentier<sup>144</sup> à la Charité-sur-Loire en 2005 (Figure 19). Il propose deux catégories de sélection : un mot de l'année du public et un mot de l'année du jury<sup>145</sup>. Nous avons obtenu sur le site du festival<sup>146</sup> beaucoup plus d'informations que pour le précédent : les mots élus de chaque année de 2005 à 2018 ainsi que la liste des mots candidats. Pour certaines années, des explications portant sur le choix et les pourcentages de vote sont fournies.

---

<sup>143</sup> « se faire électronique » en 2002, « humanicide » en 2003, « chaudard » en 2004, « désoiffer » en 2005, « ordinosore » en 2006, « photophoner » en 2007, « bonjoir » en 2008, « aimeuse » en 2009, « phonard » en 2010, « attachiant » en 2011, « watture » en 2012, « Plénior » en 2013, « médicalmant » et « casse-crotte » en 2014.

<sup>144</sup> Il est président de l'association Mot-et-MOTS, association francophile qui « encourage la diffusion de paroles venues d'ailleurs, les langues étrangères enrichissant la langue française et ses accents savoureux », selon sa biographie publiée sur *France-Inter* : <https://www.franceinter.fr/personnes/marc-lecarpentier>.

<sup>145</sup> Composé de journalistes, d'écrivains et de spécialistes de la langue française.

<sup>146</sup> Suite à des problèmes financiers et de direction, l'activité « Aux quatre coins du Mot » remplace le Festival du mot depuis 2019. De ce fait, l'ancien site n'existe plus et nous n'avons plus d'accès aux plus de détails de l'activité et aux résultats des sélections que nous n'avions pas tous sauvegardés.



Figure 19 : Affiche du Festival du Mot (édition 2018)

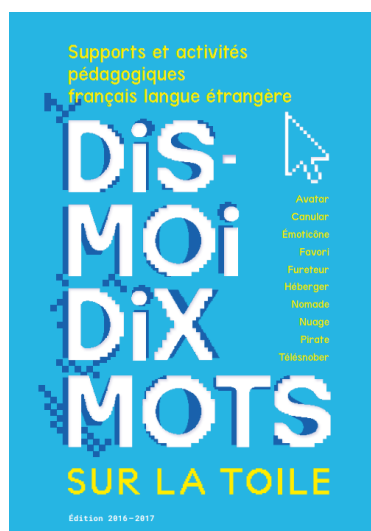
Après avoir pris connaissance de ces éléments, nous avons compris que les résultats concernaient plutôt le mot le plus significatif de l'année par rapport à des questions sociétales. Par exemple, le mot du jury de l'année 2018 est « femme », ou encore en 2016, le public et le jury élisent le mot « réfugiés ». Donc les critères de fréquence, de popularité et de relation au contexte numérique ne sont pas présents, ni même la néologie malgré la présence de mots comme « selfie », « twitter ». Les organisateurs du festival n'ont pas donné suite à notre demande d'information.

3) « Dis-moi dix mots » est une opération nationale qui relève d'un niveau institutionnel puisqu'elle est organisée par le Ministère de la Culture et de la Communication en collaboration avec plusieurs institutions et partenaires francophones<sup>147</sup>. Elle se déroule tout au long de l'année scolaire et a pour but de sensibiliser les élèves à la langue française d'une manière ludique grâce à son concours, en France et à l'étranger dans une perspective de rayonnement de la

---

<sup>147</sup> Voici quelques-unes des institutions : l'Agence pour l'enseignement français à l'étranger, la Mission laïque française, la Ligue de l'enseignement, le Réseau Canopé, les éditions Le Robert, les éditions Larousse, l'École des lettres, etc. Et les partenaires francophones du réseau OPALÉ : la France, la Belgique, le Québec, la Suisse et l'Organisation internationale de la Francophonie (qui représente 84 États et gouvernements).

francophonie<sup>148</sup>. La thématique est différente chaque année<sup>149</sup>. C'est l'édition 2016-2017 intitulée « Dis-moi dix mots sur la toile »<sup>150</sup> (Figure 20) qui a attiré notre attention car elle se donnait pour objectif de montrer « la manière dont le français s'adapte à des technologies et à des usages numériques en constante évolution, en s'enrichissant des sens nouveaux, sans qu'il soit nécessaire d'emprunter à une autre langue »<sup>151</sup>.



**Figure 20 : Couverture de la brochure pédagogique « dis-moi dix mots » (édition 2016-2017)**

Donc, ce sont dix mots nouveaux ou sens nouveaux « purement français » associés aux pratiques d'Internet et aux nouvelles technologies. Cette liste semble pour nous la plus proche de notre sujet de recherche. Mais il est à noter que cette activité est organisée à l'initiative de DGLFLF<sup>152</sup> qui est chargée de favoriser la diversité linguistique dans une perspective

---

<sup>148</sup> Concours de création littéraire ouvert aux élèves des écoles, collèges et lycées, y compris ceux relevant de l'Agence pour l'enseignement français à l'étranger (AEFE) et de la mission laïque française ainsi qu'aux classes d'établissements étrangers (FLE-FLS). Le concours est ouvert aux inscriptions jusqu'à fin janvier et la publication du résultat est au cours du février. Une remise du prix est organisée en mars lors de la Semaine de la langue française et de la francophonie.

<sup>149</sup> Selon les informations sur le site officiel : <http://www.dismoidixmots.culture.fr/>, il se peut que la première édition date de 2002. Voici quelques thématiques précédentes :

2004-2005 : « Les mots scientifiques, hommage à Jules Verne » ;

2006-2007 : « Les mots migrants » ;

2008-2009 : « Des mots pour dire demain » ;

2012-2013 : « Dis moi dix mots semés au loin ».

<sup>150</sup> Les élus de l'édition 2016-2017 sont les suivants : avatar, canular, émoticône, favori, fureteur, héberger, nomade, nuage, pirate, télésnober.

<sup>151</sup> Voir : <http://www.dismoidixmots.culture.fr/2017/ressources/thematique-dix-mots-2016-2017.html>.

<sup>152</sup> La Délégation Générale à la Langue Française et aux Langues de France.



normative. Nous nous demandons si ces créations qui montrent la vivacité du français sont plutôt des néologismes imposés par des commissions terminologiques, ou encore des néologismes « attestés » approuvés officiellement qui ne couvrent pas de ce fait tous nos intérêts de recherche. Mais surtout le thème non-consécutif ne peut nous fournir qu'une seule liste de dix choix pour une seule année alors que notre période d'étude est de neuf années. De plus, cette activité ne renseigne pas une période de sélection précise qui indiquerait l'année d'apparition du mot. Enfin, mais c'est le cas pour la plupart des activités, les mots nouveaux sont liés aux nouvelles technologies de communication et aux RSN mais comme thématique pas forcément comme espace de création.

Comme les mots choisis ne répondaient pas forcément aux critères de popularité et de fréquence d'utilisation et ils sont proposés par des organes différents auprès desquels il n'est pas aisé pour nous de nous renseigner, nous n'avons finalement pas pris contact avec ces organes concernés pour obtenir des renseignements.

Au terme de l'exploration infructueuse de ces pistes, nous avons constaté que les activités de sélection en France étaient peu similaires à celles de Chine que nous avons retenues et nous avons décidé de travailler à partir d'autres ressources : les ressources encyclopédiques partagées et les articles de presse. Pour ce qui est des ressources encyclopédiques destinées au grand public et des sites ou dictionnaires partagés et accessibles en ligne<sup>153</sup>, nous avons remarqué que ces ressources et publications comportent une grande quantité de propositions éditées par les internautes. Il faudrait mener une réflexion approfondie sur le fait de savoir si de telles sources sont recevables d'un point de vue scientifique. Le problème du tri et du traitement des entrées que nous pouvions recueillir par ce moyen se posait car les propositions apparaissent par ordre alphabétique (comme c'est usuel dans les dictionnaires) sans être accompagnées de critères de catégorisation qu'il nous aurait fallu établir et appliquer. Et comme nous l'avons déjà

---

<sup>153</sup> Le dictionnaire *BLAZZ* se présente comme un dictionnaire d'argot français, créé par les internautes pour tenter de comprendre la jeunesse d'aujourd'hui : <https://www.blazz.fr/>;

Le dictionnaire *Urban* propose des mots argotiques sur Internet : <http://www.urbandico.com/>;

Le dictionnaire *Sensagent* qui propose une liste de termes d'argot Internet est hébergé par le journal *le Parisien* : <http://dictionnaire.sensagent.leparisien.fr/index.html>;

« Le dico des mots, le dictionnaire vraiment pas académique: un dictionnaire collaboratif créé par les internautes pour proposer des expressions qui font rage chez les potes » : <https://dico-des-mots.com/>;

Dico2rue - Le dictionnaire collaboratif de l'argot français : <http://www.dico2rue.com/>;

La Parlure - Le dictionnaire collaboratif du français parlé : <https://www.laparlure.com/>.

signalé ni la popularité, ni la fréquence ni le caractère néologique ne peuvent être conjointement et absolument vérifiés.

Nous ne pouvions nullement en tant que chercheuse individuelle constituer ni traiter pour le français des données quantitatives comparables à celles que les institutions chinoises traitent. D'ailleurs notre orientation méthodologique est celle du qualitatif-significatif et non du représentatif. Nous nous sommes donc résolue à procéder différemment pour le français mais en gardant un filtre médiatique et institutionnel et en combinant deux types de publications qui s'inscrivent dans un métadiscours autorisé sur la langue.

Nous avons été attirée par les articles de presse généraliste qui présentent régulièrement lors de la parution des éditions annuelles des dictionnaires de langue en France qui actualisent leur contenu, certaines nouvelles entrées de ces dictionnaires (Figure 21). L'enregistrement dans les dictionnaires de ces nouvelles entrées avec la rigueur et les moyens de la lexicographie font le lien entre la science et ses méthodes et le grand public. Cette actualisation annuelle suppose l'examen rapide de données quantitativement importantes et qualitativement significatives, et manifeste à la fois l'influence de l'évolution de la langue sur les publications qui la décrivent, et l'influence des dictionnaires sur la langue et ses représentations chez les locuteurs. En effet, les dictionnaires sont une publication autorisée et de référence. La perspective néologique et les examens et analyses qu'elle suppose sont prises en compte et reconnues par les experts chargés de la tâche de sélection et de description. La périodicité annuelle convenait bien au créneau temporel que nous avons défini.

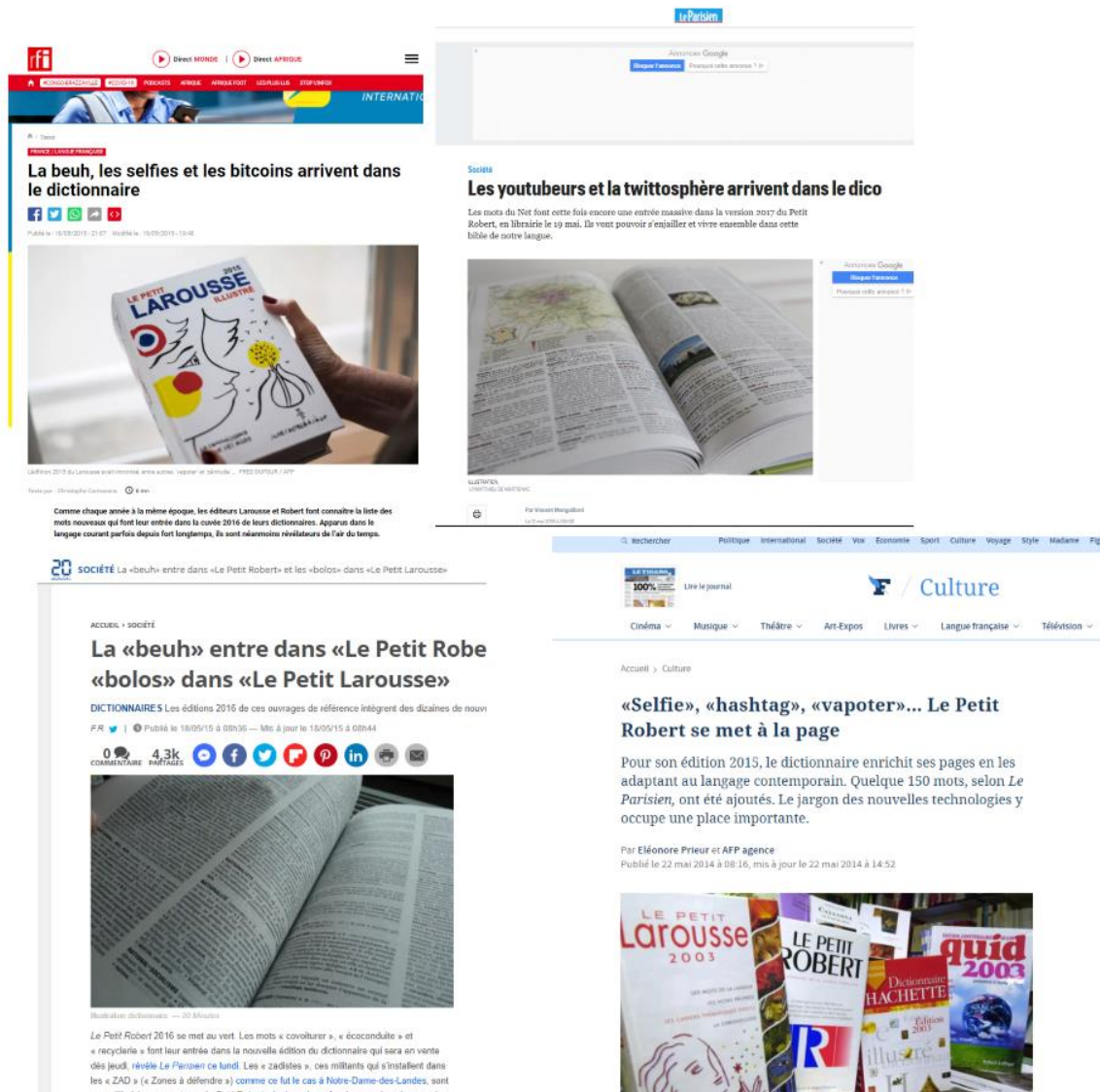


Figure 21 : Captures d'écran d'articles de presse qui présentent les nouvelles entrées des dictionnaires

Pour accéder directement aux nouvelles entrées des dictionnaires pour les années correspondant aux bornes temporelles de notre corpus, nous nous sommes appuyée sur le travail réalisé par le club d'orthographe de Grenoble<sup>154</sup> qui dresse des listes et les met à disposition du public pour montrer d'une manière synthétique cet enrichissement qui varie selon les éditions (Tableau 18).

<sup>154</sup> Né initialement de la rencontre des étudiants pour préparer le rendez-vous annuel des Dicos d'or depuis 2001, le club se constitue en association en 2006 et est hébergé par l'université Stendhal-Grenoble 3 à partir 2007. L'objectif du club est de jouer avec les mots d'une manière joyeuse. Il organise régulièrement les activités de dictée. Le lien du site est le suivant : <https://orthogrenoble.net/>. Les informations sur les nouvelles entrées des dictionnaires ont été obtenues en comparant intégralement les éditions successives d'un même dictionnaire.

## Petit Larousse

▼ Mots nouveaux
Petit Larousse 2021 (91 articles + 2 mots cachés)
Petit Larousse 2020 (81 + 1)
Petit Larousse 2019 (81 + 7)
Petit Larousse 2018 (96 + 2)
Petit Larousse 2017 (101 + 5)
Petit Larousse 2016 (113 + 13)
Petit Larousse 2015 (89 + 12)
Petit Larousse 2014 (99 + 0)
Petit Larousse 2013 (87 + 1)
Petit Larousse 2012 (1882 + 19)
Petit Larousse 2011 (82 + 2)
Petit Larousse 2010 (94 + 0)
Petit Larousse 2009 (109 + 1)
Petit Larousse 2008 (65 + 0)
Petit Larousse 2007 (113 + 1)
Petit Larousse 2006 (103 + 1)
Petit Larousse 2005 (301 + 0)
Petit Larousse 2004 (56 + 0)
Petit Larousse 2003 (103 + 0)
Petit Larousse 2002 (134 + 0)
Petit Larousse 2001 (113 + 0)
Petit Larousse 2000 (288 + 5)
Petit Larousse 1999 (37 + 0)
Petit Larousse 1998 (1451 + 19)

## Petit Robert

▼ Mots nouveaux
Petit Robert 2021 (87 articles + 22 mots cachés)
Petit Robert 2020 (107 + 27)
Petit Robert 2019 (131 + 11)
Petit Robert 2018 (122 + 9)
Petit Robert 2017 (118 + 15)
Petit Robert 2016 (134 + 13)
Petit Robert 2015 (105 + 10)
Petit Robert 2014 (109 + 30)
Petit Robert 2013 (101 + 7)
Petit Robert 2012 (93 + 4)
Petit Robert 2011 (151 + 26)
Petit Robert 2010 (172 + 18)
Petit Robert 2009 (202 + 16)
Petit Robert 2008 (423 + 39)
Petit Robert 2007 (421 + 92)
Petit Robert 2006 (70 + 3)
Petit Robert 2005 (64 + 3)
Petit Robert 2004 (79 + 7)
Petit Robert 2003 (130 + 12)
Petit Robert 2002 (96 + 12)
Petit Robert 2001 (113 + 14)
Petit Robert 2000 (41 + 12)
Petit Robert 1999 (48 + 3)
Petit Robert 1998 (72 + 8)

**Tableau 18 : Récapitulatif synthétique des nouvelles entrées de deux dictionnaires selon le club d'orthographe de Grenoble**

Nous avons pu constater que d'une part les nouvelles entrées étaient très nombreuses, et d'autre part qu'elles variaient selon les dictionnaires<sup>155</sup>. Et nous avons eu l'idée de conserver le recours aux articles de presse qui signalent les publications des dictionnaires avec une série d'exemples en les utilisant à nouveau comme des filtres afin, là encore, d'avoir affaire à des échantillons comme dans le cas des nouveautés langagières chinoises et d'intégrer le discours médiatique au processus de repérage, et éventuellement d'évaluation de l'évolution lexicale.

En effet, les organes de presse qui consacrent presque rituellement des articles à l'actualisation annuelle des dictionnaires au moment de la parution des nouvelles éditions ne manquent pas de faire de ces articles une occasion de réflexion sur la société et ses évolutions et une occasion de réflexion sur la langue. De ce fait les auteurs des articles citent et expliquent certaines entrées.

---

<sup>155</sup> Nous reviendrons de manière plus précise sur ces points dans le chapitre V. Nous avons retenu deux dictionnaires *Le Petit Robert* et *Le Petit Larousse*.

Ainsi cette diffusion, via les articles de presse qui rendent compte du travail des dictionnaires, en mettant en relief certaines entrées nous permet d'être assurée du caractère néologique. Certes, elle ne prouve pas absolument la popularité et la fréquence d'utilisation de ces nouvelles entrées mais en donne tout de même des indices forts.

Une question reste irrésolue : nous nous demandons s'il peut arriver que certaines productions néologiques soient exclues de la sélection des dictionnaires en raison des normes linguistiques qui sont une des considérations qui intervient en Chine, notamment dans les sélections du magazine 《咬文嚼字》(yǎo wén jiáo zì). Nous avons l'impression que selon les institutions éducatives ou culturelles, les médias et même les dictionnaires, il existe plusieurs systèmes de normes possibles. Nous n'approfondirons pas cette question qui mériterait un développement à part entière.

#### **4.3. Une démarche parallèle mais différenciée pour un double corpus**

Nous avons donc résolu de manière empirique et adaptée la démarche de collecte de productions langagières en ne procédant pas de la même manière pour le contexte chinois et le contexte français, mais en retenant globalement les principes de nouveauté, fréquence, popularité et de relation forte avec les dispositifs numériques faute de pouvoir détecter absolument ces dispositifs comme espace de création originelle. Comme nous l'avons dit en introduction de ce chapitre, la collecte doit déboucher sur deux types de corpus (et pour chacun des types un sous-corpus en français et un sous-corpus en chinois) : un corpus qui établit une liste d'entrées lexicales, de façon à fournir un échantillon et d'en faire l'analyse linguistique et communicationnelle ; un corpus qui rassemble des discours en ligne comportant une partie des entrées lexicales retenues et analysées, de façon à fournir un échantillon qui témoigne de la circulation et permettre une analyse discursive et communicationnelle.

Pour constituer dans les deux langues un corpus d'échantillon de nouveautés langagières qui sera le point de départ de la constitution dans les deux langues et sur les deux RSN d'un corpus d'échantillon discursif de tweets, nous avons donc opté pour deux démarches spécifiques.

Nous considérons que les activités annuelles organisées par des institutions chinoises visant à élire de nouveaux mots et expressions liés à l'Internet répondent aux critères de nouveauté, de

popularité et de fréquence d'utilisation (ce qui va de pair avec une large diffusion), en nous proposant des choix représentatifs parmi les nouveautés langagières existantes. Sur cette base nous avons pu relever une liste de 175 résultats<sup>156</sup> de sélection pour les années qui nous intéressent, venant de trois institutions : Le magazine 《咬文嚼字》 (yǎo wén jiáo zì), l'activité 《汉语盘点》 (hàn yǔ pán diǎn), 《人民日报》 (rén mín rì bào, le journal *Le Quotidien du Peuple*). Il nous appartiendra, selon une démarche de tris successifs motivés que nous expliquerons, de sélectionner, au sein de ce corpus d'échantillon, un autre échantillon pour étudier des tweets.

Dans la mesure où nous n'avons pas pu trouver d'activités institutionnelles équivalentes en France, en conservant les trois critères de nouveauté, de popularité et de fréquence mentionnés *supra*, nous nous sommes orientée vers une constitution de listes d'entrées à partir des nouvelles entrées de deux dictionnaires de langues (*Petit Robert* et *Petit Larousse*) telles que rapportées par les articles de quatre médias français (*Le Figaro*, *20 Minutes*, *Le Parisien*, *RFI*)<sup>157</sup>. Nous avons ainsi constitué deux sous-corpus : l'un recueille 138 nouvelles entrées du dictionnaire *Petit Robert* et l'autre 234 résultats du dictionnaire *Petit Larousse*. Nous allons justifier le choix de ces résultats et donner les détails du corpus dans le chapitre suivant.

Le passage d'un corpus de listes à un corpus de tweets dans une perspective qualitative dont nous détaillerons les principes dans le chapitre VI consacré à l'analyse des tweets exige un nouveau tri. En effet nous avons une liste de 175 entrées pour le chinois et de 371 pour le français.

Vu la quantité importante des nouveautés langagières recueillies dans notre corpus de liste, il est impossible d'observer la circulation de chaque entrée de la liste dans tous les tweets qui l'utilisent. Il nous faudra donc opérer une double sélection : celle des entrées du corpus listes que nous conservons et celle des tweets dont le texte comporte ces entrées que nous retenons. Pour choisir le nombre des nouveautés langagières à observer sur les microblogs, comme le corpus de listes est constitué de plusieurs sous-corpus provenant de sources différentes, nous

---

<sup>156</sup> Voir l'annexe n°2 et les tableaux du chapitre V.

<sup>157</sup> Voir l'annexe n°6c.

avons privilégié les entrées communes aux deux sous-corpus chinois toutes années mêlées<sup>158</sup> et les entrées des deux sous-corpus français (les entrées de deux dictionnaires) qui sont remarquées par leurs procédés de création toutes années mêlées, en visant, pour l'un et pour l'autre, une vingtaine d'entrées. Une fois les entrées retenues pour le chinois et pour le français nous les avons examinées de près sur le plan linguistique et communicationnel de façon à affiner les critères de choix définitifs en à n'en sélectionner qu'une demi-douzaine environ nous paraissant à la fois diversifiées et pertinentes.

Enfin une dernière phase de tri est intervenue, cette fois pour collecter des tweets contenant les nouveautés langagières retenues, autrement dit, des occurrences d'utilisation des nouveautés langagières sur les microblogs. Pour trouver ces occurrences, nous nous sommes servi du moteur de recherche interne de *Sina Weibo* et de *Twitter*. En saisissant chaque entrée dans le moteur de recherche, celui-ci affiche tous les résultats en sachant que la plateforme du microblog actualise en temps réel. Le nombre de résultats n'est pas stable et on peut imaginer qu'il augmente sans cesse. En considérant qu'il est impossible d'épuiser un corpus dynamique ni d'analyser toutes les occurrences d'une nouveauté langagière employée à haute fréquence et de retracer les parcours, nous avons admis de limiter également le nombre de tweets retenus à 10 par entrée. Nous y reviendrons dans le chapitre VI.

## **Conclusion**

Nous nous sommes attachée dans ce chapitre à reparcourir le cheminement qui nous a permis de constituer les corpus que nous allons analyser dans les chapitres suivants. Nous avons évoqué nos questionnements, nos difficultés et avons tenté de justifier nos choix et d'expliquer comment nous avons procédé notamment pour établir les conditions d'une comparaison valide et pour disposer d'échantillons significatifs. Nos corpus sont de deux types, corpus de listes et corpus de tweets mais nous avons surtout évoqué ici les démarches de collecte des corpus de listes car c'est l'étape préalable indispensable à la constitution des corpus de tweets. Nous avons notamment expliqué comment la perspective comparative a dû s'ajuster à des contextes différents et comment nous avons fait intervenir pour la collecte des nouveautés langagières des

---

<sup>158</sup> Cette opération de sélection a dégagé 21 entrées parmi lesquelles nous avons fait un nouveau choix de sept entrées.

instances autorisées (institutions, dictionnaires, médias) qui sont à la fois pour nous une caution et un moyen d'accéder à une collecte déjà réalisée selon les critères de nouveauté, de fréquence et de popularité que nous avons retenus.

Cet aperçu sommaire et chronologique de nos démarches successives sera repris pour certains de ses aspects et développé dans les trois prochains chapitres. Le chapitre V sera consacré à la présentation des corpus de listes et à leur analyse pour le français et le chinois. Le chapitre VI procèdera tout d'abord à une mise au point théorique et méthodologique sur l'étude des tweets en sciences du langage puis à une description de la démarche de constitution des corpus pour le chinois et le français. Le chapitre VII développera l'analyse des deux corpus de tweets.





## CHAPITRE V

### Constitution et analyse des corpus de liste

#### Introduction

Dans ce chapitre, nous allons présenter le contenu du corpus de liste dans les deux langues. Il est à noter que la constitution du corpus français – aussi bien pour le corpus de listes que pour le corpus de tweets qui fera l’objet des chapitres suivants –, intervient après l’établissement du corpus chinois. Etant une locutrice non native du français et ayant été en contact plus tôt avec les nouveautés langagières chinoises et le contexte numérique chinois – en tant que témoin et usagère à la fois –, nous sommes guidée par nos expériences et avons débuté en premier la constitution du corpus chinois. C’est pourquoi nous suivons cet ordre (chinois puis français). Les démarches adoptées, les activités sélectionnées n’échappent certes pas à la subjectivité du chercheur mais cela nous a aussi permis de nous appuyer sur certains critères déjà établis tout en les ajustant ensuite pour le choix du corpus français.

Les corpus d’échantillons une fois établis seront ensuite mobilisés pour une analyse linguistique. Nous nous interrogeons sur leurs caractères néologiques. De ce fait, nous allons observer les procédés de création de ces nouveautés langagières et essayons de leur affecter une catégorie en nous référant aux travaux existants, en particulier une typologie des néologismes. Ensuite, une comparaison entre les deux corpus du point de vue des langues chinoise et française sera effectuée pour vérifier s’il existe les mêmes tendances de formation malgré les différences linguistiques.

#### 5.1. Constitution du corpus de liste chinois

Nous allons présenter respectivement les trois sous-corpus qui sont issus des activités de sélection en Chine. Les nouveautés langagières chinoises collectées seront présentées dans un tableau par sous-corpus (Tableau 19, Tableau 20, Tableau 21), écrites en caractères et en *pinyin* et accompagnées d’une proposition de traduction en français de notre part. Ces traductions sont

informatives et n'ont pas été élaborées avec une grande précision et profondeur car le but n'était pas de traduire mais de rendre intelligibles ces mots et expressions<sup>159</sup>.

### **5.1.1. Corpus A : 《咬文嚼字》十大流行语 (les 10 mots et expressions les plus populaires de l'année)**

Nous avons déjà abordé le choix de la période d'observation dans le chapitre IV : cette période est délimitée de 2009 à 2017 en prenant en considération la mise en service de *Sina Weibo* en 2009 et une modification technique importante sur *Twitter* (la suppression de la limite des 140 signes dans le message) en 2017. Le magazine a commencé à proposer une sélection annuelle successive depuis 2008. Nous avons pu collecter 90 résultats pour la période délimitée.

---

<sup>159</sup> Pour faciliter la compréhension et rendre compte des aspects néologiques, nous avons ajouté dans les tableaux 19, 20, 21 une traduction littérale en français et/ou une explication contextuelle pour certaines nouveautés langagières qui nous paraissaient l'exiger. Quand il y a lieu de distinguer sens littéral et sens figuré, la traduction du sens littéral est présentée en premier entre crochets et le sens figuré se trouve en dessous, suivi par l'explication contextuelle mise elle, entre parenthèses (si celle-ci est indispensable). Précisons que pour éviter des redondances, nous n'avons pas renseigné les 21 éléments communs aux trois sous-corpus qui sont l'échantillon de départ pour la constitution du corpus de tweets et qui font l'objet d'une présentation très détaillée au chapitre VI (Tableau n°27 du chapitre VI). Pour ces éléments, nous nous bornons à indiquer le numéro du terme qui renvoie au tableau du chapitre VI.

2009	2010	2011	2012	2013	2014	2015	2016	2017
不差钱 <i>bù chà qián</i>  ne manque pas d'argent	给力 <i>gěi lì</i>  c'est génial	亲 <i>qīn</i>  [chéri(e)] <sup>160</sup> cher client (conversation commerciale en ligne) <sup>161</sup>	正能量 <i>zhèng néng liàng</i>  qui donne une énergie positive	中国梦 <i>zhōng guó mèng</i>  rêve chinois (ambition pour le pays issu du discours politique du président Xi <sup>162</sup> )	顶层设计 <i>dǐng céng shè jì</i>  traiter d'un point de vue général en visant l'efficacité (disc.politique)	获得感 <i>huò dé gǎn</i>  satisfaction d'avoir obtenu , gagné qqch. (pour le peuple : disc.politique)	洪荒之力 <i>hóng huāng zhī lì</i>  Voir n° 12 * <sup>163</sup>	不忘初心 <i>bù wàng chū xīn</i>  ne pas oublier les objectifs initiaux dans l'action (disc.politique)
躲猫猫 <i>duǒ māo māo</i>  [jeu de cache cache] (connotation grave : suspicion en lien avec actualité du décès d'un suspect en prison)	神马都是浮云 <i>shén mǎ dōu shì fú yún</i>  rien n'a d'importance	伤不起 <i>shāng bù qǐ</i>  difficile à supporter	元芳，你怎么看 <i>yuán fāng, nǐ zěnme kàn</i>  Voir n° 1*	光盘 <i>guāng pán</i>  finir toute l'assiette (initiative gouvernementale pour éviter gaspillage)	新常态 <i>xīn cháng tài</i>  nouvelle norme (notion économique)	互联网+ <i>hù lián wǎng +</i>  l'autoroute de l'information (disc. politique 1er ministre chinois Li keqiang)	吃瓜群众 <i>chī guā qún zhòng</i>  Voir n° 13 *	砥砺前行 <i>dǐ lì fèn jìn</i>  avancer malgré les difficultés (disc. politique)

<sup>160</sup> Les éléments entre crochets correspondent à une traduction littérale quand celle-ci est différente du sens global ou du nouveau sens du mot ou de l'expression.

<sup>161</sup> Les éléments entre parenthèses indiquent brièvement le contexte ou le domaine thématique s'il est nécessaire à la compréhension.

<sup>162</sup> La plupart des éléments issus du discours politique viennent des propos du président Xi Jinping ou du gouvernement, tels qu'ils sont repris par les médias.

<sup>163</sup> Les éléments suivis d'astérisque sont étudiés de manière développée dans le chapitre VI *infra*, Tableau n°27.

<p>低碳 dī tàn</p> <p>baisse du carbone (notion écologique ; disc. politique)</p>	<p>围脖 wéi bó</p> <p>petite écharpe ; cache-cou  (nom du microblog dans l'usage courant)</p>	<p>Hold住 Hold zhù</p> <p>tenir qqch. sous contrôle</p>	<p>舌尖上 shé jiān shàng</p> <p>Voir n° 2 *</p>	<p>倒逼 dào bāi</p> <p>Stimuler d'une manière inverse</p>	<p>打虎拍蝇 dǎ hǔ pāi yíng</p> <p>[frapper les tigres comme les mouches]  (disc. pol. ; anticorruption concerne tout le monde, les cadres comme les agents)</p>	<p>颜值 yán zhí</p> <p>évaluation d'après une belle apparence</p>	<p>工匠精神 gōng jiàng jīng shén</p> <p>être créatif, professionnel et sérieux dans le travail  [L'esprit d'artisanat]  (disc. politique maîtrise des techniques)</p>	<p>共享 gòng xiǎng</p> <p>partager  (notion économique; système de location ; pratique sociale)</p>
<p>被就业 bèi jiù yè</p> <p>être employé à son insu  (commentaire ironique d'actualité)</p>	<p>围观 wéi guān</p> <p>regarder en tant que spectateur ou témoin passif</p>	<p>我反正信了 wǒ fǎn zhèng xìn le</p> <p>[je suis convaincu]  je ne crois pas  (antiphrase)</p>	<p>躺着也中枪 tǎng zhe yě zhōng qiāng</p> <p>Voir n° 3*</p>	<p>逆袭 nì xí</p> <p>Voir n° 5 *</p>	<p>断崖式 duàn yá shì</p> <p>comme une falaise qui s'écroule  (disc. politique à propos de qch qui baisse soudainement, ex. : les prix)</p>	<p>宝宝 bǎo bǎo</p> <p>Voir n° 10*</p>	<p>小目标 xiǎo mù biāo</p> <p>Voir n° 14  *</p>	<p>有温度 yǒu wēn dù</p> <p>qqch. ou qqn qui procure le bien- être  (disc. politique à propos de la construction d'une ville agréable, plus humaine)</p>

<p>裸</p> <p>luǒ</p> <p>[être nu]</p> <p>ne rien préparer ; être dégage de l'aspect matériel</p>	<p>二代</p> <p>èr dài</p> <p>[2° génération]</p> <p>qui ne fait pas d'effort mais bénéficie d'avantages qu'offre la famille</p>	<p>坑爹</p> <p>kēng diē</p> <p>[tromper son père]</p> <p>Quelle tromperie !</p>	<p>高帅富</p> <p>gāo shuài fù</p> <p>Voir n° 4 *</p>	<p>微 XX</p> <p>wēi XX</p> <p>micro, notamment pour la taille</p>	<p>你懂的</p> <p>nǐ dǒng de</p> <p>c'est sûr que tu as compris</p> <p>(pour indiquer qu'il ne convient pas d'en parler en public)</p>	<p>创客</p> <p>chuàng kè</p> <p>qqn de très doué pour la technique qui aime partager</p> <p>(trad.anglais <i>culture maker</i>)</p>	<p>一言不合就 XX</p> <p>yī yán bù hé jiù XX</p> <p>si on n'est pas d'accord, tout de suite on va faire XX (qqch.)</p>	<p>流量</p> <p>liú liàng</p> <p>acteurs ou célébrités qui ont beaucoup de fans et d'influence (valeur attraction éventuellement commerciale)</p>
<p>纠结</p> <p>jiū jié</p> <p>[plantes ou fils qui s'entrelacent]</p> <p>hésiter</p>	<p>拼爹</p> <p>pīn diē</p> <p>[comparer les pères]</p> <p>profiter des avantages des pères</p>	<p>卖萌</p> <p>mài méng</p> <p>montrer qu'on est mignon</p> <p>(cf. n° 9*)</p>	<p>中国式</p> <p>zhōng guó shì</p> <p>à la chinoise</p> <p>(titre d'un feuilleton ex. : « Divorce à la chinoise »)</p>	<p>大 V</p> <p>dà V</p> <p>comptes vérifiés sur microblog qui ont une influence importante</p>	<p>断舍离</p> <p>duàn shě lí</p> <p>il faut abandonner qqch., mener un style de vie simple</p> <p>(emprunt au jap.)</p>	<p>脑洞大开</p> <p>nǎo dòng dà kāi</p> <p>[il y a un grand trou dans le cerveau]</p> <p>libérer l'imagination</p>	<p>友谊的小船，说 翻就翻</p> <p>yǒu yì de xiǎo chuán, shuō fān jiù fān</p> <p>Voir n°15*</p>	<p>可能 XXX 假 XXX</p> <p>kě néng XXX jiǎ XXX</p> <p>possible que qqch. XXX soit une contrefaçon (plaisanterie sur la mauvaise qualité ou qqch. qui n'est pas conforme à l'attente)</p>

钓鱼 diào yú  [pêcher] Tendre un piège à qqn	控 kòng  aimer beaucoup qqch. (emprunt phonétique au jap.)	吐槽 tù cáo  [vomir ; ustensile pour gaver] se plaindre ; commenter (emprunt au jap.)	压力山大 yā lì shān dà  [Alexandra : paronyme]  stresser beaucoup	女汉子 nǚ hàn zi  Voir n°6*	失联 shī lián  perdre contact (dialecte)	任性 rèn xìng  [capricieux] faire comme on veut, être désinvolte	供给侧 gòng jǐ cè  [fournir un côté] entité qui offre  (notion économique)	油腻 yóu nì  [gras] Voir n°19 *
秒杀 miǎo shā  [seconde tuer] offre de dernière minute	帝 dì  [L'empereur] qqn talentueux	气场 qì chǎng  [champ d'énergie autour des personnes] charme ou caractère de qqn (notion de physique)	赞 zàn  [génial]  j'aime ( <i>like</i> sur microblog)	土豪 tǔ háo  Voir n°7 *	神器 shén qì  [objet sacré] outils performants efficaces qui facilitent une tâche	剁手党 duò shǒu dǎng  [gang qui coupe les mains] qqn qui achète trop de choses, qui aime acheter	葛优躺 gē yōu tǎng  Voir n° 16 *	尬 gà  gênant

蜗居 wō jū  [maison d'escargot]  Se loger dans un espace exigu pour des raisons financières, souvent dans des grandes villes  (conditions de vie des jeunes)	达人 dá rén  [qqn de rationnel]  qqn qui a du talent	悲催 bēi cuī  triste	最美 zuì měi  [le plus beau]  lié à actualité pour décrire le caractère positif héroïque de qqn	奇葩 qí pā  [jolie fleur] qqn de bizarre  (ironie)	高大上 gāo dà shàng  Voir n°8 *	网红 wǎng hóng  [Internet rouge]  influenceur (rouge associé à célébration)	套路 tàò lù  Voir n°17 *	怼 duǐ  Voir n°20 *
蚁族 yǐ zú  [fourmis] jeunes diplômés en situation précaire vivant regroupés comme dans une fourmilière en banlieue	穿越 chuān yuè  [traverser un espace physique]  voyager dans le temps (science-fiction)  être perdu (sentiment)	忐忑 tǎn tè  s'inquiéter de qqch.	接地气 jiē dì qì  [toucher la terre]  connaître qqch. c'est le connaître dans le réel au quotidien	点赞 diǎn zàn  [j'aime, je like]  dire des choses positives sur qch	萌萌哒 méng méng da  Voir n°9 *	主要看气质 zhǔ yào kàn qì zhi  Voir n°11 *	蓝瘦，香菇 lán shòu, xiāng gū  Voir n°18 *	打 call dǎ call  Voir n°21*

Tableau 19 : Liste du corpus A



A première vue, en ce qui concerne la structure, la plupart des nouveautés langagières sont composées de deux ou trois (parfois quatre) caractères. Les caractères simples et les longues expressions sont présents également mais ne sont pas nombreux. Il est intéressant de constater que le magazine a choisi de présenter la version compacte plutôt que la version complète d'une nouveauté langagière. Par exemple, nous pouvons repérer 高大上 (haut de gamme) (2014) dans le tableau et non la version 高端大气上档次; 舌尖上 (sur le bout de la langue) (2012) au lieu de 舌尖上的中国 (la Chine sur le bout de la langue) et 小目标 (petit objectif) (2016) au lieu de 定个小目标 (établir un petit objectif). Cela pourrait s'expliquer par l'un des critères de sélection que nous avons repéré dans leurs billets de blog : la brièveté .

La diversité thématique est constatée : société (蜗居 wō jū, 2009), quotidien (纠结 jiū jié, 2009), cultures et loisirs (穿越, chuān yuè, 2010 ; 控 kòng, 2010), politique (打虎拍蝇 dǎ hǔ pāi yíng, 2014), économie (共享, gòng xiǎng, 2017), informatique (互联网+, hù lián wǎng +, 2015), etc. Il est surtout remarquable qu'une partie des sélections soient à l'origine des discours officiels du gouvernement chinois, par exemple : l'expression 中国梦 zhōng guó mèng (2013) qui signifie « le rêve chinois » est un slogan lancé par le président Xi Jinping. En nous appuyant sur les critères de sélection précisés dans la publication en 2008, nous avons constaté que la prise en considération des discours médiatiques est beaucoup soulignée par le magazine. Cela constitue une ressource importante de l'activité.

D'ailleurs, en dépit de son rôle de « défenseur du chinois » et de l'exclusion des créations grammaticalement incorrectes, nous pouvons constater que des propositions échappent à l'exclusion. Par exemple, la forme mixte Hold 住 (2011) ne peut qu'être considérée comme une expression ludique, propice à un usage familier parce qu'elle ne correspond ni à la norme linguistique chinoise ni anglaise. Nous supposons que sa présence dans la liste de sélections s'explique par son grand succès d'usage.

En résumé, les critères établis par le magazine ont conduit à une sélection avec des caractéristiques qui lui sont propres. Mais cette liste prend en compte en même temps une grande diversité non seulement des domaines thématiques comme des créateurs, individuels ou institutionnels.

### 5.1.2. Corpus B : 《汉语盘点》十大网络用语 (le top des 10 termes utilisés en ligne)

L'année de lancement de l'activité 《汉语盘点》 (hàn yǔ pán diǎn) remonte à 2006. Mais ce n'est qu'à partir de 2012 qu'est proposé le top des 10 termes utilisés en ligne. De ce fait nous avons pu regrouper 60 résultats pour les années 2012 à 2017.

2012	2013	2014	2015	2016	2017
<p>中国好声音</p> <p>zhōng guó hǎo shēng yīn</p> <p>[la belle voix chinoise]</p> <p>nouveau mode de sélection avec jury</p> <p>(programme <i>The Voice</i> en Chine)</p>	<p>中国大妈</p> <p>zhōng guó dà mā</p> <p>[les dames chinoises seniors]</p> <p>groupe de dames chinoises au pouvoir d'achat élevé qui pratiquent par ex. l'achat de l'or</p> <p>(sens social et économique)</p>	<p>我也是醉了</p> <p>wǒ yě shì zuì le</p> <p>[je suis ivre]</p> <p>sentiment de ne rien y pouvoir</p>	<p>重要的事情说三遍</p> <p>zhòng yào de shì qing shuō sān biàn</p> <p>il faut dire trois fois les choses importantes</p>	<p>洪荒之力</p> <p>hóng huāng zhī lì</p> <p>Voir n°12 *</p>	<p>打 call</p> <p>dǎ call</p> <p>Voir n°21*</p>
<p>元芳, 你怎么看</p> <p>yuán fāng, nǐ zěn me kàn</p> <p>Voir n°1*</p>	<p>高端大气上档次</p> <p>gāo duān dà qì shàng dàng cì</p> <p>Voir n°8 *</p>	<p>有钱就是任性</p> <p>yǒu qián jiù shì rèn xìng</p> <p>avec de l'argent on peut faire n'importe quoi</p>	<p>世界那么大, 我想去看看</p> <p>shì jiè nà me dà, wǒ xiǎng qù kàn kàn</p> <p>[le monde est grand et je voudrais le visiter]</p> <p>(à partir du cas d'une jf qui démissionne de son travail pour voyager,</p>	<p>友谊的小船</p> <p>yǒu yì de xiǎo chuán</p> <p>Voir n°15 *</p>	<p>尬聊</p> <p>gà liáo</p> <p>bavarder de façon futile</p>

			renvoie à attitude de liberté et de bien-être)		
高富帅，白富美 gāo fù shuài, bái fù měi  Voir n°4 *	爸爸去哪儿 bàba qù nǎ'er  [tu vas où Papa ?] on va où ? (titre émission télé-réalité)	蛮拼的 mán pīn de  travailler dur	你们城里人真会玩 nǐ men chéng lǐ rén zhēn huì wán  [Vous les citadins savez bien vous amuser] renvoie à des pratiques surprenantes insolites (ironie)	定个小目标 dìng gè xiǎo mù biāo  Voir n°14 * [se fixer un petit objectif] gagner une grosse somme d'argent (ironie)	你的良心不会痛吗 nǐ de liáng xīn bù huì tòng ma  [tu te sens mal à ton bon cœur ?] tu as mauvaise conscience ? (ironie)
你幸福吗 nǐ xìng fú ma  [est-ce que tu es heureux ?] (ironique)	小伙伴们都惊呆了 xiǎo huǒ bàn men dōu jīng dāi le  [mes copains et moi nous sommes tous étonnés] c'est étonnant	挖掘机技术哪家强 wā jué jī jì shù nǎ jiā qiáng  [quelle école donne la meilleure formation pour une belle mécanique hydraulique ?] (publicité pour recruter ou indiquer qualité)	为国护盘 wèi guó hù pán  [protéger la Bourse pour le pays] aider le pays à garder sa stabilité économique	吃瓜群众 chī guā qún zhòng  Voir n°13 *	惊不惊喜，意不意外 jīng bù jīng xǐ, yì bù yì wài  c'est une surprise (fin de message)

<p>江南 style</p> <p>jiāng nán style</p> <p>Gangnam style (titre chanson coréenne)</p>	<p>待我长发及腰</p> <p>dài wǒ zhǎng fā jí yāo</p> <p>[quand j'aurai les cheveux longs jusqu'à la ceinture...]</p> <p>j'attends un moment pour faire qqch.</p>	<p>保证不打死你</p> <p>bǎo zhèng bù dǎ sǐ nǐ</p> <p>[je t'assure je ne vais pas te battre à mort]</p> <p>viens pour que je te tape (plaisanterie)</p>	<p>明明可以靠脸吃饭,却偏偏要靠才华</p> <p>míng míng kě yǐ kào liǎn chī fàn, què piān piān yào kào cái huá</p> <p>[qqn qui peut gagner de l'argent simplement avec sa beauté mais a choisi de travailler avec ses talents]</p> <p>qqn qui est beau mais a aussi des talents</p>	<p>葛优躺</p> <p>gě yōu tāng</p> <p>Voir n°16 *</p>	<p>皮皮虾, 我们走</p> <p>pí pí xiā, wǒ men zǒu</p> <p>[les squillidae crevettes-mantes, on y va]</p> <p>au revoir</p> <p>(image crevette signale l'humour)</p>
<p>躺着也中枪</p> <p>tǎng zhe yě zhōng qiāng</p> <p>Voir n°3 *</p>	<p>喜大普奔</p> <p>xǐ dà pǔ bēn</p> <p>c'est une très bonne nouvelle et je vais courir pour que tout le monde le sache</p> <p>(structure qui imite une locution en 4 caractères ; sens très compact)</p>	<p>萌萌哒</p> <p>méng méng da</p> <p>Voir n°9*</p>	<p>我想静静</p> <p>wǒ xiǎng jìng jìng</p> <p>Laisse-moi tranquille</p>	<p>辣眼睛</p> <p>là yǎn jīng</p> <p>[Ça pique les yeux] se dit de choses peu intéressantes ou horribles</p>	<p>扎心了, 老铁</p> <p>zhā xīn le, lǎo tiě</p> <p>[mon frère tu as blessé mon cœur]</p> <p>je suis triste, tu me fais mal</p> <p>(humour pour refuser une conversation)</p>

屌丝，逆袭 <i>diǎo sī, nì xí</i>  Voir n°5 *	女汉子 <i>nǚ hàn zi</i>  Voir n°6 *	时间都去哪了 <i>shí jiān dōu qù nǎ le</i>  [Où va le temps] (titre chanson) gaspiller le temps sans en avoir conscience	吓死宝宝了 <i>xià sǐ bǎo bǎo le</i>  Voir n°10 *	全是套路 <i>quán shì tàolù</i>  Voir n°17 *	还有这种操作 <i>hái yǒu zhè zhǒng cāo zuò</i>  [il y a une opération pareille ?] c'est surprenant
舌尖上的中国 <i>shé jiān shàng de zhōng guó</i>  Voir n°2*	土豪金 <i>tǔ háo jīn</i>  Voir n°7 *	我读书少你别骗我 <i>wǒ dú shū shǎo nǐ bié piàn wǒ</i>  [je ne lis pas beaucoup et tu ne me trompes pas]  j'ai bcp de connaissances et tu n'arrives pas à me tromper (antiphrase)	内心几乎是崩溃的 <i>nèi xīn jī hū shì bēng kuì de</i>  [au fond du cœur c'est presque écrasant]  je suis dépassé	蓝瘦，香菇 <i>lán shòu xiāng gū</i>  Voir n°18 *	怼 <i>duǐ</i>  Voir n°20 *

<p>最炫民族风 zuì xuàn mín zú fēng</p> <p>[le meilleur style folklorique]</p> <p>(titre chanson)</p>	<p>摊上大事了 tān shàng dà shì le</p> <p>rencontrer un terrible événement, être frappé par le malheur</p>	<p>画面太美我不敢看 huà miàn tài měi wǒ bù gǎn kàn</p> <p>[l'image est très belle et je n'ose pas regarder] je ne peux pas supporter de regarder ça (antiphrase pour qqch. de choquant)</p>	<p>我妈是我妈 wǒ mā shì wǒ mā</p> <p>[ma mère c'est ma mère] (ironique surtout à l'égard des administrations qui demandent de prouver cela)</p>	<p>老司机 lǎo sī jī</p> <p>[vieux chauffeur]  qqn qui a de l'expérience dans un domaine</p>	<p>你有 freestyle 吗 nǐ yǒu freestyle ma</p> <p>est-ce que tu as le free style ?</p>
<p>给跪了 gěi guì le</p> <p>[s'agenouiller] admirer qqn</p>	<p>涨姿势 zhǎng zīshì</p> <p>j'ai appris qqch. (paronyme)</p>	<p>且行且珍惜 qiě xíng qiě zhēn xī</p> <p>[on marche tout en chérissant qqch.] il faut être compréhensif</p>	<p>主要看气质 zhǔ yào kàn qì zhì</p> <p>Voir n°11 *</p>	<p>厉害了我的哥 lì hài le wǒ de gē</p> <p>bravo mon ami,e</p>	<p>油腻 yóu nì</p> <p>Voir n°19*</p>

Tableau 20 : Liste du corpus B

Dans un premier temps, un contraste est immédiatement repérable si on compare la liste du corpus B à celle du corpus A. Nous avons constaté une importante présence des longues expressions allant jusqu'à l'échelle de la phrase. Par la suite, un aperçu rapide du contenu fait remarquer une deuxième différence : les nouveautés langagières imposées par le gouvernement ne constituent pas celles les plus utilisées en ligne et sont donc absentes dans cette liste. Rappelons que les propositions du corpus B sont basées principalement sur la fréquence d'utilisation et que nous avons spécialement pris en considération la catégorie qui concerne les termes néologiques utilisés en ligne.

Outre ces différences entre le corpus A et le corpus B, nous avons constaté certaines caractéristiques des entrées du corpus B comme suit :

- la structure libre des nouveautés se rapproche de celle de la langue orale. Par exemple : dans l'expression 厉害了我的哥 (lì hài le wǒ de gē, « qui sert à apprécier quelqu'un »), contrairement à la norme linguistique chinoise (sujet + particule + complément), le sujet 我的哥 (wǒ de gē, « mon ami.e ») est placé à la fin de la phrase avec absence de la particule ;

- l'interprétation de ces nouveautés langagières demande la connaissance du contexte de création pour saisir le sens figuré ; en effet, nous avons constaté l'emploi de différentes figures dans les expressions. Par exemple : l'effet ironique est net dans l'expression 你们城里人真会玩<sup>164</sup> (nǐ men chéng lǐ rén zhēn huì wán, « pour renvoyer à des pratiques surprenantes insolites ») ; 画面太美我不敢看 (huà miàn tài měi wǒ bù gǎn kàn) est une antiphrase pour indiquer des images choquantes et non le sens littérale de « très belles images ».

Cette liste nous permet d'obtenir des résultats qui sont étroitement liés au contexte numérique auquel nous nous intéressons dans cette recherche.

---

<sup>164</sup> Traduction littérale : « Vous les citadins savez bien vous amuser ».

### 5.1.3. Corpus C : 2014 网络流行语排行 (les mots populaires de l'Internet de l'année 2014)

Le troisième sous-corpus regroupe les 25 nouveautés langagières les plus utilisées durant l'année 2014 sur la plateforme de microblog chinois *Sina Weibo*.

1	萌萌哒 <i>méng méng da</i>	voir n°9 *
2	逗比 <i>dòu bǐ</i>	quelqu'un de drôle
3	我也是醉了 <i>wǒ yěs hì zuì le</i>	[je suis ivre] sentiment de ne rien y pouvoir
4	且行且珍惜 <i>qiě xíng qiě zhēn xī</i>	[on marche tout en chérissant quelque chose] il faut être compréhensif
5	涨姿势 <i>zhǎng zī shì</i>	j'ai appris qqch. (paronyme) j'ai pris connaissance
6	心塞 <i>xīn sāi</i>	[le cœur est bloqué] je ne me sens pas bien
7	暖男 <i>nuǎn nán</i>	[chaud homme] homme très attentionné
8	买买买 <i>mǎi mǎi mǎi</i>	acheter
9	也是蛮拼的 <i>yě shì mán pīn de</i>	travailler dur
10	逼格 <i>bī gé</i>	avoir de la classe
11	画面太美我不敢看 <i>huà miàn tài měi wǒ bù gǎn kàn</i>	[l'image est très belle et je n'ose pas regarder] je ne peux pas supporter de regarder ça (antiphrase pour qqch. de choquant)
12	小鲜肉 <i>xiǎo xiān ròu</i>	[petite viande fraîche] jeune homme
13	那么问题来了 <i>nà me wèn tí lái le</i>	voici que viennent les questions ?
14	No zuo no die	il ne faut pas jouer avec le feu
15	脑洞大开 <i>nǎo dòng dà kāi</i>	[il y a un grand trou dans le cerveau] libérer l'imagination
16	有钱就是任性 <i>yǒu qián jiù shì rèn xìng</i>	avec de l'argent on peut faire n'importe quoi
17	时间去哪儿了 <i>shí jiān qù nǎ'èr le</i>	[Où va le temps ?]



		gaspiller le temps sans en avoir conscience (titre chanson)
18	保证不打死你 <b>bǎo zhèng bù dǎ sǐ nǐ</b>	[je t'assure je ne vais pas te battre à mort] viens pour que je te tape (plaisanterie)
19	我读书少你别骗我 <b>wǒ dú shū shǎo nǐ bié piàn wǒ</b>	[je ne lis pas beaucoup et tu ne me trompes pas] j'ai bcp de connaissances et tu n'arrives pas à me tromper (antiphrase)
20	习大大 <b>xí dà dà</b>	[Xi grand grand] Oncle Xi (surnom du président Xi, sens positif)
21	上天台 <b>shàng tiān tái</b>	[aller sur le toit] se suicider (plaisanterie pour se moquer quand on subit un échec)
22	什么仇什么怨 <b>shén me chóu shén me yuàn</b>	[quelle haine, quelle hostilité] se demander pourquoi
23	你家里人知道吗? <b>nǐ jiā lǐ rén zhīdào ma?</b>	[les membres de ta famille sont au courant?] ce n'est pas bien (ironie) ou c'est super (sens positif)
24	挖掘技术哪里强 <b>wā jué jì shù nǎ lǐ qiáng</b>	[quelle école donne la meilleure formation pour une belle mécanique hydraulique?] (publicité pour recruter ou indiquer qualité)
25	APEC 蓝 <b>APEC lán</b>	[le bleu APEC] la bonne qualité de l'air (référence au ciel bleu rare à Pékin lors de l'APEC Chine 2014 en raison de la campagne de réduction des émissions dirigée par le gouvernement chinois.)

Tableau 21 : Liste du Corpus C

Selon nous, deux points remarquables sont à noter après avoir examiné cette liste du corpus C. Tout d'abord, nous retrouvons des nouveautés langagières du corpus A et B dans les 25 sélections élues. Ce fait nous permet d'être assurée de la pertinence des deux premiers corpus qui n'ont pas été collectés directement sur le dispositif du microblog. Le repérage à la fois sur le microblog et sur les autres supports qui constituent les ressources des autres d'activités met en relief un usage important des nouveautés langagières. C'est aussi ce constat qui nous a incitée pour constituer notre corpus de tweets – que nous présenterons dans le chapitre suivant –, à prendre les mots communs entre les corpus comme critère de sélection.

Deuxièmement, il est aussi intéressant de constater que de longues expressions sont au premier rang des nouveautés langagières utilisées sur un dispositif qui met pourtant en avant un message limité.

## **5.2. Procédures de constitution du corpus français**

### **5.2.1. Les nouvelles entrées des dictionnaires citées dans les articles de presse**

L'impossibilité de trouver des activités françaises identiques aux chinoises, que nous avons abordées dans le chapitre IV, nous confronte à plusieurs difficultés relatives notamment à la quantité, à la période d'observation, à la question de la nouveauté pour constituer le corpus de listes français. Comme nous l'avons déjà indiqué sommairement, nous nous sommes ainsi orientée vers une constitution du corpus de liste à partir des nouvelles entrées des dictionnaires de langue citées dans les articles de presse. Nous allons maintenant justifier cette option.

#### ***Le choix des dictionnaires à observer***<sup>165</sup>

Durant nos premières recherches<sup>166</sup> visant à identifier les productions langagières néologiques, nous avons déjà constaté que certains articles de presse française font état des nouvelles éditions annuelles des dictionnaires et citent plusieurs nouvelles entrées. Faute de pouvoir trouver en France des activités équivalentes à celles de Chine relatives aux créations lexicales, nous nous sommes intéressées à d'autres métadiscours disponibles (ressources encyclopédiques, dictionnaires partagés, articles de presse qui commentent les nouveaux phénomènes langagiers, etc.) et décidé d'observer les nouvelles entrées des dictionnaires de langue. Les dictionnaires nous ont paru d'une bien plus grande fiabilité que les ressources encyclopédiques et collaboratives.

Cette fiabilité repose sur la possibilité pour le public de connaître les démarches et les principes

---

<sup>165</sup> Pour les critères de sélection des deux dictionnaires que nous avons retenus, nous avons eu recours à des références encyclopédiques en ligne :

<https://www.lerobert.com/nos-metiers> (consulté le 22 décembre 2019)

[https://www.lemonde.fr/les-decodeurs/article/2015/05/20/comment-le-petit-larousse-choisit-ses-nouvelles-entrees\\_4635608\\_4355770.html](https://www.lemonde.fr/les-decodeurs/article/2015/05/20/comment-le-petit-larousse-choisit-ses-nouvelles-entrees_4635608_4355770.html) (consulté le 22 décembre 2019)

<http://www.academie-francaise.fr/le-dictionnaire/commission-du-dictionnaire> (consulté le 22 décembre 2019)

<sup>166</sup> Voir chapitre 3.2.2, les recherches par métatermes.

de sélection et sur le professionnalisme de la rédaction des dictionnaires. Au sein d'une équipe éditoriale, les lexicographes s'appuient d'abord sur les propositions des linguistes et spécialistes de domaines différents (informatique, sciences, gastronomie, économie...) afin de connaître les mots « candidats » à l'ajout dans une nouvelle édition. Ensuite, les lexicographes sont chargés de leur côté de cerner les divers usages des ajouts proposés sur des supports différents (journaux, radios, télévision, livres, Internet, etc.), c'est-à-dire d'exercer une veille néologique. Il faut souligner que l'objectif de la veille est de vérifier si le terme a connu un usage répandu et durable pendant une période donnée. Cela indique bien que les nouvelles entrées des dictionnaires de langue sont présélectionnées selon leur fréquence d'utilisation, critère qui coïncide avec un de ceux que nous avons suivis durant la constitution du corpus chinois.

Par la suite, un processus de vote a lieu pour obtenir une liste finale en fonction de critères supplémentaires qui sont propres à chaque maison d'édition. Nous avons constaté que dans le *Petit Robert* et le *Petit Larousse*, les choix sont guidés par les faits d'actualité et les tendances sociales dans l'objectif d'être le reflet de la société. Ils effectuent ainsi une mise à jour chaque année et se montrent plus ouverts aux termes accueillis que par exemple le dictionnaire de l'Académie française qui assure une mission de préservation de langue française. La prise en compte de cet aspect de popularité des nouveautés langagière nous paraît un critère essentiel. Nous pensons que le recours aux nouvelles entrées de ces deux dictionnaires présente l'avantage de constituer le corpus de listes français en partant des mêmes critères (nouveau, popularité et fréquence d'utilisation) que pour le corpus chinois malgré les différences des procédures de collecte. En même temps, leur publication régulière nous offre un choix abondant. À part le cas d'une refonte (celle du *Petit Larousse 2012* qui compte plus d'un millier d'entrées), le nombre moyen des nouvelles entrées est d'une centaine<sup>167</sup>. Ajoutons que les deux dictionnaires destinés au grand public que sont le *Petit Larousse* et le *Petit Robert*, ne proposent pas forcément les mêmes mots en même temps. Il serait intéressant, à ce propos, de comparer leurs nouvelles entrées annuelles respectives pour examiner la diversité des choix.

---

<sup>167</sup> Voir le Tableau n°Tableau 18 : Récapitulatif synthétique des nouvelles entrées de deux dictionnaires selon le club d'orthographe de Grenoble dans le chapitre 4.

### *Le choix des organes des presse*

Si nous prenons la même période d'observation que pour le corpus chinois, de 2009 à 2017, nous pouvons arriver à une quantité d'environ 5000 entrées. Or cette quantité d'entrées est beaucoup plus importante que celle du corpus chinois. Pour des raisons d'équilibre entre les deux corpus, nous trouvons nécessaire d'établir un autre filtre de sélection pour constituer un échantillon comparable et nous avons trouvé une solution en nous référant aux articles qui signalent les nouvelles éditions des dictionnaires avec des exemples. Ce sont les mentions des exemples (éventuellement commentés) qui constituent un sous-ensemble. Cela nous permet d'une part de résoudre la question du nombre en diminuant considérablement la quantité et d'autre part d'être assurée du caractère néologique de ces exemples.

Pour le choix des articles, nous nous sommes appuyée premièrement sur le moteur *Google* pour repérer les articles en utilisant « nouvelles entrées du dictionnaire » comme mots clés de recherche. Nous avons ainsi pu relever de différentes sources de publication : des organes de presse, des sites encyclopédiques ou des billets personnels. Nous avons ensuite appliqué des filtres de temps (pour la période 2009 à 2017) afin de faire un bilan des publications de chaque année. Les résultats montrent que ce sont les organes de presse qui accordent d'une manière régulière leur attention aux nouvelles publications des dictionnaires, en faisant en quelque sorte la promotion et l'occasion de commenter l'évolution de la langue et de la société. De ce fait, nous avons choisi de collecter nos entrées de liste à partir des articles de presse.

En prenant en compte les caractéristiques des médias parmi les résultats disponibles et dans l'objectif d'assurer une certaine diversité au corpus sans viser l'exhaustivité ni chercher à systématiser une étude du contenu de ces articles – ce qui n'est pas notre objectif –, nous avons choisi quatre médias : *Le Figaro*, *20 Minutes*, *Le Parisien* et *RFI*. Ces médias jouissent tous les quatre d'une diffusion importante en visant des publics différents de part leurs orientations, leur couverture et leur nature. Il s'agit de trois quotidiens généralistes et d'une radio à vocation internationale, désormais développant un site multimédia avec des articles écrits. Les journaux sont publiés à la fois en version papier et en ligne : un quotidien national (*Le Figaro*) et un

quotidien régional (*Le Parisien*) qui ont des tirages importants<sup>168</sup> ; *20 Minutes* est un quotidien d'information gratuite diffusé en France et dans d'autres pays avec une édition locale et nationale<sup>169</sup>. Quant à *RFI*<sup>170</sup>, c'est une radio française tournée vers l'international qui constitue une plateforme multimédia et met également l'accent sur la langue française et la francophonie.

### ***La constitution du corpus de liste***

Sur ces bases nous avons commencé à constituer notre corpus de liste français avec les nouvelles entrées du dictionnaire *Le Petit Robert* (ci-après dénommé corpus Af) et celles du dictionnaire *Le Petit Larousse* (ci-après dénommé corpus Bf) mentionnées dans les articles.

Nous avons en effet pu collecter des articles qui traitent des nouvelles éditions des dictionnaires et citent des entrées mais la quantité de nouvelles entrées rapportées n'est pas identique d'une année à l'autre, d'un journal à l'autre et ne traitent pas forcément des deux dictionnaires. Par exemple, les articles de *Figaro* relevés, datant respectivement des années 2011 et 2016, évoquent les nouveautés du dictionnaire *Le Petit Robert* et non du *Petit Larousse*. En 2014, *RFI* s'intéresse seulement aux nouvelles entrées du *Petit Larousse* et non à celles du *Petit Robert*. Même pour *le Figaro*, nous n'avons pas trouvé d'articles publiés plus tôt que l'année 2011.

Nous pensons que notre choix de mots-clés appliqués dans la recherche a restreint la récolte que nous avons faite et qu'il aurait sans doute été possible de trouver d'autres articles sur des contenus semblables avec d'autres mots-clés, mais pour les objectifs que nous visons nous l'avons jugée suffisante et n'en avons pas relancé une autre.

---

<sup>168</sup> Le tirage en 2019 est de plus de 300 000 exemplaires en 2019 d'après [https://fr.wikipedia.org/wiki/Le\\_Figaro](https://fr.wikipedia.org/wiki/Le_Figaro). *Le Figaro* est un très ancien journal, fondé en 1826. Il a une tradition littéraire et propose beaucoup d'articles sur les questions de langue.

*Le Parisien* a un tirage de plus de 180 000 exemplaires en 2019 d'après [https://fr.wikipedia.org/wiki/Le\\_Parisien](https://fr.wikipedia.org/wiki/Le_Parisien); diffusé en Ile-de-France. Il a aussi une édition nationale sous le titre *Aujourd'hui en France*. Il s'appelait à l'origine, en 1945 *Le Parisien libéré* en référence à la fin de la Seconde guerre mondiale.

<sup>169</sup> Son tirage est de près de 900 000 exemplaires en 2018 d'après [https://fr.wikipedia.org/wiki/20\\_Minutes\\_\(France\)](https://fr.wikipedia.org/wiki/20_Minutes_(France)). Il est financé par la publicité et a été fondé en 2002.

<sup>170</sup> Radio France Internationale, créée en 1975 émet en français et dans une quinzaine d'autres langues dans le monde. Elle a plus de 40 millions d'auditeurs d'après [https://fr.wikipedia.org/wiki/Radio\\_France\\_internationale](https://fr.wikipedia.org/wiki/Radio_France_internationale) et est très écoutée en Afrique. Sa vocation est d'être tournée vers l'extérieur et de valoriser la francophonie. Elle assure une information en continu. Ses programmes font une place importante aux questions de langue. Le site de *RFI* <https://www.rfi.fr/fr/> diffuse des contenus audio, vidéo et écrits.

Non seulement nous ne pouvions pas constituer la liste française d'une manière identique à celle du corpus de liste chinois (à partir d'activités d'institutions et en suivant les sélections sur plusieurs années consécutives avec une dizaine d'entrées par année) ni parvenir à un nombre comparable mais les entrées mentionnées par les organes de presse ne sont pas un palmarès : c'est un choix variable d'entrées jugées intéressantes, surprenantes, conformes aux tendances et révélatrices d'un état social. En quelque sorte, c'est une autre manière de récapituler l'actualité pour la presse.

Nous avons choisi finalement de négliger le critère de sélection consécutive pour un même média mais de maintenir le principe de regrouper les nouvelles entrées sur des années différentes et de nous baser sur un nombre total qui ne soit pas inférieur à la quantité d'entrées du corpus chinois (175 sélections collectés). Notre analyse comparative ultérieure ne consiste pas en effet à distinguer les sélections annuelles mais à observer d'un point de vue général les nouveautés langagières en Chine et en France pendant une période donnée.

Cette période que nous avons définie au départ à partir de deux bornes (2009 : création de *Sina Weibo* / 2017 : suppression de la limite de 140 signes pour *Twitter*) a été circonscrite un peu différemment par les résultats de recherche. Comme observé dans le corpus chinois où il existe des nouveautés langagières en commun qui sont sélectionnées par une institution telle année, mais intègrent plus tard la sélection d'une autre institution, la situation d'intégration des nouveautés une année plutôt qu'une autre dans un des deux dictionnaires se retrouve. Il faut aussi tenir compte du fait que l'édition d'un dictionnaire porte la date de l'année de la publication (par exemple 2017 pour une collecte réalisée en 2016) et pas de l'année de création des nouvelles entrées qui mettent plus ou moins de temps à s'imposer en fonction du sentiment néologique que nous avons évoqué *supra* et des critères des équipes d'experts du dictionnaire.

La période de sélection pour notre corpus de liste français a donc été redéfinie de 2011 à 2017. Nous prenons en considération les nouvelles entrées rapportées par l'un ou l'autre des quatre médias durant cette période pour qu'elles puisse se compléter. Parmi les entrées citées, les noms propres de personnalités sont exclus de notre sélection car il ne s'agit pas d'une nouvelle création lexicale mais de la notoriété acquise par des personnes. Comme le montrent les tableaux récapitulatifs ci-dessous (Tableau 22, Tableau 23), nous obtenons donc un corpus Af (*Le Petit Robert*) d'un nombre de 140 sélections, collecté sur les années 2011, 2013, 2015, 2016

et 2017, et un corpus Bf (*Le Petit Larousse*) constitué de 231 sélections, collecté respectivement sur les années 2011, 2014, 2015, 2016 et 2017.

		les organes de presse			
		Figaro	20 Minutes	Le parisien	RFI
Année d'édition Petit Robert	Nombre des nouvelles entrées				
		2017			16
	2016	17	15		10
	2015				6
	2014				
	2013			29	24
	2012				
	2011	23			

**Tableau 22 : Récapitulatif des nouvelles entrées cités par les organes de presse  
(Dictionnaire *Le Petit Robert*) - Corpus Af**

		les organes de presse			
		Figaro	20 Minutes	Le parisien	RFI
Année d'édition Petit Larousse	Nombre des nouvelles entrées				
		2017	40	18	13
	2016	32	4		11
	2015	37	5		4
	2014	18	22		
	2013				
	2012	2			
	2011	12			

**Tableau 23 : Récapitulatif de nouvelles entrées citées par les organes de presse  
(Dictionnaire *Le Petit Larousse*) - Corpus Bf**

Précisons que les listes des nouveautés langagières classées par média et classées par dictionnaire sont disponibles en annexe n°3<sup>171</sup>.

### 5.3. Analyse linguistique comparée des corpus de listes

Nous avons donc constitué deux types de corpus, en chinois d'une part, en français d'autre part.

<sup>171</sup> Il s'agit des annexes 3a et 3b. Le nombre total des entrées lexicales et le fait qu'il s'agit de termes français qui ne posent pas de problème de compréhension dans une recherche rédigée en français, contrairement aux termes chinois nous a décidée à les faire figurer uniquement en annexe et non dans le corps du présent chapitre.

Le premier que nous nous nommons corpus de liste est élaboré à partir de listes de nouveautés langagières, comme nous venons de l'expliquer. Le deuxième regroupe des tweets reliés à une partie restreinte des nouveautés langagières sélectionnées au sein de ces listes et nous l'aborderons au chapitre suivant (chapitre VI).

Nous allons observer maintenant les corpus dans une perspective linguistique. L'objectif est de dégager les principaux procédés de création dans les deux langues (chinoise et française). Nous rappelons que nous avons fait l'hypothèse que des mêmes tendances de création devraient se manifester dans les deux langues.

### **5.3.1. Une grille d'analyse commune pour comparer les procédés de création**

Pour les corpus de liste aussi bien en chinois qu'en français, afin d'identifier les procédés de création de chaque nouveauté langagière en rendant possible la comparaison, nous allons partir d'une référence commune, le tableau de la typologie des néologismes selon Sablayrolles (2019) que nous avons présenté au chapitre III et que nous reproduisons ici une nouvelle fois pour commodité de consultation (Tableau 15b).



m a t r i c e s  i n t e r n e s	morpho- sémantiques	construction	affixation	préfixation	détatouer, rediaboliser
				suffixation	statuesque, tontoniser
				dérivation inverse	dissider, prester
				parasynthétique	désidéologisé
				flexion	ils closirent, lesbien une représaille
		composition	composition « rég. »	voiture-bélier	
			synapsie	lanceur d'alerte	
			composition savante	ptochophobie	
			hybride	nounoursothérapie	
			compoaction	glamping, héliport	
imitation et déformation	fractocomposition	télespectateur, xgate			
	mot valise	peopolitique, quinquado			
	factorisation	optipessimiste, hépathétique			
	substitution	monokini, trilogie			
	onomatopée fausse coupe ou paronymie redoublement déf. syst. : verlan...	dzoing, whishh la nesthésie, pestacle tuture (voiture) reuch, loufoque			
m a t r i c e s  e x t e r n e	syntactico- sémantiques	changement de fonction	conversion	la gagne, bande-dessiner	
			conversion verticale	son ex, les chic ouf	
			déflexivation	le manger, le pleurer	
			combinatoire syntax <sup>2</sup> / lexicale	ça craint, encourir la liberté	
			changement de sens	métaphore	souris (inform.), bounty
		métonymie	passoire thermique		
		autres figures	déficient auditif (sourd)		
		purement morpho- logiques	réduction de la forme	truncation	ricain, blème, déj'...
				siglaison	LMD
				acronyme	ecue, pacs, perdir
phraséologiques	création d'expression	détournement	faire du huit mégabits à promesses		
		emprunt	quantified self, fact checking, fioul, redingote		

Tableau 15 : Tableau raisonné et hiérarchisé des matrices (Sablayrolles 2019:127)

Comme le tableau est hiérarchisé sur plusieurs niveaux, nous allons observer les proportions qu'occupent les procédés des niveaux supérieurs pour avoir une idée des tendances de création. Nous allons nous intéresser tout d'abord à la répartition des nouveautés langagières dans les matrices internes et externe (soit le 1<sup>er</sup> niveau). Par la suite, nous allons classer les nouveautés langagières selon cinq grands sous-ensemble au sein des matrices (soit le 2<sup>e</sup> niveau). Quatre matrices internes sont distinguées: matrice morpho-sémantique (MS ci-après), matrice syntactico-sémantique (SS ci-après), matrice purement morphologique (PM ci-après), matrice phraséologique (PH ci-après). L'emprunt constitue le seul procédé de la matrice externe.

Pour les autres niveaux hiérarchisés inférieurs, nous n'allons pas nous intéresser aux proportions qu'ils représentent mais nous allons mobiliser les procédés pour comprendre les opérations linguistiques et rhétoriques qui produisent les nouveautés langagières, sans les traiter une par une mais en examinant des cas particulièrement saillants ou exemplaires. Sur ces bases et après avoir traité séparément les deux corpus, nous tenterons de comparer les procédés de

création dans les deux langues.

Les annexes 2d pour le corpus chinois et 3c pour le corpus français présentent le récapitulatif de l'intégralité des nouveautés langagières classées par matrices<sup>172</sup>.

### **5.3.2. Le corpus de listes chinois**

Le corpus de listes chinois est composé par 3 sous-corpus<sup>173</sup> dont le total donne 175 résultats. Mais, bien que les critères des activités de sélection des nouveautés langagières ne soient pas identiques, certaines entrées lexicales sont communes aux sous-corpus (voir chapitre VI). Cela ramène le nombre des entrées du corpus de liste chinois à 140. C'est à partir de ce nombre que nous avons effectué des calculs de proportion présentés ci-dessous. Chaque entrée répertoriée est étiquetée selon les catégories de la typologie des néologismes, le détail figurant en annexe 2d.

#### **5.3.2.1 Matrices internes et matrice externe**

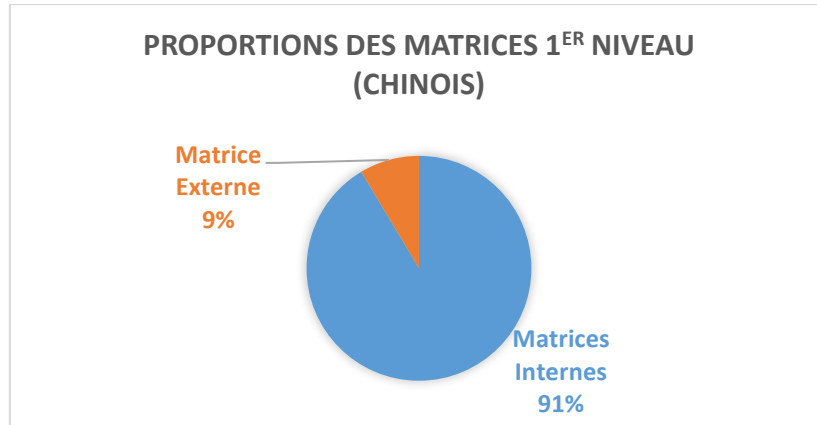
##### *Proportions*

Le premier résultat que montre la Figure 22 concerne la proportion du 1<sup>er</sup> niveau des matrices internes et externe

---

<sup>172</sup> L'annexe 3f quant à elle, présente uniquement en version électronique du volume 2 pour des raisons de lisibilité, contient un Tableau Excel préparatoire au classement qui précise pour chaque entrée du corpus de liste français de manière détaillée la ou les matrices à l'origine du classement.

<sup>173</sup> 1. Corpus A : les 10 termes annuels (de 2009-2017) choisis par le magazine 《咬文嚼字》 (yǎo wén jiáo zì) ;  
2. Corpus B : les 10 termes sélectionnés annuellement (de 2012-2019) à travers l'activité 《汉语盘点》 (hàn yǔ pán diǎn) ;  
3. Corpus C : les 25 termes proposés dans un rapport de 2014 portant sur le langage de l'Internet en Chine.



**Figure 22 : Proportions des matrices internes et externe dans le corpus de listes chinois**

Le diagramme de la Figure 22 ci-dessus montre les proportions qu’occupent respectivement la matrice externe (9%) et les matrices internes (91%) dans les procédés de création des nouveautés langagières chinoises. Il est important de rappeler que dans la typologie qui nous sert de référence, l’emprunt constitue le seul procédé de la matrice externe. En fait, le statut de l’emprunt au sein des typologies de néologismes fait débat depuis longtemps. Certains chercheurs optant pour une trichotomie classique<sup>174</sup> considèrent l’emprunt comme un procédé indépendant à côté du néologisme formel et du néologisme sémantique puisque la nouveauté de l’emprunt vient des langues étrangères. Mais des avis divergents plaident pour considérer l’emprunt comme un sous-type à la fois des néologismes formels et des néologismes sémantiques dans la mesure où les altérations sont constatées aussi bien au niveau de la morphologique que du sens. Dans la typologie établie à laquelle nous nous référons, l’influence des éléments étrangers sur le français est mise en valeur par Sablayrolles. Suivant cette interprétation et en nous intéressant ici au chinois, dans notre corpus nous avons relevé 12 emprunts sur les 140 résultats soit moins de 10 % de l’ensemble. La présence des langues étrangères serait donc peu importante.

En fait, l’identification des emprunts ne va pas de soi. La formation des emprunts est étroitement liée au contact des langues qui est un processus linguistique, social et culturel. Non seulement les caractéristiques linguistiques de la langue prêteuse, mais aussi les relations qu’entretiennent les langues en contact ainsi que les pratiques concernées sont à prendre en

---

<sup>174</sup> Voir chapitre III 3.1.4.

considération. Il faut aussi souligner que la formation d'un néologisme donné ne dépend pas toujours d'un seul procédé de création. Par exemple, le mot « geeker » est créé par suffixation sur le mot « geek » qui est emprunté à l'anglais. Nous nous posons les questions suivantes : les éléments étrangers constituent-ils vraiment le critère déterminant pour identifier les emprunts ? Comment distinguer l'emprunt des procédés des matrices internes quand les influences étrangères ne sont pas manifestées directement ou sont peu visibles sur le plan morphologique ? En dépit du petit nombre d'entrées classées dans les emprunts, nous avons trouvé une diversité de cas. Nous allons aborder successivement les caractéristiques des emprunts relevés et commenter certains cas qui provoquent notre réflexion.

N°	Emprunts identifiés dans le corpus de listes chinois	Langues sources
1	控 / kòng	japonais
2	达人 / dá rén	japonais
3	卖萌 / mài méng	japonais
4	吐槽 / tù cáo	japonais
5	逆袭 / nì xí	japonais
6	打 call / dǎ call	japonais
7	断舍离 / duàn shě lí	japonais
8	颜值 / yán zhí	japonais
9	低碳 / dī tàn	anglais
10	新常态 / xīn cháng tài	anglais
11	江南 style / jiāng nán style	coréen
12	暖男 / nuǎn nán	coréen

**Tableau 24 : Emprunts identifiés dans le corpus de listes chinois et leurs sources**

#### *Discussion sur les emprunts : langues étrangères*

Les emprunts dans notre corpus (Tableau 24) ont été examinés sous différents angles : source, forme, histoire, pratiques. Tout d'abord, nous avons identifié trois langues étrangères qui constituent les sources de création. Il s'agit du japonais (8), du coréen (2), de l'anglais (2). Ces emprunts réfèrent à des notions qui sont populaires ou nouvelles dans leurs langues d'origine et dont les équivalents en chinois manquent. En effet, les emprunts au japonais sont issus de

différents domaines culturels dont la culture ACGN<sup>175</sup> et la culture des fans. Les deux emprunts au coréen sont liés respectivement à une chanson connue *Gangnam Style* et à la création des fans (concernant les feuilletons coréens). Les emprunts à l'anglais concernent des notions économique 新常态 *xīn cháng tài* (*new normal*) et écologique 低碳 *dī tàn* (*low carbon*) qui ont aujourd'hui un emploi international. Le résultat montre que les échanges culturels restent un élément majeur de la formation des emprunts dans notre corpus et que la culture japonaise a une influence non négligeable.

En ce qui concerne les aspects formels, à première vue, dix emprunts sur douze sont constitués des caractères chinois alors que deux sont mixtes (caractères chinois combinés avec les mots anglais). L'influence des langues étrangères se manifesterait davantage sur le plan sémantique mais il faut nuancer cette assertion en raison des relations entre les langues prêteuses et le chinois : la formation originelle des langues japonaise et coréenne a été fortement influencée par le chinois archaïque.

Aujourd'hui, dans les signes idéographiques de l'écriture japonaise, les *kanjis*, certains caractères chinois ont été conservés sans modification morphologique, alors que dans la langue coréenne, les traces chinoises sont plus faciles à repérer sur le plan phonétique. Nous avons constaté que cet aspect historique a beaucoup joué sur la formation des emprunts dans notre corpus. Par exemple, les emprunts japonais « 达人 » (*dá rén*) et « 逆袭 » (*nì xí*) sont les résultats de l'association directe des *kanjis* « 達人 (たつじん) », « 逆袭 (ぎゃくしゅう) » aux caractères chinois simplifiés (voir Chapitre I) en transposant en même temps le nouveau sens. « 达人 » et « 逆袭 » constituent des caractères chinois assimilés à la langue japonaise. Ils signifient respectivement « quelqu'un de talent » et « réussir malgré les désavantages ». Les expressions « 达人 » et « 逆袭 » qui existent également dans le chinois contemporain sont d'ailleurs dotés d'un sens différent<sup>176</sup> de celui du japonais. Dans ce cas-là, convient-il de les considérer seulement comme des emprunts qui se traduisent par des modifications sémantiques ou comme des créations chinoises ?

La nouveauté langagière issue du coréen « 暖男 » (*nuǎn nán*), désignant un homme très

---

<sup>175</sup> Sigle du Animation, Comic, Game, Novel

<sup>176</sup> « 达人 » signifie originellement « personne rationnelle » en chinois et « 逆袭 » signifie « contre-attaquer ».

attentionné, constitue un autre exemple significatif. Ce terme a été créé à partir d'un néologisme coréen « hum-nam » (la romanisation du terme 훈남) qui sert à décrire les caractères des personnages masculins dans les feuilletons coréens romantiques. Le caractère chinois « 暖 » est un adjectif qui signifie originellement « chaud et doux ». Il est employé d'une manière métaphorique pour décrire psychologiquement les personnages. Le caractère « 男 » signifie « homme » en chinois. Nous supposons que son emploi n'est pas seulement motivé par la recherche d'un équivalent en chinois, mais que sa prononciation, similaire à celle du coréen, est aussi prise en compte. Cet aspect phonétique est aussi constaté dans les emprunts japonais « 控 », « 打 call » et l'emprunt coréen « 江南 style ». Il est également intéressant de remarquer la forme mixte de l'emprunt « 打 call »<sup>177</sup>. Dans la translittération du terme japonais « コール », le terme anglais « call » est utilisé par rapprochement phonétique avec la prononciation japonaise.

Ces exemples illustrent une formation des nouveautés langagières issues des langues étrangères asiatiques et réalisée par des procédés mixtes (d'ordre graphique, sémantique et/ou phonétique), avec parfois une structure combinatoire mixte. Cela est étroitement lié aux relations entre ces langues. Ces relations qu'entretiennent les langues nous ont conduite à privilégier l'influence des langues étrangères, surtout quand les procédés de création se superposent en considérant que ces nouveautés langagières relèvent de la matrice externe.

#### *Discussion sur les emprunts : formes mixtes*

Dans notre corpus de liste, nous avons relevé plusieurs autres formes mixtes sans les classer comme emprunts même si les éléments étrangers, surtout des mots anglais, sont repérables. Nous les considérons comme des créations chinoises basées sur des éléments étrangers. Cette décision est inspirée par les travaux en collaboration de Sablayrolles et Jacquet-Pfau (2008) portant sur le repérage des emprunts. Selon eux, si les nouvelles créations n'existent pas et ne peuvent pas exister dans les langues d'où viennent les éléments étrangers, il ne convient pas de les traiter comme des emprunts. Ainsi, ils classifient les hybrides (combinaison d'éléments de

---

<sup>177</sup> Le sens contextualisé est précisé dans le chapitre VI, le Tableau n°27 intitulé « Liste des nouveautés langagières en commun et leurs traductions ».

la langue<sup>178</sup> et d'éléments étrangers) *a priori* comme des créations internes de la langue française tout en ouvrant la possibilité de les considérer comme emprunts. Dans les cas relevés, les éléments étrangers sont ajoutés dans la structure combinatoire pour créer un effet selon la situation donnée. Ni la structure figée ni le sens connotatif ne peuvent être retrouvés dans les langues prêteuses. Nous les distinguons pour le chinois en trois sous-ensembles :

- *Caractère(s) chinois + éléments anglais.*

Cette combinatoire est la forme la plus souvent constatée. L'emplacement des éléments étrangers est libre dans la structure. Nous parlons de la présence d'« éléments anglais » au lieu de « mots anglais » puisque nous trouvons soit des mots soit des lettres anglais dans la structure. Dans l'expression « hold 住 » (qui signifie maintenir la situation, garder le contrôle de...), le mot anglais est placé au début de l'expression. Cependant, la désignation des comptes vérifiés (et accompagnés d'un nombre de suivi important) sur *Sina Weibo*, « 大 V » est composé par une lettre anglaise « V » (il s'agit de la première lettre du mot *verified*) qui est mise à la fin de l'expression. En nous référant à leurs contextes de création, nous pouvons comprendre que ces deux exemples sont des créations propres de la langue chinoise et que les expressions équivalentes n'existent pas encore en anglais au moment de création. Pour « hold 住 », il est employé lors d'un sketch présenté à la télévision en 2011. L'objectif du comédien est de créer un personnage qui aime à se faire remarquer par la maîtrise d'une langue étrangère. Donc, l'utilisation de la forme mixte correspond au caractère du personnage et peut ajouter à la fois un effet humoristique et ironique. Quant à l'expression « 大 V », elle a été créée pour nommer d'une manière courante et pratique les comptes vérifiés sur *Sina Weibo*. L'utilisation de la lettre V est d'une part associée au mot *verified*, et d'autre part rappelle le badge spécifique de la plateforme (le V) pour identifier un compte certifié. Cela constitue une pratique caractéristique de la culture de l'Internet chinoise sachant que *Twitter* n'a pas adopté le même signe de vérification. En plus, la création de telles expressions montre l'influence augmentée des comptes individuels sur la pratique de communication en ligne.

---

<sup>178</sup> Rappelons que nous appliquons au chinois une typologie élaborée avec des exemples français mais sans restriction exclusive à cette langue.

- *Pinyin chinois + mot (s) anglais.*

Nous avons précisé dans le Chapitre I que le pinyin constitue le système de transcription phonétique du chinois contemporain. Cependant, l'utilisation du pinyin dans la communication écrite (non normative) devient une pratique très constante indépendamment du niveau de maîtrise de la langue et de son écriture. Plusieurs raisons possibles peuvent expliquer cette pratique : i) la saisie du pinyin est plus rapide et directe que si l'on cherche les caractères correspondants à travers le clavier numérique et les systèmes de saisie numérique ; pour réaliser une communication instantanée, le pinyin est parfois adopté mais il se peut aussi que l'apparition du pinyin soit dû à une fausse saisie ; ii) pour s'exprimer en dialectes, la transcription phonétique pouvant combler la lacune de caractères pour écrire les dialectales; iii) pour transmettre indirectement un message à un destinataire averti ou en connivence avec l'auteur<sup>179</sup>. Rappelons que le chinois contemporain ou mandarin est une langue tonale dont les tons servent à distinguer le sens de chaque syllabe pour ensuite l'associer à un caractère donné. Le pinyin employé dans la communication souvent ne comporte pas de tons, ce qui peut provoquer un obstacle pour la compréhension immédiate : le lecteur doit chercher le caractère non-dit soit selon le contexte soit par sa connaissance préalable. C'est pourquoi il est très rare de trouver des expressions longues ou des phrases en pinyin et il est difficile de retrouver les caractères correspondants dans ce cas-là pour pouvoir saisir le sens.

Prenons un exemple dans notre corpus, *no zuo no die*. Il s'agit d'une création à partir de la phrase 不作死就不会死 (*bù zuō sǐ jiù bù huì sǐ*) qui signifie : il ne faut pas jouer avec le feu. Le pinyin *zuo* est l'abréviation du terme *zuō sǐ* (jouer avec le feu) qui est issu des dialectes de l'Est de la Chine. Sa structure *no...no...* est inspirée de la locution anglaise *no pain, no gains* alors que l'ensemble de l'expression est grammaticalement incorrecte en anglais. Cette forme mixte est créée suite à la popularité de la phrase chinoise. Nous supposons que l'utilisation d'une forme mixte correspond aux raisons suivantes : i) une forme plus concise et pratique à appliquer ; ii) un effet spécifique d'imitation de la structure anglaise. Ainsi, l'emploi du pinyin au lieu de caractères peut

---

<sup>179</sup> En effet, la compréhension du message en *pinyin* n'est pas directe dans la mesure où il faut restituer les tons pour atteindre le sens du message. Ainsi cette pratique peut correspondre à une volonté de contourner la censure ou de jouer sur des tons voisins.



se traduire comme un acte volontaire pour chercher le rapprochement graphique. La structure est tellement réussie que l'expression est choisie comme une nouvelle entrée du *Urban Dictionary*<sup>180</sup> en 2014. De ce fait, nous considérons que l'apparition de cette forme mixte peut être associée à une pratique chinglish<sup>181</sup> plutôt qu'à un acte d'emprunt.

- *Jeux des paronymes*<sup>182</sup> *et traduction littérale.*

Dans ces cas-là, les éléments étrangers ne sont pas visibles graphiquement. C'est souvent à travers leurs structures graphiques longues que nous pouvons premièrement les repérer. Ainsi un mot anglais plurisyllabique sera transcrit en plusieurs caractères chinois. La traduction littérale d'un mot étranger demande toujours à être explicitée en chinois bien que le sens échappe. Mais les nouvelles expressions phraséologiques deviennent de plus en plus nombreuses comme nous le constatons dans notre corpus (c'est le cas que nous allons aborder dans la partie suivante). Pour saisir cette altérité non évidente, des connaissances sur les langues étrangères sont probablement nécessaires. Il faut souligner que l'influence des langues étrangères porte sur la structure formelle et que la part chinoise de la création se manifeste plutôt à travers le sens connotatif. De ce fait, nous considérons ces cas comme créations chinoises et non comme emprunts. Par exemple, la nouveauté langagière 压力山大 (yā lì shān dà, « trop stressé ») est le paronyme d'un prénom 亚历山大 (yà lì shān dà). Ce dernier est lui-même la transcription chinoise du prénom anglais « Alexandra ». Cet paronyme n'est pourtant plus utilisé comme un prénom mais sert à décrire le stress et la pression lourde en employant une métaphore. Le niveau de stress est comparé au poids d'une montagne insupportable pour un être humain. Cette nouveauté langagière a pour but de montrer la pression sociale de plus en plus ressentie par les Chinois ; elle a rencontré un grand écho parmi les utilisateurs qui l'ont beaucoup partagée. Pour le cas de la traduction littérale, un autre exemple est le terme 友谊的小船, 说翻就翻<sup>183</sup> (yǒu yì de xiǎo chuán, shuō fān jiù fān) qui signifie « la fragilité de la relation interpersonnelle ». La première partie du terme 友谊的小船 s'agit de la traduction littérale du mot anglais *friendship*. 友

---

<sup>180</sup> Un dictionnaire collaboratif en ligne anglophone lancé en 1999, et dont les définitions sont écrites par les internautes inscrits au site.

<sup>181</sup> Traductions littérales du chinois vers l'anglais.

<sup>182</sup> Voir note n°110.

<sup>183</sup> Le sens contextualisé est exposé dans le Tableau n°27 du Chapitre VI. Il s'agit de la nouveauté langagière n°15.

谊 (yǒu yì, « amitié » en français) est la traduction du composant *friend* et *ship* est traduit comme 小船 (xiǎo chuán, « petit bateau ») en chinois.

Les cas que nous venons d'évoquer montrent que les traces des éléments étrangers ne se manifestent pas d'une manière simpliste. Il ne convient pas de considérer tous les néologismes qui comportent de telles traces comme des emprunts. Sablayrolles et Jacquet-Pfau proposent une méthode « pratique » pour identifier un emprunt : il y a emprunt quand une analyse interne (morphologique et sémantique) n'est pas possible. Cette méthode est opportune mais n'est pas suffisante : nous pensons qu'une analyse à la fois linguistique et extralinguistique est nécessaire pour chaque nouveauté langagière, quelle que soit la langue. Parmi les éléments extralinguistiques (le contexte, les relations entre les langues, etc...), les pratiques qui font naître les créations langagières sont également à ne pas négliger et nous avons notamment réfléchi aux pratiques de traduction.

#### *Discussion sur les emprunts : la traduction*

Dans les explications détaillées sur les matrices de sa typologie, Sablayrolles (2019) indique que les traductions ne doivent pas être considérées comme des emprunts quand des procédés de création sur le plan sémantique sont également constatés. Il donne l'exemple de *souris/mouse*. Selon lui, le nouveau sens d'ustensile informatique de *souris* est formé grâce à une matrice interne, la métaphore. Il a choisi alors de privilégier cet aspect rhétorique au lieu de sa source de création anglaise. Nous ne sommes pourtant pas complètement convaincue par cet exemple. En effet l'emploi métaphorique du mot « souris » ne crée pas tant une nouvelle expression en français qu'il ne montre, par la transposition d'un sens attesté en anglais dans un mot français, l'effet stylistique de la langue prêteuse. La métaphore n'est pas un procédé initiatif et créatif dans cette pratique de traduction.

Pour les cas de traductions souvent mêlés à plusieurs autres procédés dans notre corpus, nous pensons qu'il convient d'investiguer la situation de traduction, ses objectifs et ses stratégies. Par exemple, nous distinguons d'une part la pratique consistant à introduire une notion usuelle ou nouvelle de la langue prêteuse et d'autre part une pratique de traduction occasionnelle et connotative pour des produits culturels. Nous considérons le premier cas comme un emprunt et le dernier comme une création chinoise. C'est ainsi que les traductions (新常态 xīn cháng tài

et 低碳 dī tàn) des notions anglaises *new normal* et *low carbon* sont pour nous des emprunts qui introduisent deux notions issues des domaines économique et écologique dans la langue chinoise ; la translittération est principalement adoptée pour créer une nouvelle combinatoire. En revanche, la nouveauté langagière 正能量 (zhèng néng liàng)<sup>184</sup> est selon nous une création incluse dans les matrices internes. L'expression 正能量, est utilisée d'une manière créative lors de la traduction du titre du livre anglais *Rid it up* dont l'auteur, Richard Wiseman, est un psychologue. Le terme qui signifie « énergie positive » correspond à la compréhension synthétique du traducteur sur le contenu du livre. En outre il n'existe pas une telle expression spéciale dans le livre ou dans la langue anglaise. Les objectifs sont bien différents dans ces deux exemples.

Nous pouvons résumer notre réflexion en disant que l'emprunt n'est pas un procédé linguistique simple. Les éléments étrangers ne constituent pas un critère déterminant pour l'identification des emprunts. Ainsi, la distinction entre les matrices internes et externe peut varier d'une nouveauté langagière à l'autre. Aucun critère universel ne peut être appliqué pour cette distinction. Il convient d'analyser chaque cas relevé avec un point de vue dynamique et ouvert. Voyons maintenant ce qu'il en est des matrices internes dans le corpus chinois.

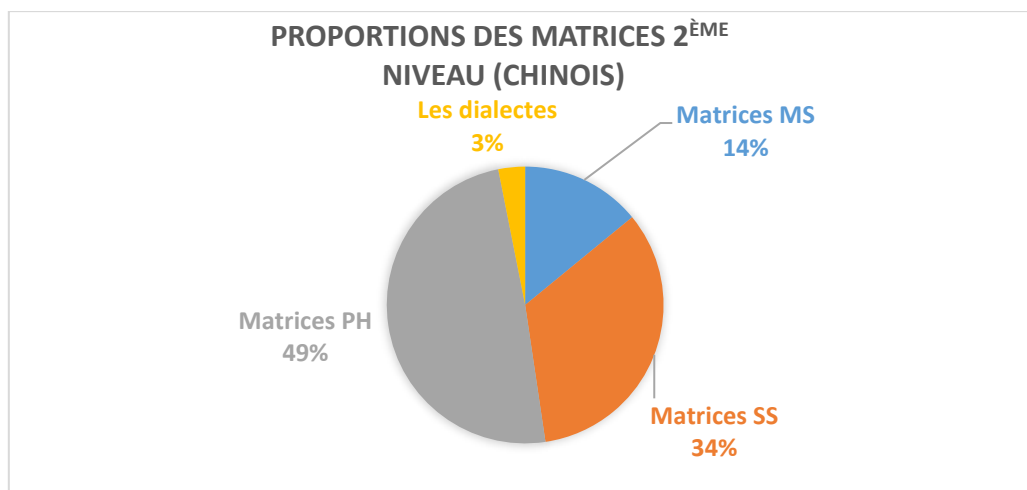
### ***5.3.2.2 Matrices internes***

#### *Proportions des différentes matrices internes*

Le diagramme (Figure 23) que nous présentons ci-dessous nous permet d'observer les proportions des différents procédés de 2<sup>e</sup> niveau, au sein de l'ensemble des matrices internes pour notre corpus.

---

<sup>184</sup> 正 (zhèng) signifie positif ; 能量 (néng liàng) peut être traduit directement comme énergie.



**Figure 23 : Proportions des différentes matrices internes dans le corpus de listes chinois**

Nous constatons que dans les cas de création internes, environ la moitié des nouveautés langagières chinoises est issue des procédés phraséologiques. La part des changements syntaxiques et sémantiques de sens est aussi remarquable avec plus d'un tiers de l'ensemble. En revanche, nous avons une proportion bien plus faible de procédés concernant le plan morphosémantique. Nous avons ajouté une catégorie supplémentaire qui est celle des dialectes sur laquelle nous reviendrons.

Nous avons choisi une typologie basée sur les analyses des néologismes français comme référence. En classant les nouveautés langagières chinoises, nous avons vu intervenir les différences linguistiques entre le français et le chinois et nous avons dégagé les caractéristiques des nouveautés langagières chinoises collectées : ces deux éléments importants conditionnent la décision d'appartenance à telle ou telle catégorie. Nous avons alors été amenée à réfléchir à deux questions principales. La première repose sur l'adéquation de l'application des matrices morphosémantiques aux nouveautés langagières chinoises. La deuxième est la suivante : comment traiter les procédés de création mixtes constatés ?

En fait, les observations montrent que les nouveautés langagières sont souvent le résultat des procédés mixtes. Au cours de l'identification, afin d'aboutir à une proportionnalité, nous avons choisi de privilégier telle ou telle matrice mais nous avons conscience que d'autres choix

auraient été possibles.<sup>185</sup>

*Les matrices morphosémantiques sont-elles adéquates pour décrire les créations chinoises ?*

Selon les définitions des matrices, les néologismes simultanément observés sur le plan morphologique et sémantique sont inclus dans les matrices morphosémantiques. Donc, le changement formel constitue un des critères indispensables. Vient alors la question de savoir si les nouveautés langagières chinoises sont toutes exclues de ces matrices puisque les caractères chinois sont morphologiquement invariables. Il est opportun de rappeler ici notre discussion dans le Chapitre III sur l'unité de néologisme en chinois. En partant de la distinction et de la relation entre les notions 字 (caractère) et 词 (unité minimale associée à un sens), nous avons montré dans un premier temps que l'analyse de l'opération néologique doit partir de la dimension du « 词 » et non du caractère. Cette dimension peut être comparable à la notion de « lexie ». Ainsi, la nouvelle forme se traduit plutôt dans la langue chinoise par une nouvelle combinaison structurelle des caractères quand l'expression en comporte plusieurs. Mais cette combinaison est moins perceptible que le nouveau sens généré puisqu'aucun changement formel n'intervient sur chaque caractère concerné. Nous arrivons donc à la première réponse suivante : il ne convient pas de privilégier à la fois les changements morphologiques (ou plutôt combinatoires) et sémantiques pour les néologismes chinois que pour les néologismes français. Dans les cas où plusieurs matrices interviennent, nous choisissons de privilégier les matrices PH.

La difficulté de retrouver des procédés morphologiques définis pour le français dans les productions chinoises, ne tient pas seulement au fait que les deux systèmes de langue sont très différents, mais aussi à des spécificités observées sur ces nouveautés langagières chinoises. Le système de la langue chinoise distingue les procédés de construction et les procédés d'imitation et de déformation comme dans notre typologie de référence appliquée au français.

Les caractéristiques linguistiques du chinois et les spécificités des nouveautés langagières nous conduisent à opter pour la décision suivante : pour intégrer un élément à la matrice morphosémantique, l'aspect signifiant doit être prioritairement pris en compte. Si cet aspect

---

<sup>185</sup> Voir l'annexe n°2d

n'est pas visible ou difficile à définir, il convient de privilégier l'aspect signifié et d'effectuer l'analyse selon les autres matrices disponibles. C'est pourquoi nous rangeons dans les matrices syntactico-sémantiques et phraséologiques la plupart des nouveautés langagières.

Toutefois, il existe deux catégories que nous considérons comme morphologiquement remarquables et qu'il est possible selon nous de placer dans la matrice morphosémantique. Ce sont les formes mixtes avec la présence d'éléments étrangers dont nous avons parlé plus haut et les créations issues de la paronymie que nous allons aborder.

### *Les cas de formes mixtes et de la paronymie*

En ce qui concerne les influences des éléments étrangers, nous avons distingué précédemment les emprunts des créations chinoises que nous avons placées dans les matrices internes en soulignant leur originalité. Nous avons constaté principalement trois manifestations différentes. Nous nous demandons maintenant si ces manifestations peuvent être analysées par une même matrice et quel critère il convient d'appliquer. En nous centrant sur la présence directe ou non des éléments étrangers, nous pouvons diviser ces créations en deux catégories qu'elles apparaissent en caractères chinois ou en pinyin. Les mots et la structure peuvent être concernés. Par exemple, dans APEC 藍, les sigles anglais APEC sont présents alors que dans 友谊的小船, l'influence du terme anglais *friendship* n'est pas perceptible à première vue. La forme mixte provoque une altérité visible dans la combinaison des éléments étrangers. L'ajout des mots et/ou de la structure étrangère répond de la part des créateurs à des objectifs variés. Rappelons que l'exemple hold 住 mentionné *supra* est destiné à produire des effets comiques. Nous trouvons de ce fait qu'il est nécessaire de souligner l'aspect formel et c'est pourquoi ces formes mixtes sont placées prioritairement dans les matrices morphosémantiques. Il est à noter que les éléments linguistiques étrangers ne sont pas le seul composant des formes mixtes. Des signes typographiques comme +, @ sont aussi concernés.

Un autre cas qui joue visiblement sur le signifiant dans les nouveautés langagières chinoises concerne les créations paronyme. Ces créations ne sont pas nombreuses dans notre corpus mais elles sont significatives pour le changement formel. Est-il pertinent de placer toutes ces créations dans les matrices morphosémantiques ? La réponse n'est pas aisée car les procédés de

création sont souvent mixtes. Par exemple, l'expression 围脖 (wéi bó)<sup>186</sup> qui est une appellation familière du microblog *Weibo* est créée non seulement par rapprochement phonétique avec la prononciation du terme chinois 微博 (weibo)<sup>187</sup> mais aussi par métaphore. Le message limité à 140 caractères est en effet comparé à un cache-cou qui a une taille réduite par rapport à l'écharpe. L'association d'un nouveau produit numérique à une expression déjà existante par le procédé de la paronymie a pour objectif de mettre en lumière la caractéristique du microblog, à savoir la contrainte de brièveté. Vu cela, il est intéressant d'identifier les effets créés par paronymie. Nous avons ainsi constaté les cas suivants :

- pour produire un rapprochement sémantique avec d'autres caractères de la langue afin de mettre en évidence certaines caractéristiques, comme dans l'exemple 围脖 (wéi bó) qui vient d'être exposé;
- pour illustrer l'accent du dialecte, voir l'exemple de 蓝瘦, 香菇 (lán shòu, xiāng gū)<sup>188</sup> ;
- pour créer une nouvelle combinaison des caractères qui est associée à un sens nouveau, voir l'exemple 压力山大 (yā lì shān dà, qui signifie stressé) présenté page 190.

Ces distinctions nous ont conduite à apercevoir que c'est dans le premier cas que la nouveauté formelle est privilégiée. La création joue sur la recherche des différences formelles pour souligner l'altérité d'expression. Donc, nous les avons mises dans la matrice morphosémantique.

#### *Certains cas contestables pour les matrices PH*

Les réflexions sur les matrices MS nous ont amenée à regrouper prioritairement sous les matrices PH les longues expressions ou les phrases pour lesquelles nous n'avons pas trouvé d'analyse morphologique pertinente. Cependant, nous avons indiqué précédemment (voir chapitre III) que les phrases ne peuvent pas être considérées comme des « lexies » et traitées comme des néologisme selon Sablayrolles. Rappelons que notre choix de nommer notre sujet

---

<sup>186</sup> L'expression signifie originellement « cache-cou » en chinois.

<sup>187</sup> Prononciation proche mais caractères différents.

<sup>188</sup> « 蓝瘦, 香菇 » (lán shòu, xiāng gū, *litt.* : « bleu et mince, champignon shiitake ») est un paronyme de « 难受, 想哭 » (nán shòu, xiǎng kū « trop triste et avoir envie de pleurer »). Le sens contextuel est développé en détail dans le Tableau n°27 (terme n°18) du chapitre VI.

de recherche « nouveauté langagière » et non « néologisme » répond notamment au souhait de prendre en compte des possibilités dynamiques de création, éventuellement en dehors du cadre néologique reconnu. Les nouvelles phrases constituent une de ces possibilités selon nous.

D'une part, leur quantité importante dans les créations langagières chinoises montre qu'elles deviennent une pratique fréquente et non négligeable. D'autre part, nous avons constaté qu'elles sont plutôt employées comme des syntagmes en perdant leur valeur de phrase. Elles deviennent ainsi fonctionnelles (un des critères qui détermine une lexie) dans la communication. Donc, les phrases relevées ne doivent pas être exclues de notre recherche. Nous remarquons également le figement de sens et de structure. C'est en fonction de ces deux dimensions de figement que nous avons relevé dans la partie suivante certains cas à préciser et à déplacer dans les autres matrices.

#### *Une nouvelle structure en détournement constant*

Un phénomène de création bénéfique en chinois d'une dénomination également nouvelle : 体 (tǐ)<sup>189</sup>. Il s'agit d'une structure partiellement figée dont certains éléments peuvent être remplacés selon la pratique des usagers et le contexte de communication. Par exemple, dans la structure « 一言不合就 XX », les signes XX indiquent la place où le remplacement est possible. La traduction littérale est la suivante : « Si on est en désaccord, on fera quelque chose immédiatement ». Le sens global est « une action soudaine ». XX peut être associé à des actions diverses, par exemple : chanter, danser, se disputer, etc. C'est à travers le changement d'élément que le nouveau sens connotatif connaît une utilisation et une circulation élargies. Voici un autre exemple, celui de la structure 舌尖体 (shé jiān tǐ). Cette structure est créée à partir de l'expression 舌尖上的中国 (shé jiān shàng de zhōng guó, *litt.* « La Chine sur le bout de la langue ») qui est le titre d'un documentaire chinois destiné à présenter les spécialités culinaires chinoises. Le grand succès du documentaire rend l'expression populaire et 舌尖 (shé jiān, « le bout de la langue ») est ainsi associé à la gastronomie. Nous pouvons constater les variations suivantes : 舌尖上的快乐 (shé jiān shàng de kuài lè), 舌尖上的爱情 (shé jiān shàng de ài qíng). Dans le premier cas, 中国 (zhōng guó, « la Chine ») est remplacé par 快乐 (kuài lè,

---

<sup>189</sup> Il signifie originellement un genre ou style littéraire.



« la joie »). La nouvelle expression signifie « le plaisir de goûter les plats ». Dans le deuxième cas, 爱情 (ài qíng, « l'amour ») est utilisé à la place de 中国 (zhōng guó, « la Chine ») pour décrire une relation amoureuse nouée par la gastronomie. Mais, il se peut que le remplacement ne soit pas obligatoire puisque le sens connotatif ne sera pas affecté. La structure qui s'appelle 元芳体 (yuán fāng tǐ) <sup>190</sup> en est un exemple représentatif. Elle est formée à partir de l'expression 元芳, 你怎么看 (yuán fāng, nǐ zěn me kàn) qui signifie littéralement « Yuan fang, qu'en pense tu ? ». Cette expression est originellement utilisée par le procureur pour demander à son assistant Yuan fang de donner ses analyses d'un procès dans un feuilleton télévisé policier. Elle est ensuite employée par les internautes pour exprimer ironiquement leurs doutes sur un résultat constaté, notamment dans les faits d'actualité. Dans cette structure, le prénom Yuan fang peut être remplacé par les autres prénoms selon les usagers. Mais le changement de prénom ne va pas modifier le sens global de l'expression. Dans tous les cas mentionnés ci-dessus, l'interprétation de la structure demande préalablement une connaissance des expressions et phrases d'origine.

Nous avons d'abord analysé ces nouveautés langagières<sup>191</sup> dans les matrices phraséologiques comme un sens figé directement associé à une expression ou phrase dont la structure est issue d'une création nouvelle. Or, la possibilité de changement combinatoire favorisée par la structure ne peut pas rester inaperçu. S'agit-il de ce fait d'une combinatoire syntaxique ? Il faut souligner que le changement combinatoire ne forme qu'une application occasionnelle qui contribue en réalité à la mémorisation du figement d'origine (sur le sens et sur la structure). La transgression combinatoire est donc absente. Vu l'importance de la structure qui prend le pas sur le figement de sens, nous pensons qu'il est pertinent d'analyser ces nouveautés langagières selon les matrices morphosémantiques en privilégiant cet aspect formel.

#### *Les cas de concurrences entre les matrices PH et SS*

Les matrices syntactico-sémantiques disposent de deux grands sous types : le changement de fonction et le changement de sens. Dans notre corpus, c'est le dernier sous-type « le changement

---

<sup>190</sup> Il s'agit du terme n°1 du Tableau n°27 des 21 termes communs dans le chapitre VI où nous avons exposé en détail le sens contextualisé.

<sup>191</sup> 一言不合就 XX, 可能 XXX 假 XXX, 微 XX, 友谊的小船, 舌尖上, 待我长发及腰, 被就业.

de sens » qui constitue le cas le plus constaté. Or, dans le classement temporaire des matrices PH, nous avons aussi remarqué des expressions et phrases dont le sens d'origine diffère de leurs sens figés associés. Ces origines sont multiples. Par exemple, 最炫民族风 (zuì xuàn mín zú fēng) et 时间都去哪儿了 (shí jiān dōu qù nǎ'èr le) sont originellement des titres de chanson. Les titres d'émissions télévisuelles ou les expressions issues des campagnes de publicité ou des productions culturelles sont aussi concernés<sup>192</sup>. Le sens d'origine va avec des usages dans un certain cadre contextuel prévu. Mais dans une phase ultérieure, les usagers ayant la connaissance de ce cadre contextuel leur ont attribué un sens connotatif selon leurs intentions de communication (exprimer un sentiment, commenter un fait social). La chanson « 时间都去哪儿了 » (shí jiān dōu qù nǎ'èr le, *litt.* « Où va le temps ») a été diffusée pour la première fois au Gala de Nouvel An 2014 de CCTV<sup>193</sup>. Les paroles de cette chanson transmettent l'idée suivante : « Chérissez vos parents et prenez du temps pour eux car plus rien au monde ne remplace leur amour lorsqu'ils ne sont plus là ». Une fois diffusée, la chanson est très vite devenue populaire en Chine. Le titre de chanson est ensuite employé dans les communications en ligne comme une expression figée pour signifier « le gaspillage inconscient du temps ».

Il est à noter que le nouveau sens intervient dans un nouveau contexte mais que l'interprétation est également déterminée par le contexte d'origine. Dans le cas de l'expression 时间都去哪儿了, la chanson souligne le fait que « le temps passe très vite » et c'est pourquoi le titre de cette chanson est associé ultérieurement au « gaspillage du temps ». Il faut ainsi distinguer ces nouveautés langagières de celles qui sont employées dans une situation donnée et ensuite popularisées. Par exemple, 我妈是我妈 (wǒ mā shì wǒ mā, *litt.* « ma mère c'est ma mère ») est originellement associée à la situation où il est demandé à un citoyen chinois de fournir une attestation complémentaire qui atteste que « sa mère est sa mère » lors de démarches administratives. L'expression est ensuite largement utilisée avec une ironie à l'égard du processus administratif.

Deux critères peuvent nous aider à la distinction i) il s'agit une structure déjà existante ; ii)

---

<sup>192</sup> Voici ces éléments en chinois. Le titre d'une émission télévisuelle : 爸爸去哪儿 ; les expressions issues des campagnes de publicité ou des productions culturelles : 挖掘机技术哪家强, 那么问题来了, 躺着也中枪, 我读书少你别骗我.

<sup>193</sup> Voir note n°128 dans le chapitre IV.

l'interprétation dépend de plusieurs contextes. Si la réponse est négative pour ces deux questions, ces nouveautés langagières sont plutôt considérées comme des créations phraséologiques. Cette distinction nous a conduit à opter pour un classement dans les matrices SS de certaines nouveautés langagières.

Nous abordons maintenant la question de l'extension et de la restriction du sens : Sablayrolles a exclu ces deux cas des changements sémantiques. Selon lui, ce sont plutôt des évolutions de sens à travers le temps. Ce processus de formation est progressif et non perceptible par les usagers, alors que pour les nouveautés langagières, même s'il existe un décalage entre le sens d'origine et le sens nouvellement figé, nous pouvons découvrir que le nouveau sens est apparu d'une manière imprévisible et brutale, ce qui constitue leur caractéristique (Voir Chapitre III). Ainsi, sans fournir un effort, les locuteurs auront du mal à les comprendre. De ce point de vue, il est opportun d'analyser ces expressions et phrases dans les matrices SS.

*Comment traiter les influences de dialectes ?*

N°	Termes liés aux dialectes	Sources
1	不差钱 / bù chà qián	dialecte du nord de la Chine
2	坑爹 / kēng diē	dialecte du nord de la Chine, notamment de la province de Shandong
3	失联 / shī lián	dialecte du sud de la Chine
4	怼 / duǐ	dialecte de la province Henan

**Tableau 25 : Liste des termes**

Pour certains cas de créations influencés par les dialectes, nous hésitons toujours sur la matrice à laquelle les rattacher (voir Tableau 25). En effet, leurs procédés de formation sont complexes. La transposition des expressions dialectales à l'écrit concerne les dimensions phonétique, sémantique et combinatoire. Par exemple, l'expression (4) 怼 (duǐ)<sup>194</sup> est associée au dialecte

<sup>194</sup> Le sens contextualisé est expliqué dans le tableau n°27 du chapitre VI. C'est le terme n°20 qui signifie « haïr, offenser, critiquer ».

de la Province Henan à la fois pour la prononciation et le sens. Il faut souligner que la prononciation originale du caractère (duì) est différente de la prononciation dialectale en ce qui concerne le ton qui sert en chinois à distinguer les sens.

Pour le cas de l'expression (1) 不差钱 (bù chà qián), c'est la structure combinatoire des caractères et le sens qui sont liés au dialecte du nord de la Chine. Il est difficile de privilégier un des procédés de création plutôt qu'un autre car le plus visible c'est leurs effets convergents pour mettre en relief le dialecte.

Il faut encore préciser que les influences des dialectes en Chine sont souvent perçus comme comiques. Reprenons l'exemple 不差钱 (bù chà qián) (1), reconnu par le public après avoir été utilisé dans un sketch diffusé à l'échelle nationale pendant le Nouvel An chinois. Ce genre de sketch, souvent présenté en dialecte du nord de la Chine, est un programme du Gala de Nouvel An de CCTV. Regarder le Gala télévisé est la routine des Chinois la veille de Nouvel an comme nous l'avons vu *supra* dans le chapitre IV. Donc, en inspirant par exemple un sketch, les dialectes sont employés par les Chinois pour créer des effets comiques. Avec les contacts d'internautes de régions différentes, d'autres dialectes que le dialecte du Nord ont été employés dans la communication.

Notre solution est que nous mettons temporairement les dialectes dans une catégorie indépendante pour le calcul des proportions.

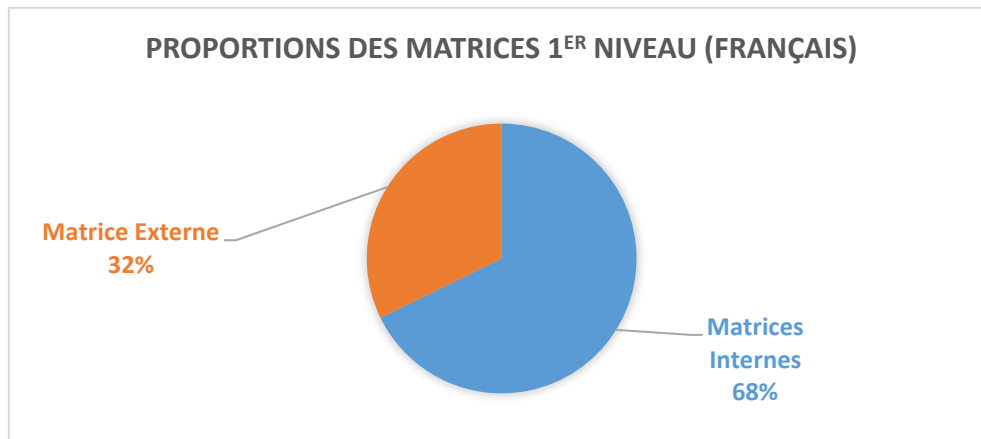
### **5.3.3. Le corpus de listes français**

Le corpus de listes français est composé par deux sous-corpus dont le total donne 371 résultats. Après avoir exclu les répétitions, nous avons obtenu 273 résultats effectifs. C'est à partir de ce nombre que nous avons effectué des calculs de proportion présentés ci-dessous. Chaque entrée répertoriée est étiquetée selon les catégories de la typologie des néologismes, le détail figurant en annexe 3d.

### 5 3.3.1. Matrices internes / matrice externe et différentes matrices internes

*Proportions d'ensemble : première approche de l'emprunt*

Le premier résultat que montre la Figure 24 concerne la proportion du 1<sup>er</sup> niveau des matrices internes et externe dans la première approche des emprunts qui a été la nôtre.



**Figure 24 : Proportions des matrices internes et externe dans le corpus de listes français**

Le diagramme de la Figure 24 ci-dessus montre les proportions qu'occupent respectivement la matrice externe (32%) et les matrices internes (68%) dans les procédés de création des nouveautés langagières françaises. La relation est d'un tiers à deux tiers et il faut d'emblée reconnaître la part importante de l'emprunt qui est le seul procédé de la matrice externe. Celui-ci n'est pas sans poser de problèmes de délimitation, comme nous l'avons vu à propos du corpus chinois *supra*. Au terme de la discussion sur cette délimitation, nous allons proposer une autre répartition proportionnelle.

Pour parvenir à ce premier résultat proportionnel (Figure 24) nous avons d'abord identifié les emprunts d'une manière intuitive en prenant en compte tout ce que nous percevions dans notre corpus comme influence d'éléments étrangers. Cette décision provisoire est motivée par deux raisons :

- certaines nouveautés langagières procèdent à la fois des matrices internes et externe et pour décider il faut choisir de privilégier un aspect plutôt qu'un autre, et faire à cette fin une analyse

approfondie ; avant d'effectuer cette démarche, ce sont les influences étrangères, visibles, qui sont pointées ;

- le statut d'élément étranger ou non étranger est problématique et recoupe celui des variations internes à une même langue : d'un côté, les nouveautés langagières témoignent de contacts de langue entre le français et beaucoup d'autres langues (aussi bien européennes qu'asiatiques par exemple) ; d'un autre côté, les variations du français, régional ou non hexagonal sont aussi présentes.

En intégrant tous ces phénomènes, nous nous sommes appuyée temporairement sur une définition large de l'emprunt<sup>195</sup> et c'est ainsi que nous avons obtenu le résultat de 32% du nombre total pour la matrice externe.

#### *Définition plus restrictive de l'emprunt et incidence sur les résultats*

Mais dans un deuxième temps, nous nous sommes ralliée au point de vue de Sablayrolles (2019:175-183) pour qui les influences étrangères ne sont pas manifestées seulement à travers la matrice externe et pour qui il est convenable d'adopter une conception restreinte de l'emprunt au lieu d'assimiler toutes les manifestations étrangères à des emprunts comme le font certains chercheur.e.s qui s'intéressent à l'innovation lexicale (Cartier 2019 ; Tallarico 2020).

L'emprunt est un processus régulier de l'enrichissement du lexique. Mais les caractéristiques formelles peuvent être influencées par le temps et le degré d'intégration et engendrent éventuellement des critères d'identification variés. Partant d'un point de vue synchronique et pour souligner une influence directe des éléments étrangers dans le français contemporain, Sablayrolles considère ainsi les unités lexicales étrangères conservées sous forme initiale et/ou peu adaptées comme de véritables emprunts à distinguer des faux emprunts, des calques (morphologiques et sémantiques) et des synthèses néologiques<sup>196</sup> sous influence étrangère.

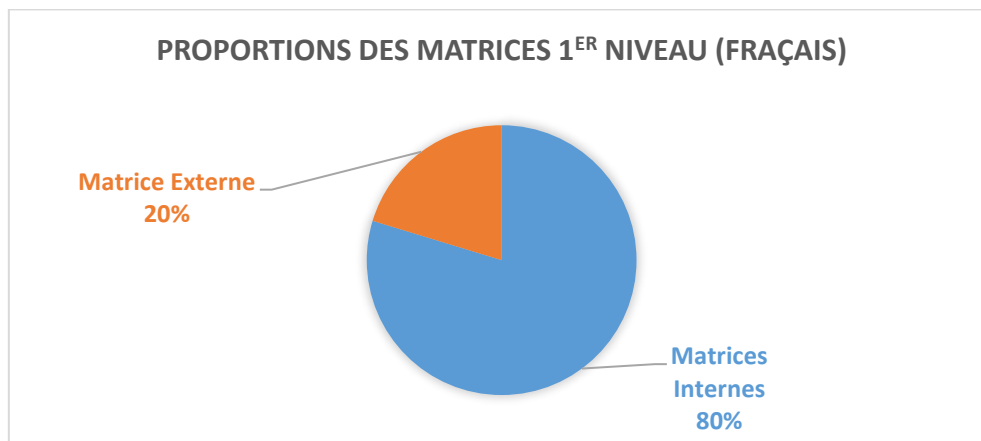
En adhérant à cette conception restreinte qui s'applique à notre classification en annexe 3c,

---

<sup>195</sup> « L'emprunt est une forme d'expression qu'une communauté linguistique reçoit d'une autre communauté » (Deroy 2013).

<sup>196</sup> Ces catégories seront abordées dans le sous chapitre 5.3.3.1.

nous avons obtenu un nouveau résultat que nous présentons dans la Figure 25 ci-dessous.

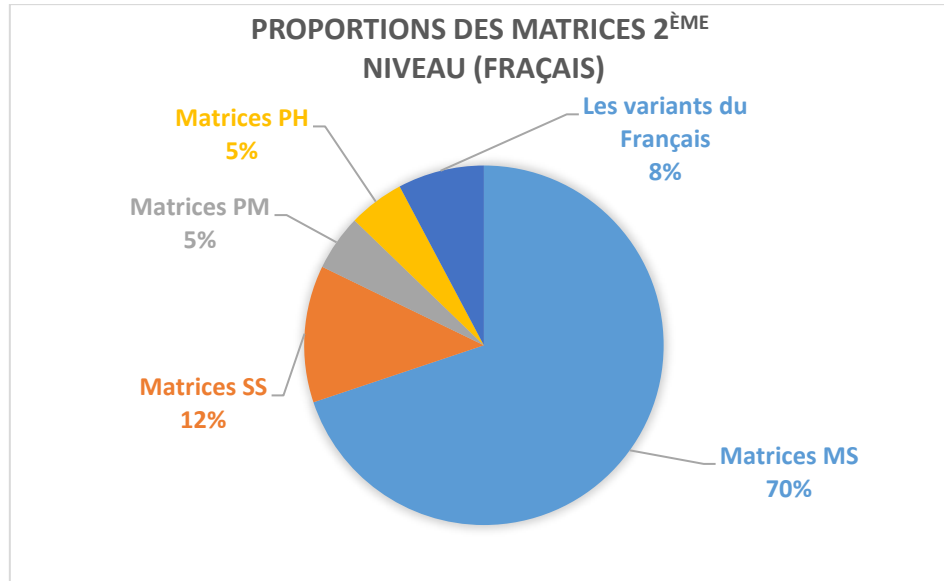


**Figure 25 : Proportions des matrices internes et externe dans le corpus de listes français à partir d'une définition restrictive de l'emprunt**

Ce nouveau diagramme (Figure 25) qui présente le résultat de la répartition entre matrices internes et externe à partir d'une définition plus étroite de l'emprunt marque une nette diminution de la proportion de la matrice externe par rapport à l'ensemble (de 32% à 20 %), mais cette matrice externe représente tout de même encore près du quart des nouveautés langagières. Comme une partie des éléments sous influence étrangère est désormais intégrée aux matrices internes, les cas significatifs rencontrés et intégrés ou écartés dans les processus d'identification seront abordés dans le sous chapitre suivant consacré au niveau des différentes matrices internes.

#### *Différentes matrices internes*

Les matrices internes rassemblent plus de trois quarts des entrées. Nous voyons avec la Figure 26 quels procédés de création sont prédominants au sein de ce nouvel ensemble :



**Figure 26 : Proportions des différentes matrices internes dans le corpus de listes français**

Au sein des matrices internes, les quatre matrices principales de deuxième niveau sont toutes représentées (MS, SS, PH, PM) et nous avons ajouté une cinquième rubrique sur laquelle nous reviendrons plus loin. Les matrices morphosémantiques sont très nettement majoritaires (70%) avec une tendance de formation par affixation qui l'emporte très nettement sur les divers procédés de composition qui sont un choix secondaire. Nous n'avons relevé que quelques créations qui correspondent aux procédés d'imitation et déformation.

Les matrices syntactico-sémantiques occupent loin derrière en termes de proportion la deuxième place parmi les matrices internes avec 13%. Mais nous pouvons remarquer que sur le plan quantitatif, elles sont moins représentées que la matrice externe des emprunts (22%). Les matrices purement morphologiques et phraséologiques sont minoritaires et comparables avec chacune une proportion de 5%. La rubrique supplémentaire des formes lexicales régionales ou non hexagonales compte un pourcentage un peu plus important que les catégories mentionnées précédemment.

Une fois les proportions précisées aux deux niveaux hiérarchiques supérieurs (1<sup>e</sup> et 2<sup>e</sup> niveaux), nous allons maintenant aborder la question des influences étrangères pour elle-même, en précisant, au fur et à mesure de l'analyse, quelle est la ligne de partage entre emprunt et non emprunt mais sans faire de cette ligne de partage le principe organisateur de notre développement.



### 5 3.3.2. *Les influences étrangères*

Tout d'abord, qu'il s'agisse d'emprunts ou d'autres procédés relevant des matrices internes, nous allons nous intéresser aux différentes langues concernées et aux thématiques représentées.

Quelles sont les langues étrangères fournissent ou inspirent le plus de nouveautés langagières et comment celles-ci interviennent-elles, en tant qu'emprunts ou autres procédés ?

Pour la catégorie des emprunts, l'influence de l'anglais ou de l'anglais américain est remarquables quantitativement (40 sur 56) aussi bien que par la diversité des domaines concernés : domaines culinaire (*cheese-cake*), culturel (au sens large dans lequel nous englobons la mode *fashionista*) ; les productions culturelles (*fanfiction*, *spin-off*) ; les phénomènes sociaux (*speed dating*, etc.), sportifs (*fixie*), économiques (*bitcoin*), numérique (*notebook*, *mème*) ; les termes informatiques (*pop-up*) ; les tendances culturelles (*selfie*...). Cette présence importante de l'anglais pourrait s'expliquer par son rôle véhiculaire dans la communication internationale.

Pour l'italien et l'espagnol, la proximité géographique de pays limitrophes de la France (Italie, Espagne) et l'appartenance à une même famille linguistique, celle des langues romanes, sont susceptibles de favoriser l'enrichissement du lexique français sans nécessiter d'adaptation particulière et dès lors, sur cette base morphologique, il s'agit plutôt d'emprunts conservés tels qu'ils sont dans les langues d'origine. Par exemple, nous avons relevé le mot espagnol *piquillo* qui signifie « type de poivron ».

Mais en ce qui concerne les langues asiatiques, même si nous avons catégorisé des emprunts, nous constatons toutefois que des modifications formelles même minimales sont indispensables à cause de la spécificité graphique et/ou phonique des langues. On a donc au moins une transposition phonétique pour le chinois (*wushu*), le japonais (*wakamé*), le vietnamien (*phô*), le thaï (*pad thai*) et aussi pour le sanscrit (*dosha*). La plupart de ces cas concernent le lexique alimentaire en désignant des spécialités culinaires et en conservant ainsi un certain exotisme culturel. Outre les langues asiatiques, on trouve aussi des mots émanant de l'arabe (*makrout*) et du portugais brésilien (*cachaça*) qui sont introduits, l'un par l'adaptation phonétique et l'autre sans modification.

La plupart des emprunts anglais ont été introduits sans modification formelle dans la langue française comme les cas de *big data*, *open source*. Mais il existe des cas pour lesquels une adaptation morphologique et graphique intervient. Par exemple : le mot anglais *prequel* est modifié en *préquelle* pour désigner en français « une épisode d'une œuvre dont l'action se situe avant celle des épisodes précédents ». Les observations montrent, de façon peu surprenante, que l'anglais constitue non seulement la langue étrangère qui fournit le plus d'emprunts dans le corpus (voir les exemples cités *supra*), mais aussi que de nombreuses nouveautés langagières répertoriées dans les matrices internes sont également étroitement liées à l'influence de l'anglais.

Nous présentons maintenant quelques cas significatifs d'influence de l'anglais mais exclus de la matrice externe en suivant les propositions de Sablayrolles.

- *Faux anglicisme* :

Il s'agit d'un mot formé dans une langue donnée à partir d'éléments d'origine anglaise mais dont l'appartenance à la langue anglaise est un leurre. La distinction entre un faux anglicisme et un emprunt relève de ces possibilités : i) la forme dans la langue d'accueil donnée n'existe pas en anglais ; ii) un équivalent en anglais peut être proche de cette forme mais pas identique ; iii) la forme existe en anglais mais le sens dans la langue d'accueil ne correspond pas à celui de anglais.

Voici un exemple de notre corpus : *scrapbooking* est formé à partir de deux mots anglais *scrap* « morceau » et *book* « livre » avec le suffixe anglais *ing* qui s'emploie souvent pour faire un substantif d'un verbe et désigner ainsi une activité. Or, ce mot composé de deux éléments lexicaux anglais et utilisant un morphème régulier en anglais – il dénomme une forme de loisir créatif consistant à introduire des photographies dans un décor selon les thèmes différents –, n'existe cependant pas en anglais. Dans ce cas-là, nous privilégions le procédé de construction par composition et nous avons placé cet exemple dans les matrice MS-Composition.

- *Calque*<sup>197</sup> :

Le calque est étroitement lié aux processus de transposition et de traduction des éléments étrangers. Plusieurs matrices internes sont concernées par la formation des calques. Dans le corpus, nous avons relevé deux types principaux de calques : calques morphologiques et calques sémantiques. Le premier type consiste à transposer un sens étranger dans une nouvelle forme issue souvent d'une traduction littérale, en adaptant ainsi une notion. Par exemple, le mot composé *mémoire flash* (qui désigne une technologie de stockage de données) est créé à partir de la traduction directe de l'expression anglaise *flash memory*. C'est dans les matrices MS-composition (MS-CC) que nous avons placé ce terme. Le deuxième type associe le sens étranger à une forme déjà existante dans la langue emprunteuse. C'est plutôt dans les matrices syntactico-sémantiques (SS) que nous pouvons repérer des exemples. Ainsi le mot *troll* (lui-même un emprunt) désignant la créature nocturne de la mythologie scandinave est doté actuellement du sens suivant : « un individu adoptant un comportement qui vise à générer des polémiques en ligne, à perturber une interaction ». Ce nouveau sens est initialement apparu dans le contexte de l'Internet en langue anglaise.

Il est intéressant d'observer au sein de la langue française des tentatives de création en parallèle des créations conventionnelles. En plus, les résultats issus de ces tentatives ont connu souvent un usage répandu. Par exemple, *prioriser* qui est la francisation de l'anglais *prioritise* est considéré comme un barbarisme<sup>198</sup> par l'Académie française. Le terme de *scrapbooking* a connu son succès en France mais son équivalent *collimage* est proposé et utilisé au Québec pour avoir un substitut français du terme (pseudo) anglais.

---

<sup>197</sup>Un calque est un emprunt assorti d'une traduction littérale des éléments et/ou d'une reproduction de certains traits formels. Voir par exemple la définition suivante :

« [...]a new word modeled after a word in another language. While, in the case of borrowing, a foreign word and its meaning are adopted wholesale into the other language as a loan word, a calque emerges when the language is adapted to new concepts. This can happen in several ways: (a) by way of a borrowed meaning through change and expansion of the meaning of native words—write (originally 'to scratch') influenced by Lat. scribe; (b) through neologisms loosely based on a foreign concept— Ger. Sinnbild for symbol; (c) through word-for-word loan translation—crispbread from Ger. Knäckebröt, accomplished fact from Fr. fait accompli, Span. rascacielos for skyscraper; (d) through a loose loan translation—brotherhood for Lat. fraternitas. Bussmann, Hadumod (dir.). *Routledge Dictionary of Language and Linguistics* [« Dictionnaire de la langue et de la linguistique »], Londres – New York, Routledge, 1998.

<sup>198</sup> <http://www.academie-francaise.fr/prioriser>, consulté le 12 décembre 2020.

- *Les hybrides :*

Il s'agit de créations combinant les éléments étrangers et français. Ce cas est souvent rencontré dans les deux grands sous-types des matrices (2<sup>e</sup> niveau) morphosémantiques, les matrices d'affixation et de composition. Par exemple, *geeker* est formé par le mot anglais *geek* (« personne passionnée par les nouveautés techniques ») et le suffixe verbal français *-er* pour former un verbe signifiant « pratiquer en fan des activités technologiques ». Outre des élément anglais, nous avons aussi constaté que des mots étaient ainsi construits en français à partir de l'influence d'autres langues. Dans la nouveauté langagière *zénifiant* (« qui rend zen »), nous retrouvons la base japonaise *zen* qui peut être considérée comme un emprunt intégré et le suffixe *-ifier* (ici le mot répertorié est au participe présent du verbe).

Nous pouvons constater qu'environ une moitié des entrées des matrices purement morphologiques (c'est-à-dire de procédés modifiant le signifiant sans affecter le signifié) sont créées à partir de fondements anglais. La troncation et la siglaison ou l'acronymie constituent les deux types principaux. La nouveauté langagière *vegan*<sup>199</sup> est un exemple de troncation. Elle est formée à partir de la suppression de la partie centrale du mot anglais *vegetarian*. Mais il est vrai que cette troncation produit un mot qui n'a pas tout à fait le même sens car un végétarien ne consomme pas de produits carnés mais ne bannit pas tous les produits (alimentaires ou autres) d'origine animale comme le fait un vegan. *MOOC* constitue un autre exemple. Il s'agit d'un acronyme formé des initiales de l'expression anglaise *massive open online course* qui signifie en français « cours en ligne ouvert et massif » et qui désigne des formations à distance.

À travers ces différents exemples, nous avons voulu mettre en relief l'influence de l'anglais qui occupe une place prééminente dans la formation des nouveautés langagières françaises. Pour ce qui est des emprunts proprement dits, l'anglais domine par rapport aux autres langues étrangères dans la matrice externe. Par ailleurs, il contribue aussi aux divers procédés de créations des matrices MS, SS et PM au sein de la langue française.

---

<sup>199</sup> Pratique qui exclut toute consommation et utilisation de produits issus de l'exploitation animale.

### *Discussion sur le classement des nouveautés au sein des matrices internes*

Nous avons été embarrassée par la frontière indécise pour nous entre matrice externe et matrices internes d'une part et par l'éventuelle pluralité des procédés relevant de différentes matrices internes pour certaines créations d'autre part. Le constat d'influences étrangères avec importation directe de lexiques étrangers peut faciliter le repérage de l'emprunt ou déboucher sur la mise en évidence de procédés de création morphologique, sémantique et syntaxique, comme nous l'avons vu avec la complexité de la notion d'emprunt et l'extension ou non de sa définition.

Par ailleurs, quand plusieurs matrices internes sont concernées, nous avons souvent hésité à en choisir une comme essentielle pour catégoriser une nouveauté avec un étiquetage unique (voir annexe 3d) et permettre ainsi un calcul quantitatif proportionnel.

Il est parfois délicat de distinguer une création morphologique d'un emprunt partiellement adapté. C'est pourquoi nous avons d'abord hésité à classer les nouveautés langagières comme *préquelle*, *se loguer* et *prioriser* qui ont des traits formels familiers du français mais qui viennent de l'anglais. Nous avons opté pour la matrice morphosémantique et non pour l'emprunt car ces mots anglais adaptés à la langue française n'existent pas dans la langue source.

Il est aussi difficile de privilégier une matrice interne par rapport aux autres quand plusieurs procédés internes interviennent. Lorsque les influences étrangères sont aussi présentes, nous nous sommes demandé s'il n'était pas plus pratique de mettre en valeur le procédé de l'emprunt (en mettant à part le cas des faux anglicismes). Par exemple, nous avons choisi d'analyser *nomophobe* prioritairement à travers la matrice PM puisqu'il s'agit de l'abréviation d'une expression *no mobile-phone phobia* avec une condensation très remarquable. Cette expression anglo-saxonne intègre elle-même un mot d'origine grecque (*phobia*) qui existe en français sous la forme autonome ou composée *phobie* et a une forte productivité par l'intermédiaire du suffixe savant – *phobe*. Cet élément présente la particularité d'être partiellement homophone de *phone* (mot anglais « téléphone ») pour la première syllabe. *Nomophobe* est à la fois une abréviation ou plutôt une composition d'éléments abrégés, une construction par affixation et un calque modifié de l'anglais.

Nous avons repéré des phénomènes langagiers qui suscitent des hésitations lors de l'analyse à travers des matrices internes. On peut les considérer en fonction du territoire pris en compte comme des variantes du français ou comme des variantes linguistiquement apparentées au système de la langue française mais extérieures parce que régionales ou étrangères et en tous les cas inusitées en français standard. Des mots venus de Lorraine (*chouille*), de la Réunion (*cari*), du français de Suisse (*chneuquer*), de Belgique (*prépension*), du français québécois (*taxage*) et du français d'Afrique (*yoyette*) sont observés. Pour une raison de conformité au système de la langue, Sablayrolles et Jacquet-Pfau (2008) proposent d'étudier ces cas traditionnellement considérés comme emprunts dans les matrices internes. Mais pour repérer le procédé, il faut connaître les formes originelles (affixation de *taxe* à *taxage* par exemple) alors que ce n'est pas toujours possible (*cari* et *yoyette* par exemple restent mystérieux) et aussi si ces éléments sont placés dans les différentes matrices internes leur source reste inaperçue. En outre, comme il s'agit d'une introduction directe (proche des emprunts par ce trait) nous trouvons qu'il ne convient pas de les analyser avec telle ou telle matrice de création. C'est pourquoi nous avons proposé de regrouper ces nouveautés langagières dans une case à part des matrices internes.

Il est certain que les procédés de formation les plus usuels dans le corpus français relèvent des matrices MS et dans ces matrices une quantité importante de mots sont formés par construction. La construction par affixation et composition domine incontestablement. Dans le premier cas (l'affixation), une base et un affixe sont associés pour la formation du mot. Les préfixes ou suffixes permettent de produire des verbes, des noms et des adjectifs. Divers domaines sont concernés, par exemple, certains domaines scientifiques comme l'écologie (*biogaz*, *écoconduite*), la médecine (*virologie*) ; les phénomènes sociaux (*locavore*, *vapoter*) ; et culturels au sens large (*microfiction*, *cinégénique*) ; l'informatique (*se loguer*, *geeker*). Dans le deuxième cas (la composition), plusieurs unités lexicales / bases se combinent. La plupart des mots composés de notre corpus sont formés à partir de l'association de deux (ou plus de deux) unités lexicales indépendantes. Nous avons remarqué plus particulièrement parmi les noms composés : des termes économiques (*dette souveraine*, *économie collaborative*) ; politiques (*État-parti*, *enfants soldats*) ; ou bien des termes techniques (*réalité augmentée*, *imprimante 3D*). Les mots valises constituent également un des sous-types des procédés de composition. Ce procédé met en jeu l'amalgamation de deux lexies qui ont une ressemblance partielle sur le plan morphologique. Nous avons choisi d'évoquer ce procédé pour deux raisons : tout d'abord,

il s'agit d'un procédé assez productif dans le corpus (par exemple *adulescent*, *bistronomie*, *émoticône*, etc.) ; ensuite, la création de mots-valises est souvent accompagnée d'une recherche d'effet particulier, ludique, ironique... Le terme *adulescent* par exemple initialement créé dans les milieux publicitaires a une portée ludique et légèrement critique pour décrire les adultes qui tentent de s'identifier aux jeunes par certains comportements de consommation ou de jeunes adultes qui ont du mal à mûrir. Cette sorte de nouveauté langagière détient une richesse stylistique qui ne laisse pas les créations inaperçues.

D'ailleurs, en nous référant aux critères de notre typologie de référence, nous avons rencontré dans le corpus deux cas de composition susceptibles d'être traités comme des cas d'affixation.

Sablayrolles s'attache à distinguer les fractolexèmes des affixes. Il s'agit « de fragments des lexies qui valent pour la lexie entière » (Sablayrolles 2019). Ainsi, deux éléments qui doivent être considérés comme des fractolexèmes et non comme des affixes sont observés : *cyber* (Raus 2001) pour tout ce qui concerne l'information (cybernétique) et *télé* pour *télévision*. Nous les trouvons dans les termes suivants de notre corpus : *cybercrime*, *cyberdépendance*, *téléverser*, etc. Pourtant, nous pensons la frontière de distinction est délicate. Tout d'abord, si *télé* est bien un fragment du mot *télévision* valant pour ce mot dans l'exemple *télespectateur* cité par Sablayrolles, ce n'est le cas du terme *téléverser* qui vient d'être mentionné ; *télé* signifiant « à distance » et emprunté du grec ancien serait un affixe plutôt qu'un fractolexème ; c'est cependant le même élément, formellement et sémantiquement. Il serait nécessaire d'étudier au cas par cas. Ensuite, nous trouvons que la définition de fractolexème ne nous permet pas d'obtenir un critère assez précis. Nous nous demandons si les formants comme *bio* et *éco* peuvent être considérés comme des fractolexèmes pour *biologie* et *écologie* ou si ce sont des préfixes ?

Le deuxième point concerne le risque de confusion entre affixation savante et composition savante lorsque plusieurs formants des langues anciennes sont associés. La propriété spécifique des formants savants est qu'ils sont libres dans leurs emplacements au sein de la composition, alors que les affixes ont une place fixe. En prenant compte cette propriété et l'absence d'autonomie des formants savants, certains chercheurs ont proposé le concept de *confixe*. Mais Sablayrolles considère que cette dénomination ne prend en compte que l'aspect grammatical qui rattache ces éléments aux procédés d'affixation. Il faut aussi souligner que Sablayrolles

considère la combinaison de plusieurs affixes comme un procédé de dérivation qui se trouve toujours au sein des procédés d’affixation et non une catégorie à part. Pour citer un exemple, le terme *anosognosie* (« absence de conscience d’une maladie ou d’un handicap») est issu d’une composition savante : il est construit à partir du grec *nosos*, « maladie » et *gnosis*, « connaissance ».

En bref, il est clair que les procédés de construction au sein des matrices MS constituent les procédés prédominants dans la formation des nouveautés langagières françaises avec des mots simples et des mots composés en grand nombre dans le corpus. Nous nous demandons si ce résultat n’est pas influencé par notre méthodologie de collecte qui aurait introduit un biais. En effet, le fait de choisir notre corpus dans une sélection d’entrées des dictionnaires tels que *Le Petit Larousse* ou *Le Petit Robert* établit préalablement un filtrage. Les nouvelles entrées dans les dictionnaires répondent à des règles conventionnelles, des normes, et on peut penser que les créations « libres » y sont plus rares à la différence des dictionnaires collaboratifs non institutionnels. D’un autre côté, un dictionnaire est un recueil dans lequel le mot est l’unité d’index privilégiée au sein d’un dictionnaire. Cela peut être une des raisons qui explique la présence importante de mots à formation régulière dans le corpus.

#### **5.3.4. Éléments de comparaison**

Nos observations concernant le 1<sup>er</sup> niveau hiérarchique des matrices dans notre typologie de référence et la proportion occupée par la matrice externe par rapport aux matrices internes nous font faire le constat que les emprunts – et plus généralement les influences étrangères –, sont présents pour les créations des deux langues mais que leurs sources et niveaux d’influence ne sont pas les mêmes. Les échanges asiatiques dans le domaine culturel peuvent être considérés comme une source principale d’enrichissement du lexique pour le chinois, alors que pour le français, les sources étrangères sont plus diverses – même si l’anglais est en tête – ainsi que les domaines concernés.

Au 2<sup>e</sup> niveau des matrices internes, nous avons constaté des points communs et des tendances spécifiques dans chaque langue. Tout d’abord, seuls trois des quatre types de matrices internes interviennent dans la formation des nouveautés langagières chinoises (matrices MS, SS et PH) : les matrices purement morphologiques sont absentes en raison de l’invariabilité



morphologiques des unités lexicales en chinois qui minore également les procédés morphosémantiques d'affixation. En revanche, les quatre types de matrices sont repérables dans la formation des nouveautés langagières françaises.

Pour les deux langues nous avons pris en compte un procédé de création que la grille typologique ne recensait pas en tant que telle, et que nous placerions à la frontière entre matrices internes et matrice externe. Il s'agit des variations linguistiques diatopiques au sein d'un même système de langue. Nous avons observé ainsi diverses influences dialectales en chinois et l'introduction de mots du français régional et non hexagonal en français.

Comme nous l'avons déjà laissé entendre, la proportion des matrices internes n'est pas du tout la même. Les créations phraséologiques et les changements du sens qui vont de pair avec le développement de connotations des expressions sont les plus nombreuses en chinois tandis que les créations par construction dans la matrice morphosémantique sont très clairement dominantes en français, en raison de la vitalité du mécanisme de dérivation lexicale par affixation. Nous savons que le français appartient aux langues flexionnelles alors que le chinois fait partie des langues isolantes.

Du côté chinois, les nouveautés langagières nous semblent principalement liées aux divers phénomènes sociaux. C'est-à-dire qu'une situation d'énonciation donnée est imputable à l'émergence des créations. En outre, elles servent souvent à exprimer un point de vue, un sentiment. Nous pouvons aussi y constater une présence d'ironie ou de comique qui vise une certaine connivence. Les expressions longues et les structures spécifiques (mixtes et semi-figées) sont les plus remarquables dans les créations chinoises. Par exemple, 世界那么大, 我想去看看 (shì jiè nà me dà, wǒ xiǎng qù kàn kàn, *litt.* « le monde est grand et je voudrais le visiter »), constitue la seule raison exposée dans la lettre de démission d'une enseignante en Chine. L'expression est ensuite devenue populaire pour renvoyer au courage de poursuivre une vie comme on veut mais pas pour exprimer une volonté de voyager et explorer le monde. Pour donner un autre exemple 定个小目标 (dìng gè xiǎo mùbiāo, *litt.* « établir un petit objectif ») ne signifie plus un objectif facile à atteindre mais indique ironiquement que c'est une tâche impossible après avoir été employé par un homme d'affaire très connu en Chine, d'après qui gagner 100 millions est très facile à réaliser pour un jeune diplômé. Sans connaître ce contexte,

on serait incapable de saisir le vrai sens de l'expression.

Du côté français, les mots simples et composés sont dominants. La nécessité de nommer de nouvelles réalités nous semble constituer une cause principale des créations. Nous pouvons ainsi constater que le lexique vient de domaines divers (scientifique, culturel, informatique, médical, écologique, etc.) et que les influences étrangères, notamment celle de l'anglais, jouent un rôle important y compris dans les procédés des matrices internes. Cela est manifesté principalement par les calques. Les procédés néologiques ne peuvent pas être séparés des éléments contextuels. Le repérage et l'interprétation des nouveautés langagières mobilisent en même temps le recours aux aspects énonciatifs. Nous avons ainsi constaté dans le corpus français certaines situations spécifiques reliées aux procédés de création linguistiques. Par exemple, le terme *Indignés* (matrice SS) n'a pas le sens de « personnes en proie à un sentiment d'indignation » mais désigne un mouvement de contestation. Pareillement, « faucheurs volontaires » ne signifie pas que les travailleurs agricoles sont d'accord pour faucher ou sont déterminés mais désigne un mouvement de désobéissance civile consistant à détruire les plantations OGM parce que nocives pour la santé. Donc, sans référer au contexte, des difficultés de compréhension peuvent survenir.

En résumé notre hypothèse selon laquelle les grandes tendances de création sont constatables dans les deux langues n'est que partiellement confirmée puisque la présence même des matrices et la répartition des procédés varient assez fortement. En tous les cas pour prendre en considération les spécificités de création dans chaque langue, ce ne sont pas les seules propriétés linguistiques qui interviennent et le recours aux éléments extralinguistiques est nécessaire.

## **Conclusion**

En partant des mêmes principes de nouveauté, de popularité et de fréquence, nous avons procédé à deux démarches différentes pour établir notre premier type corpus, le corpus de listes dans les deux langues. Les sélections issues des activités chinoises et les nouvelles entrées citées dans des articles de presse français nous ont permis de travailler sur un échantillon pour mener une analyse linguistique comparée. Cette analyse nous permet de tirer quelques premières conclusions partielles sur les procédés de création néologique : les proportions obtenues après avoir identifié les procédés de création sur les grands niveaux des matrices sont révélatrices

d'une répartition différente dans les deux langues. Les emprunts issus des langues étrangères sont bien attestés en parallèle des créations internes mais la comparaison des nouveautés langagières chinoises et françaises a mis en évidence des disparités dans leur classement et leur nombre. Les créations phraséologiques sont plus importantes en chinois tandis que les nouveautés langagières créées par affixation et composition sont les plus nombreuses en français. Enfin la typologie permet de mettre en évidence que de nombreuses nouveautés langagières dans une langue comme dans l'autre ne peuvent figurer dans une seule case mais font intervenir plusieurs procédés pour leur formation.

## CHAPITRE VI

### Constitution d'un corpus de tweets en vue d'une analyse discursive : approche théorique et méthodologique

#### Introduction

Recueillir et analyser des corpus en contexte numérique exige une réflexion adaptée que plusieurs chercheurs en sciences du langage ont menée. Nous évoquerons les travaux de trois auteurs et auteures qui se situent dans le champ de l'analyse du discours et envisagent dans des travaux très récents la prise en compte des discours du Web et plus spécifiquement des RSN. L'une articule la sémantique à l'analyse discursive concernant à la fois le lexique des réseaux et leur configuration technique. Une autre juge nécessaire d'élaborer de nouveaux concepts et de nouvelles catégories pour le discours numérique qui transforme et développe les sciences du langage. Le troisième fait une présentation des acquis et développements de plusieurs projets portant sur les tweets politiques afin de se donner les moyens de traiter des corpus de tweets par des méthodes mixtes (l'extraction pour une analyse outillées et la contextualisation pour une analyse discursive).

Nous exposerons ensuite nos choix méthodologiques pour la constitution des corpus en ajustant nos objectifs aux fonctionnalités et spécificités des plateformes, en prenant soin de mentionner les aspects communs et différents concernant *Sina Weibo* et *Twitter*.

Nous consacrerons ensuite deux développements distincts respectivement à la constitution du corpus chinois et à celle du corpus français en retraçant les démarches que nous avons adoptées, les questions que nous nous sommes posées et les solutions que nous avons retenues.

#### 6.1. Éléments théoriques et méthodologiques pour le traitement des corpus RSN

##### 6.1.1. La créativité lexicale et *Twitter* selon Reboul- Touré

Chercheuse en analyse du discours, Sandrine Reboul-Touré a aussi pour domaine de spécialité

la sémantique lexicale et discursive. Elle a notamment travaillé sur la vulgarisation scientifique et ce qu'elle appelle les discours de l'Internet. Dans sa recherche sur la créativité lexicale de *Twitter* (Reboul-Touré 2016), elle a étudié la sphère néologique de *tweet* et les variations linguistiques autour du mot pour confirmer l'influence de l'emprunt à l'anglais et des néologismes officiels sur le vocabulaire d'Internet. D'un côté, elle montre que le terme anglais *tweet* est à l'origine de nombreux nouveaux termes de l'Internet dans des discours courants en français comme par exemple *retweeter* ou *tweeteur*. De l'autre côté, une arrivée massive des mots anglais pourrait susciter la création d'équivalents français (les néologismes officiels) proposée par des Commissions de terminologie. Elle s'est également intéressée à l'écriture spécifique du tweet qui est contrainte par la forme et a montré la créativité lexicale liée à la fonctionnalité « hashtag » de *Twitter*. À titre d'exemple, elle mentionne la création des nouveaux sigles dans le contenu hashtagué pour synthétiser les informations. Nous pouvons constater que cette recherche ne se fonde pas sur un corpus de tweets et n'aborde pas la méthodologie de constitution de tels corpus. Il s'agit plutôt d'une étude de cas concernant le mot *tweet* et l'environnement de production des tweets dans une perspective de sémantique lexicale. Sa réflexion sur le statut de néologisme et/ou d'emprunt du mot *tweet* rejoint les considérations du chapitre précédent sur les matrices néologiques qui nous ont amené à penser qu'il est nécessaire d'introduire les procédés de création des nouveautés langagières en vue d'une analyse morphosémantique.

Pourtant, il faut souligner que son intérêt pour la créativité lexicale ne se situe seulement au plan linguistique de la néologie : dans sa recherche, elle prend en considération à la fois la dimension discursive et la dimension technologique. Le mot qu'elle a choisi d'étudier est très spécifique car il concerne en même temps une nouvelle création linguistique et un nouveau moyen de communication. Elle s'interroge non seulement sur la formation des créations lexicales d'un point de vue lexicologique mais aussi sur la relation potentielle entre les nouvelles modalités d'expression et les productions lexicales néologiques. C'est dans ce cadre théorique combinant la néologie et le nouveau dispositif de communication que nous nous situons.

Pour étudier les néologismes dans leur cadre discursif, elle a choisi de repérer les premières occurrences du mot *tweet* dans une base de donnée de presse *Factiva*. Elle a ainsi montré que ces premières occurrences sont accompagnées de marques typographiques, ce qui souligne le caractère neuf du mot. Quand le mot est largement employé ultérieurement, ces marques ne

sont plus présentes. Suivant la même ligne de réflexion, dans l'objectif d'observer la circulation des nouveautés langagières sur les microblogs et leur évolution de sens, nous nous demandons si la situation de circulation des nouveautés langagières sur les microblogs diffère des occurrences du mot *tweet* dans les discours de presse et s'il existe des indices de mention lors de leurs premières circulations.

L'étude de Reboul-Touré sur l'environnement technologique de *Twitter* et son influence sur l'écriture des tweets a retenu notre attention. Pour l'auteure, les créations lexicales sur *Twitter* sont plutôt attachées à la fonctionnalité « hashtag ». Le hashtag est composé du symbole # et d'un contenu souvent synthétique et thématique. En plus, cette partie a une couleur différente de celle des autres parties d'un tweet. Elle est en même temps cliquable avec le lien hypertexte intégré qui renvoie à l'ensemble des messages regroupés sous le même thème. Reboul-Touré montre que la créativité au niveau du hashtag est souvent présentée sous forme de sigle puisque le message du *Twitter* est contraint par la limite de 140 caractères. Par exemple : #CSOJ est utilisé à la place de « *Ce soir (ou) jamais !* » ; #CSOJ pour « *Danse avec les stars* ». Cette découverte peut confirmer dans une certaine mesure son hypothèse de l'influence de nouvelles modalités d'expression et de la technologie sur la production lexicale.

Dans notre recherche qui étudie la circulation des nouveautés langagières, le choix des microblogs en Chine et en France comme lieu d'observation est lié directement à l'importance que nous accordons à la limite initiale de 140 caractères, et à l'influence de cette contrainte de brièveté pour favoriser une communication rapide. Il nous paraît nécessaire de prendre en considération les éléments extralinguistiques, notamment socio-techniques dans l'analyse et nous sommes inspirée par l'examen particulier du hashtag comme nouveau signe de l'environnement technologique susceptible d'engendrer à son tour des nouvelles créations.

### **6.1.2. La contextualisation technorelationnelle de Paveau**

Donc les démarches de Reboul-Touré qui investigate la relation entre les nouvelles technologies et les procédés lexicaux rejoignent dans une certaine mesure la notion de « technologie discursive »<sup>200</sup> de Paveau dans le cadre de son analyse du discours des réseaux sociaux numériques. Comme nous l'avons vu à la fin du chapitre III, Marie-Anne Paveau élabore une

---

<sup>200</sup> « L'ensemble des processus de mise en discours de la langue dans un environnement technologique ». (Paveau 2017:335)

théorie du discours « postlinguistique » qui articule les textes, les discours et les interactions et s'élargit aux environnements. Elle a récemment publié un dictionnaire qui explore l'analyse du discours numérique comme une discipline à part entière des sciences du langage sans exclure le dialogue interdisciplinaire (Paveau 2017).

D'après elle, une technologie discursive est « un dispositif au sein duquel la production langagière et discursive est intrinsèquement liée à des outils technologiques en ligne ou hors ligne ; la technique n'est pas un simple « support » mais bien un composant structurel des discours » (Paveau 2013:3). Cette notion invite à traiter non seulement l'énoncé comme objet d'analyse, mais aussi l'ensemble du système dans lequel il s'inscrit. Cela implique la nécessité d'appliquer pour l'analyse du discours numérique natif, une approche écologique qui prend en compte le contexte social, technique, corporel dans lequel s'élabore le discours, autrement dit, une approche contextuelle.

Cette approche écologique s'oppose à l'approche traditionnelle dite logocentrée (centrée sur le verbal) qui prend pour objet les seuls éléments verbaux dans le discours numérique natif. À travers cette approche logocentrée, les études existantes en français traitent les énoncés des RSN en les séparant de leurs contextes de production d'après Paveau. Les chercheurs collectent les contenus verbaux et les présentent sous forme de listes ou d'énoncés individuels pour ensuite procéder à une analyse hors ligne. Paveau pense que cette manière de constituer des corpus peut engendrer des problèmes concernant la compréhension du contenu, et qu'il s'agit plutôt, de la part de ces chercheurs, de prendre les RSN pour trouver du corpus et non *comme* corpus comme elle vise à le faire. Parmi les études sur *Twitter*, elle considère l'approches de Longhi que nous allons aborder *infra* comme logocentrée, même s'il fournit les contextes des contenus verbaux extraits sur un site dédié. Selon elle, les énoncés sont produits et reçus dans un environnement unique et relationnel, et en conséquence ne peuvent pas être traités de manière isolée. En plus, sur chaque compte de RSN, l'utilisateur ne voit pas le même énoncé de la même façon et en même temps. La conception d'un énoncé dépend de la manière d'explorer qu'adopte l'utilisateur et des relations construites autour d'un compte. Dans cette perspective, elle préconise la prise en compte de la capture d'écran dans la méthodologie de constitution du corpus pour l'analyse du discours numérique. Dans ses travaux portant sur *Twitter* et sur des tweets, nous pouvons constater l'application régulière de ce choix.

Nous nous demandons ce qu'on peut voir exactement sur une capture d'écran à part les éléments

linguistiques et nous prenons un exemples issu des travaux de Paveau (Figure 27). Nous pouvons constater que les captures d'écran concernent des niveaux différents : une capture d'un tweet présent dans la TL<sup>201</sup> d'un compte (en haut à gauche de la Figure 27) ou l'affichage d'un tweet dans son détail (avec les interactions autour du tweet) sur la page d'accueil d'un compte *Twitter* (à droite de la Figure 27), ou encore une capture qui montre la forme d'un tweet prêt à être inséré (en bas à gauche de la Figure 27) :



Figure 2 : le tweet dans la TL (2013)



Figure 4 : le tweet par la fonction « insérer le tweet »



Figure 3 : le tweet avec la fonction « détails »

### Figure 27 : Différents affichages de tweets en captures d'écran selon Paveau (2013:7-9)

Ces différents éléments montrent dans un premier temps qu'il est difficile de déterminer la forme exacte d'un tweet quand il existe des moyens d'affichage différents. Dans un deuxième temps, nous pouvons constater la présence de technomots comme « # » (hashtag) ou « @ »<sup>202</sup> ou des hyperliens dans le corps du message d'un tweet et divers éléments périphériques. Regardons en détail la capture d'écran d'un tweet dans la TL. Il peut être composé des éléments suivants : avatar, nom, pseudo de l'abonné, date du tweet, texte du tweet... liste des opérations possibles signalées par des mots-consignes, assortis de leurs icônes : *répondre*, *retwitter*... (Paveau 2013:7). Il faut préciser que le texte du tweet peut aussi être étudié comme

<sup>201</sup> L'abréviation de *Timeline* où l'ensemble des tweets postés par les abonnements d'un compte sont affichés par défaut dans un ordre chronologique inversé.

<sup>202</sup> C'est le technomot qui joue la fonction « mention ». Il est précédé d'un nom de compte tiers pour créer un lien vers ce compte.



texte purement linguistique ou texte intégrant des technomots ou hyperliens comme mentionné *supra*. Pour une capture d'écran qui dépasse la sphère du corps du message d'un tweet, nous pouvons aussi trouver les interactions ou les relations construites autour d'un compte.

Selon Paveau, *Twitter* est un écosystème qui présente des traits technodiscursifs particuliers permettant l'installation d'un mode relationnel spécifique. Mais en observant les captures d'écran, nous pouvons remarquer qu'il s'agit d'un choix subjectif qui peut être justifié par l'explicitation de ce qui est recherché. Cela peut soulever la question de la représentativité lors de la constitution du corpus et entraîner des difficultés à généraliser les conclusions (Longhi, 2020:6). Nous avons aussi remarqué qu'une capture d'écran a sa limite d'affichage, c'est-à-dire que le nombre de tweets contenus dans une capture d'écran est limité ou, en d'autres termes, il est impossible d'afficher sous cette forme toutes les interactions d'un tweet s'il est retweeté à une grande échelle. En plus, la capture d'écran est un support non éditable qui est difficile à exploiter en grande quantité sans extraire les éléments. De ce fait, la capture d'écran n'est peut-être pas convenable pour une analyse quantitative. Paveau a mentionné la capture d'écran comme une démarche facile à entreprendre mais elle ne l'a pas exploitée en détail. Nous pouvons tirer de cela la conclusion que, dans la mesure où les objectifs de recherche varient, appliquer une méthodologie universelle n'est pas plus adéquat pour les études de l'analyse du discours numérique que dans d'autres domaines.

Nous sommes convaincue que la formation des nouveautés langagières n'est pas un acte purement linguistique. Elle a un lien étroit avec les pratiques des inventeurs/usagers et le contexte où ces pratiques surgissent. D'un point de vue écologique (Paveau 2013) auquel nous adhérons, l'analyse de la formation et la circulation de ces nouveautés langagières doit prendre en considération un contexte plus large, autrement dit, le contexte relationnel et le contexte technologique.

### **6.1.3. La méthodologie contextuelle et outillée de Longhi appliquée aux tweets politiques**

Paveau avance que toute extraction d'éléments verbaux des tweets relève d'une approche logocentrée. Mais, nous avons remarqué que dans les cas d'extraction mentionnés par Paveau pour illustrer cette approche non contextuelle, les chercheurs n'ont pas seulement pris en compte les parties linguistiques. Par exemple, dans la recherche de Yanoshevsky (2010:4) consacrée aux échanges politiques sur *Facebook*, elle a transcrit aussi le nom du compte

(Nicolas Sarkozy), la date et l'heure de publication (Figure 28), les technomots comme « \* » et les opérations comme *report*. Pour ce qui concerne *Twitter*, les technomots et les hyperliens n'ont pas été exclus du corpus de Longhi dans l'article qu'il consacre à analyser les tweets politiques comme un genre original de discours politiques (Longhi 2013). Il a aussi mis à disposition du lecteur dans cet article, pour chaque tweet transcrit (Figure 29), un hyperlien permettant de renvoyer au contexte de chaque tweet (sous forme de capture d'écran) et ces contextes regroupés sur un site dédié. Cet outil de stockage des tweets et de leurs contextes a été beaucoup développé dans le cadre de plusieurs recherches successives<sup>203</sup>, devenant la version actuelle #Idéo2017<sup>204</sup>. Il s'agit d'une base de données regroupant tous les tweets des comptes des candidats qui ont participé à l'élection présidentielle française de 2017. Les tweets sont extraits intégralement, c'est-à-dire dans leur matérialité discursive et formelle. Les technomots comme « # » (hashtag) ou « @ » et les hyperliens sont tous pris en compte. En plus, sur la plateforme #Idéo2017, il est possible de naviguer vers tous les tweets extraits et de retourner dans l'environnement natif des tweets en cliquant sur le lien fourni. Cette interface a aussi intégré des outils de statistique textuelle et de visualisation de données pour illustrer les analyses par thème. Si son étude publiée en 2013 sur la catégorisation du tweet politique s'accompagne d'un recueil d'exemples, avec l'outil #Idéo2017, les recherches centrées sur les lexiques politiques visent à investiguer les mises en discours des candidats à l'élection de manière exhaustive (Longhi 2020) et s'appuient sur une analyse quantitative outillée<sup>205</sup> informatiquement.

---

<sup>203</sup> 1. Le projet « Humanité numériques et data journalisme : le cas du lexique politique » : soutenu par la Fondation UCP, visait à impulser une recherche sur le thème des humanités numériques, et créer des synergies entre plusieurs acteurs de l'université de Cergy-Pontoise. ;

2. Le projet CoMeRe avait pour objectif, à l'horizon 2014-2015, de créer un noyau de corpus de communication médiée par les réseaux (Computer Mediated Communication – CMC) en français. (Longhi 2020:4)

<sup>204</sup> Le projet #Idéo2017, financé par la Fondation UCP, associe des chercheurs du laboratoire AGORA et du laboratoire ETIS (ENSEA / UCP / CNRS UMR 8051). L'objectif du projet est de créer un outil d'analyse des tweets politiques lors de campagnes politiques. Le site de la plateforme : <http://patrimorph.ensea.fr/RST/>.

<sup>205</sup> Le logiciel *Iramuteq* pour l'analyse de similitudes ; la méthode *Alceste* pour l'analyse thématique.

[Nicolas Sarkozy] remercie tous ses supporters sur Facebook pour leurs messages de soutien.  
Tuesday, July 28, 2009 at 10 :08 pm  
Je vais très très bien après le petit incident de dimanche, qui m'incite à prendre du repos cet été.  
Merci pour vos encouragements et témoignages, qui me touchent beaucoup.  
Parmi tous les messages que j'ai pu recevoir – de tous mes collègues chefs d'Etat et de gouvernement étrangers notamment – les vôtres me tiennent particulièrement à cœur.  
Votre sympathie et votre spontanéité sur Facebook font vivre un lien précieux entre nous  
Je vous souhaite à tous un très bon été.  
Nicolas Sarkozy

**Figure 28 : Capture d'écran du message de Sarkozy présenté par Yanoshevsky (2010:4)**

D'autres tweets jouent sur l'analogie avec le slogan socialiste *Le changement, c'est maintenant*, permettant à M. Le Pen à la fois d'ironiser sur les promesses du candidat Hollande, et sur les actions en cours :

[http://twitter.com/MLP\\_officiel/status/228489954899152898](http://twitter.com/MLP_officiel/status/228489954899152898)

Avec #Valls, les distributeurs automatiques de nationalité, c'est maintenant ! :<http://bit.ly/Pv0LMb>

ou

[http://twitter.com/MLP\\_officiel/status/220435029308866560](http://twitter.com/MLP_officiel/status/220435029308866560)

Discours de politique générale : les poncifs c'est maintenant : <http://bit.ly/PaOTSe>

**Figure 29 : Capture d'écran de la présentation des tweets de M. Le Pen par Longhi (2013:3)**

Du point de vue des sciences du langage, comme mentionné dans ses travaux récents, Longhi se rapproche du positionnement de Paveau qui souligne l'importance de l'environnement pour l'analyse du discours numérique (Longhi 2020). Cela se manifeste par ses efforts pour recueillir l'intégralité de chaque tweet et pour fournir la possibilité de retrouver le contexte de production. Prenant en considération la méthodologie de Yanoshevsky qui opte pour une extraction des données et celle de Longhi qui procède à une démarche mixte combinant l'extraction et la contextualisation, même si les méthodes d'extraction et les outils de traitement restent classiques, il ne s'agit pas à proprement parler d'une approche décontextualisée.

#### **6.1.4. Options et questionnements**

Nous allons à présent nous situer par rapport à ces recherches en gardant à l'esprit l'objectif de la constitution de nos corpus de tweets, qui est d'observer et d'analyser l'évolution de sens des nouveautés langagières que nous avons sélectionnées dans nos corpus de liste. Pour constituer ces corpus nous devons répondre à des questions et opérer des choix. Combien de nouveautés langagières suivre dans leur circulation sur les microblogs ? Lesquelles ? Combien de tweets pour chacune des nouveautés langagières ? Quelle méthodologie utiliser pour collecter ces tweets et sous quelle forme ?

Dans le cadre d'une analyse qualitative d'échantillons, nous allons appliquer une méthode mixte combinant l'extraction du tweet et la capture d'écran. L'extraction a pour but de faciliter les analyses et la capture d'écran de renvoyer à l'environnement des tweets.

Nous retenons les enseignements de Paveau qui préconise d'adopter un point de vue écologique, puisque les éléments environnementaux sont inséparables des éléments linguistiques, concernant particulièrement le discours numérique, et que l'extraction des contenus verbaux ne permet pas d'accéder à un environnement qui fait sens. Toutefois nous ne souscrivons pas complètement à sa proposition d'appliquer systématiquement la capture d'écran pour éviter une approche logocentrée. Nous avons aussi remarqué que la capture d'écran ne convient pas pour une grande quantité de tweets et qu'une démarche d'extraction automatisée est nécessaire pour une analyse quantitative – qui n'est pas notre orientation.

Bien que nous soyons très attirée aussi par la méthode de Longhi qui pratique l'extraction sans sacrifier les éléments spécifiques<sup>206</sup>, nous ne disposons pas de moyens pour stocker les tweets en ligne comme lui avec la plateforme #Idéo2017 (Longhi 2020) et ne participons pas à un projet collectif.

Nous essayons donc appliquer une méthode mixte combinant l'extraction du tweet (pour faciliter les analyses générales) et la capture d'écran (pour renvoyer à l'environnement du tweet lors des études de cas). Afin de collecter des billets de tweet contenant les nouveautés langagières dont nous expliquerons *infra* le choix, nous nous servons des fonctionnalités des moteurs de recherche de *Sina Weibo* et *Twitter*.

### ***Fonctionnalités des moteurs de recherche***

Pour effectuer les recherches, il faut préalablement disposer d'un compte d'utilisateur. Précisons que toutes nos observations sur *Sina Weibo* et *Twitter* sont réalisées à travers sa version page Web qui offre des fonctionnalités plus complètes que les version compactes sur smartphone et tablette. Par exemple, la version page Web peut afficher tous les résultats page par page permettant d'avoir une vision globale sur la recherche et laissant la possibilité de feuilleter selon besoin, alors que les versions compactes ne donnent pas accès aux paramètres

---

<sup>206</sup> L'extraction intégrale suppose de transcrire tous les éléments extralinguistiques et pas seulement les messages verbaux.

de filtre pour réaliser une recherche précise.

On peut effectuer une recherche générale par mots clés en entrant chaque choix dans le moteur et les résultats sont affichés dans un ordre chronologiquement inversé ou selon la popularité. Cette dernière nouvelle méthode d’affichage sur la page d’accueil est appliquée par *Sina Weibo* depuis 2015 et par *Twitter* depuis 2016. La popularité pourrait dépendre de la date de publication et de l’engagement, y compris le nombre de lectures, de retweets, de « j’aime », l’ajout ou non de photos, vidéos et hashtags. Les deux RSN disposent de leurs propres critères pour le calcul concret de la popularité. On peut aussi faire une recherche avancée en appliquant des filtres pour obtenir un résultat rapproché. Sur *Sina Weibo*, on propose des filtres comme les types de tweet (les tweets initiatifs d’un compte<sup>207</sup>, les tweet classés populaires, les tweets publiés par des comptes de médias ou des comptes certifiés, ou les tweets des abonnés), les contenus (contenant ou non des photos, vidéos, liens, musiques), la date et l’endroit de publication. Sur *Twitter*, nous pouvons trouver les filtres suivants : mots (différents critères de combinaison, hashtag, etc.), langue, comptes, présence ou non des réponses ou liens, engagement (nombre de réponses, de « j’aime », de retweets), et date de publication.

Les productions sur un RSN dépendent à la fois des fonctionnalités de chaque plateforme et des pratiques des usagers. Le moteur de recherche inhérent du microblog peut nous donner accès aux archives publiées mais la quantité de résultats est contrainte. L’affichage maximum de 50 pages de résultats sur *Sina Weibo*, (quels que soient les critères de recherche), en est un exemple. Or il se peut que la quantité des tweets contenant un mot clé donné dépasse 50 pages. Sur *Twitter*, un autre problème se pose : nous ne pouvons que faire défiler la page vers le bas constamment et attendre le chargement des tweets s’il y en a. L’inconvénient est que nous n’arrivons pas à connaître la fin de l’affichage qui varie par ailleurs d’une nouveauté langagière à l’autre en fonction du nombre de tweets qui les contiennent. Dans ce cas-là, il nous est nécessaire d’appliquer des filtres pour effectuer des recherches avancées. Mentionnons aussi que la fonctionnalité offrant les choix de publication privée ou ouverte aux usagers peut modifier le résultat final. Sur *Twitter* et *Sina Weibo*, on distingue deux sortes de tweets : i) tweets publics (paramètre par défaut) que tous les utilisateurs peuvent voir, qu’ils disposent ou non d’un compte *Twitter* ; ii) tweets protégés que seuls les abonnés à *Twitter* peuvent voir. Seuls les

---

<sup>207</sup> La plateforme distingue deux catégories de tweets : i) publication initiale du compte, ii) retweet dont le contenu émane des autres comptes.

tweets publics sont ouverts à tout le monde et en conséquence les tweets obtenus à travers une recherche au sein du microblog sont des tweets publics. Le contenu publié sur *Sina Weibo* et *Twitter* est également dynamique : cela est non seulement manifesté par les publications innombrables apparues à tout moment, mais aussi par des suppressions imprévisibles. Donc, il est difficile d'avoir un résultat stable.

### ***Fonctionnalités relationnelles***

En tant que réseaux sociaux numériques, les plateformes de microblog favorisent la transmission des messages et donc la circulation de nouveautés langagières notamment par les fonctionnalités relationnelles telles que les retweets avec ou sans commentaire, les réponses. Cela implique que les termes peuvent être utilisés dans le contenu de commentaires et réponses. Il ne faut pas oublier les attachements d'un tweet, images et vidéos, où on peut aussi repérer la présence des nouveautés langagières. Il n'est pas non plus surprenant de les retrouver dans les pseudonymes des comptes. Les éléments constitutifs du microblog offrent différentes possibilités d'utilisation d'un terme. Mais les moyens de repérage de ces utilisations est principalement conditionné par la mise en relation concrète de chaque microblog, autrement dit les fonctionnalités à travers lesquelles se manifestent les interactions.

À la différence de *Twitter*, pour chercher un mot clé, les pages de résultats de *Sina Weibo* affichent par défaut prioritairement le contenu textuel d'un tweet sans prendre en considération les occurrences d'utilisation du mot clé dans les réponses. Rappelons que ces deux microblogs présentent des nuances concernant leur affichage. À partir d'une fenêtre de consultation sur *Sina Weibo*, nous pouvons directement visualiser le tweet publié et les interactions par retweet et les commentaires qui les accompagnent s'il y en a, mais à cause de la contrainte de 140 signes imposée, les personnes qui retweetent sont susceptibles de supprimer les commentaires des autres s'il ne s'agit pas d'un retweet direct à partir de la publications d'origine. Il faut cliquer sur le bouton situé en bas d'un tweet pour connaître les réponses.

Sur *Twitter*, ce n'est pas le même cas. Les réponses sont affichées de la même manière que les retweets. Voici deux exemples d'affichage du terme *attachiant*, néologisme morphosémantique de construction par composition du type mot-valise (Figure 30 et 31). Nous avons retrouvé le mot *attachiant* dans une réponse (« En réponse à ») et un retweet par la même recherche.



Figure 30 : *attachiant* se trouve dans une réponse



Figure 31 : *attachiant* se trouve dans un retweet avec commentaire

*Sina Weibo* ne propose pas non plus les filtres de recherche de *Twitter* qui permettent d'afficher sur le même plan les réponses et tweets d'origine ou de présenter uniquement les réponses.

Ces diverses caractéristiques nous amènent à nous interroger sur notre méthodologie de constitution du corpus. Nous nous demandons s'il est convenable de prendre en considération seulement le corps message d'un tweet (sans faire de distinction entre un tweet et un retweet) comme corpus d'analyse. Si nous prenons aussi les réponses contenant les nouveautés langagières, comment pouvons-nous les chercher et les distinguer sur *Sina Weibo* ? Il faut

souligner que les réponses sur un microblog ne concernent pas seulement la réponse à un tweet mais aussi une réponse à une autre réponse. Certes, la compréhension d'une réponse doit être associée au contenu auquel elle réagit. Ces interactions dans les réponses d'un tweet doivent-elles être considérées comme des éléments à part entière d'un corpus écologique? Si oui, faut-il faire une sélection ou non ? Comment pouvons-nous recueillir un corpus de réponses si la conversation est abondante ?

Aux questions concernant le choix des éléments, s'ajoutent aussi celles qui concernent le choix d'affichage. Il est impossible pour les chercheurs d'accéder aux mêmes interfaces d'interaction que les usagers mais plusieurs options sont possibles. En reprenant l'exemple du mot-valise *attachiant*, la Figure 32 est un résultat de recherche par le moteur de *Twitter*. Le contenu de la réponse et l'identité de l'interlocuteur constituent les éléments essentiels.



Figure 32 : Réponse contenant le mot « attachiant » présentée dans la page de résultat



Figure 33 : Conversation liée à une réponse





Figure 34 : Affichage de la réponse dans l'intégralité de l'interaction

En cliquant sur cette réponse, nous sommes renvoyés à une autre page (Figure 33) indépendante qui nous montre une conversation englobant un tweet d'origine et une autre réponse. Nous

pouvons comprendre que « crazyyy » est la réponse au tweet initial « *c'est quoi vos types de personnalité* ». Le terme « attachiant » est employé par l'auteur du tweet initial pour interagir avec la réponse « crazyyy ». Si nous cliquons sur le tweet initial, nous pouvons accéder à la page (Figure 34) où nous pouvons connaître l'intégralité des interactions autour de ce tweet. Nous pouvons constater que la réponse de la Figure 34 se trouve à la fin de toutes les interactions et qu'il s'agit de la réponse la plus récente. Il est intéressant de se demander si la réponse « crazyyy » est influencée par les autres réponses dans le fil de conversation. Par cet exemple, nous voulons montrer que les différents affichages peuvent fournir des informations variées aussi bien pour les réponses que pour les tweets et retweets associés aux interactions.

Face à la circulation favorisée par de multiples pratiques relationnelles et en prenant en considération les fonctionnalités du microblog et les différences entre *Sina Weibo* et *Twitter*, nous décidons de collecter les tweets (y compris les retweets s'il y en a) dont le corps du message contient les termes que nous choisirons d'observer et d'éliminer les occurrences d'utilisation des termes dans la rubrique « réponse » et les noms et pseudonymes des comptes. Pour trouver un autre principe de limitation et en raison de son importance sociale et sémantique, nous considérons comme centrale la fonctionnalité « hashtag », ce qui nous permettra de limiter le corpus aux tweets contenant les nouveautés marquées par le hashtag.

### ***La fonctionnalité hashtag***

Selon le Dictionnaire d'analyse du discours numérique:

Le hashtag est un segment langagier précédé du signe #, utilisé principalement sur le réseau de microblogging *Twitter*. Cette association en fait un tag cliquable, inséré manuellement dans un tweet et permettant d'accéder à un fil. Le hashtag est une convention mise en place par les usagers du réseau (Kwak et al. 2010). C'est une forme langagière dont la fonction est essentiellement sociale, permettant l'affiliation diffuse (ambient affiliation) des usagers, la technoconversationalité et l'investigabilité (searchability) du discours (Paveau 2013<sup>208</sup>).

La fonctionnalité du hashtag est dans un premier temps de regrouper tous les tweets contenant le même hashtag. Nous pouvons accéder à ce regroupement sur des pages d'affichage dédiées qui renvoient à la rubrique 话题 (huà tí) sur *Sina Weibo* et à la rubrique « Tendances » sur *Twitter*. Par contre, le hashtag de *Sina Weibo* est présenté par deux signes #. Il faut insérer le

---

<sup>208</sup> Source : <https://technodiscours.hypotheses.org/488>, Dictionnaire d'Analyse du Discours Numérique (DADN) en ligne, consulté le 20 juin 2017, (Paveau 22 mai 2013).

contenu entre les signes # pour pouvoir partager le contenu avec la fonctionnalité. Sur *Twitter*, le contenu à partager est précédé d'un seul signe #. Cela rend également les tweets associés investigables en cliquant sur la rubrique ou par la recherche du hashtag dans le moteur. L'ajout du hashtag est une action volontaire de la part des usagers. Il leur permet de participer à une communauté discursive par la publication au niveau du tweet seul. C'est aussi une façon d'optimiser la visibilité d'un tweet pour pouvoir provoquer plus de possibilités d'interaction.

Nous avons choisi d'observer les termes marqués par hashtag pour les raisons suivantes : le regroupement des tweets sous un hashtag peut nous donner un contenu présélectionné. En même temps, la recherche par hashtag peut nous permettre de retrouver des tweets contenant les termes exacts tandis que dans les résultats d'une recherche générale nous pouvons trouver des tweets qui contiennent seulement certains caractères ou mots qui coïncident avec les composants d'un terme. Dans ce dernier cas, les termes que nous cherchons sont en réalité absents. Un autre apport du hashtag est qu'il nous permet d'observer le message verbal et le contexte techno-relationnel sur une seule page et sur une affiche simple (comme Figure 32). Nous n'avons pas besoin de nous reporter aux autres pages pour collecter les données visant à constituer le corpus (comme Figure 34). En plus, la page où les tweets sont réunis constitue une interface commune. C'est-à-dire qu'un producteur, un usager qui consulte ou un chercheur peuvent avoir la même fenêtre de consultation suivant les critères de recherche. Tout cela nous facilite la constitution du corpus et nous pouvons ainsi disposer d'exemples significatifs pour l'analyse de l'évolution du sens.

### ***Les catégorisations des tweets***

Sur la page dédiée à l'affichage des tweets contenant le hashtag, *Sina Weibo* et *Twitter* proposent tous deux des catégories (voir les captures d'écran ci-dessous, Figure 35 et 36) pour que les usagers puissent filtrer les résultats selon leurs besoins. Cela peut nous inspirer pour choisir notre corpus.



Figure 35 : Interface d’affichage des résultats de recherche par Hashtag sur *Sina Weibo*



Figure 36 : Interface d’affichage du résultat de recherche par Hashtag sur *Twitter*

Sur *Sina Weibo*, 5 rubriques sont disponibles. La première que nous traduisons comme « l’ensemble » regroupe en effet tous les résultats. Les deux suivantes sont les rubriques « récent » et « à la une » qui classent respectivement les tweets récemment publiés par rapport à la date de recherche et les tweets qui bénéficient de popularité selon les critères de *Sina Weibo*. Les deux rubriques « vidéo » et « photos » privilégient les tweets qui partagent ces types de contenus. *Twitter* propose la même quantité de rubriques. La comparaison montre que les différences entre les deux plateformes ne sont pas très importantes. *Sina Weibo* offre un espace affichant tous les résultats alors que cette option n’est pas présentée sur *Twitter*. Nous avons aussi remarqué que *Twitter* met à disposition ici la rubrique « Personne »<sup>209</sup> pour chercher les comptes des tweets où les mots de recherche apparaissent dans les pseudonymes ou la description du profil. Mais cette option semble-t-il n’a rien à voir avec le hashtag puisque qu’on trouve rarement le signe # dans les résultats surtout dans les pseudonymes des comptes. En fait, les catégorisations pour filtrer les résultats de recherche relatifs au hashtag sur les deux microblogs ne se différencient pas des recherches de base sans hashtag. Nous allons donc écarter la catégorisation « Personne » parce que cela ne nous permet pas de retrouver des tweets correspondants.

Pour essayer d’avoir les mêmes critères de sélection sur les deux plateformes, nous nous

<sup>209</sup> *Sina Weibo* a une rubrique sur la page de recherche générale mais pas sur les résultats de recherche des hashtags.

sommes ensuite intéressée aux catégories communes proposées. Ce sont les suivantes : tweets récents, tweets à la une, tweets avec vidéo ou photo. Le choix de tweets récents est conditionné par le moment de recherche qui est subjectif. Tant que ce moment n'est pas déterminé, ces tweets constituent également un corpus dynamique. L'affichage sur cette page des productions instantanées de tous les usagers de la plateforme nous pose également la question de la quantité qui n'est pas encore résolue dans nos premières démarches de recherche. En plus, dans l'objectif d'observer une éventuelle évolution de sens des nouveautés langagières, nous n'avons pas voulu limiter notre observation à une période unique. Donc, nous passons d'abord cette catégorisation.

La catégorie des tweets contenant vidéos et photos (y compris les Gif<sup>210</sup>, images animées) sont des contenus enrichissant les tweets par des composants pluri-sémiotiques. Vidéos et photos peuvent fournir des éléments complémentaires du message textuel ou constituer le contenu essentiel d'un tweet. Pourtant, ils ne sont pas les premières fonctionnalités proposées avec le message limité à 140 signes. Ils ne font pas non plus partie de la production native du microblog, c'est-à-dire que les vidéos ou photos ne sont pas directement élaborés sur la plateforme du microblog.

Notre constitution du corpus de tweets a pour objectif d'observer le rapport entre les termes choisis et leurs contextes dans un environnement techno-discursif. Pour cela, nous collectons notre corpus de tweets par capture d'écran afin de ne pas séparer les énoncés de leurs contextes de production. Ce faisant, nous allons donc non seulement prendre en considération le contenu verbal mais aussi l'ensemble des éléments constituant un tweet, comme par exemple : le pseudo de l'abonné, la date du tweet, les opérations relationnelles marquées par des mots-consignes, ainsi que les attachements comme des vidéos et photos. Mais le recueil et l'analyse des vidéos et des images Gif nécessitent l'application de méthodes spécifiques autres que la capture d'écran. En plus, l'extraction du contenu dynamique d'une vidéo peut aboutir à une collecte volumineuse. Sachant que nous nous intéressons principalement à l'analyse de la production native du microblog et non à l'analyse du contenu audiovisuel et des images qui sont produits ailleurs, les tweets contenant les vidéos et les images Gif ne feront pas partie de nos choix. Quant aux images statiques, nous n'avons pas besoin d'effectuer des démarches supplémentaires pour leur recueil et il sera intéressant dans ce cas de pouvoir observer la

---

<sup>210</sup> Graphics Interchange Format.

relation entre le corps du message et les images jointes. Par ailleurs, pour le respect du droit à l'image, nous n'allons pas collecter les tweets contenant des selfies ou des portraits<sup>211</sup>.

Les tweets dans la rubrique « à la une » sont ceux qui bénéficient de popularité. Cette popularité est calculée sur *Sina Weibo* par les composants d'un tweet (c'est-à-dire le message textuel, les technomots comme les hashtags ou les hyperliens, les images, les vidéos, etc.) et le taux d'engagement (soit le nombre d'interactions suscitées par un tweet). Plus la forme est diversifiée (présence d'images et de vidéos), plus un tweet est susceptible d'être considéré comme populaire. Il en va de même pour les tweets qui provoquent des interactions (retweeter, aimer, etc.). Pour le mode de calcul de la popularité sur *Twitter*, nous n'avons pas trouvé de références qui l'expliquent. Mais selon nos observations et après consultation de versions de *Twitter* dans différentes langues (chinois, français et anglais), nous supposons que la popularité d'un tweet sur *Twitter* est basée sur les mêmes critères. Nous avons constaté que les tweets populaires ne sont pas classés selon un ordre chronologiquement inversé. Il nous semble dès lors plus facile d'accéder à des tweets publiés à périodes différentes et cela lève la difficulté du choix d'une période pour nous. Le taux d'engagement indique que les utilisations des nouveautés langagières que nous voulons étudier ont été reconnues par une partie des usagers de plateforme, autrement dit ces tweets sont significatifs, ce qui répond à notre objectif d'analyse.

Nous décidons donc de sélectionner notre échantillon à partir des tweets populaires. N'oublions pas que la popularité est aussi liée aux éléments qui accompagnent les messages textuels, mais comme nous l'avons dit *supra*, nous ne prendrons pas en considération les tweets populaires contenant les vidéos pour des raisons méthodologiques.

### ***Quelle quantité de tweets choisir et lesquels ?***

Les contraintes technologiques, l'imprévisibilité de la production des tweets et la pratique personnelle des usagers de microblogs nous conduisent à entreprendre une analyse qualitative d'un échantillon significatif qui permette d'observer de façon approfondie l'évolution de sens des nouveautés langagières sur les deux plateformes. De ce fait, nous jugeons pertinent de collecter pour chaque terme 10 tweets à analyser. Ce choix d'une quantité de 10, très limitée au

---

<sup>211</sup> Les photos privées et non les photos publiées par la presse.

regard de la masse des tweets, convient cependant à la démarche exploratoire que nous adoptons et aux études de cas que nous envisageons de mener. Rappelons que nous ne cherchons pas à étudier exhaustivement les occurrences d'utilisation des termes.

Nos critères de sélection sont principalement basés sur l'exclusion de certains composants d'un tweet et la prise en compte de la spécificité technique des plateformes. Sur le plan formel du tweet, nous l'avons dit *supra*, nous éliminons les tweets contenant des vidéos ainsi que ceux qui comportent des photos représentant des personnes. Concernant le corps du message des tweets, on sait que la spécificité originale et originelle du dispositif du microblog est la contrainte de 140 signes par message. Néanmoins, *Sina Weibo* et *Twitter* ont tous deux supprimé cette limite en 2016 et 2017. Comme notre recherche sur les nouveautés langagières en circulation dans un contexte techno-relationnel du microblog considère l'influence de la brièveté d'expression exigée par le dispositif sur la pratique linguistique, nous allons conserver un critère de taille. C'est-à-dire que pour les tweets de *Sina Weibo* collectés, produits après 2016, nous allons toujours appliquer la limite de caractères et écarter les tweets longs.

En outre, il faut être attentif à une pratique spécifique. *Sina Weibo* offre depuis longtemps un outil permettant de transformer un contenu long (texte et/ou image) en une seule image et de la télécharger comme image d'attachement d'un tweet. Les outils tiers comme *OneShot*, *Snaplight* permettent une pratique équivalente sur *Twitter*. Il s'agit d'une alternative pour publier un contenu sans la contrainte de 140 signes tout en restant sur la plateforme (cela est différent des hyperliens qui renvoient vers des pages extérieures au microblog). Pour les productions qui relèvent de cette pratique, même si le corps du message ne dépasse pas la longueur requise, nous n'allons pas les prendre en considération.

Maintenant, les critères de l'établissement des corpus de tweets sont établis. Nous allons revenir au choix des nouveautés langagières à observer et aux opérations de sélection effective des tweets. Nous allons ainsi présenter ces démarches qui sont premièrement réalisées sur la plateforme *Sina Weibo* et nous abordons par la suite les procédures adaptées pour la collecte sur *Twitter*.

## 6.2. La constitution du corpus de tweets chinois en 4 étapes

### 6.2.1. Choix des nouveautés langagières pour collecter des tweets : l'étape des éléments communs

Pour établir notre corpus de liste de nouveautés langagières en chinois, nous avons repris les sélections du magazine 《咬文嚼字》 (yǎo wén jiáo zì) (dénommé corpus A *supra*) et de l'activité 《汉语盘点》 (hàn yǔ pán diǎn) (dénommé corpus B *supra*) qui représentent 90% des résultats du corpus de liste chinois alors que le troisième sous-corpus (les propositions du journal *Le Quotidien du Peuple*, que nous dénommons corpus C) se limite aux 25 sélections de l'année 2014. Notre corpus de liste chinois est donc constitué de 175 termes (voir chapitre V). L'entrée dans les listes de palmarès des institutions implique déjà une forte diffusion dans la société, quels que soient les supports. Nous supposons que sur le microblog, ces nouveautés langagières ont aussi connu une grande diffusion qui est manifestée par leurs emplois dans la publication des tweets, ce qui suppose des données volumineuses. Pour choisir au sein de l'ensemble des nouveautés langagières du corpus de listes celles qui sont les plus pertinentes à observer en discours dans des tweets, nous avons considéré qu'il était opportun de nous intéresser aux éléments communs aux sous-corpus A et B comme étant significatifs. C'est-à-dire que nous pensons qu'ils sont susceptibles de répondre aux critères de popularité, de haute fréquence d'utilisation et de respect des normes linguistiques. Rappelons que le corpus A issu des propositions du magazine est sélectionné à partir des votes de lecteurs et des avis des linguistes : le corpus B émanant de l'activité 《汉语盘点》 (hàn yǔ pán diǎn) est constitué à partir de la veille sur les occurrences d'utilisation et les conseils des experts. Nous n'avons pas pris en compte le corpus C dans cette démarche parce que ces 25 termes sont aussi présents dans les corpus A et B. Donc, nous avons éliminé ce cas de répétition.

Après avoir comparé les sous-corpus A et B, nous avons obtenu 21 résultats qui correspondent dans le Tableau 26 ci-dessous, aux cases surlignées en jaune. Nous pouvons trouver, de 2012 jusqu'à 2017, des nouveautés langagières en commun pour chaque année. C'est-à-dire que depuis le commencement de son activité, 《汉语盘点》 (hàn yǔ pán diǎn) a quelques résultats qui coïncident avec les sélections de 《咬文嚼字》 (yǎo wén jiáo zì) même si la quantité se diversifie annuellement. Les cases surlignées en bleu sont celles de termes communs mais qui ne sont pas sélectionnés pour les mêmes années dans les deux sous-corpus.



YAOWENJIAOZI									HANYUPANDIAN					
2009	2010	2011	2012	2013	2014	2015	2016	2017	2012	2013	2014	2015	2016	2017
不差钱	给力	亲	正能量	中国梦	顶层设计	获得感	洪荒之力	不忘初心	中国好声音	中国大妈	我也是醉了	重要的事情说三遍	洪荒之力	打call
猕猴桃	神马都是浮云	伤不起	元芳,你怎么看	光盘	新常态	互联网+	吃瓜群众	砥砺前行	元芳,你怎么看	高端大气上档次	有钱就是任性	世界那么大,我想去看看	友谊的小船	尬聊
低碳	围脖	Hold住	舌尖上	倒逼	打虎拍蝇	颜值	工匠精神	共享	高富帅,白富美	爸爸去哪儿	蛮拼的	你们城里人真会玩	定个小目标	你的良心不会痛吗
被就业	围观	我反正信了	躺着也中枪	逆袭	断崖式	宝宝	小目标	有温度	你幸福吗	小伙伴们都惊呆了	挖掘机技术哪家强	为国护盘	吃瓜群众	惊不惊喜,意不意外
裸	二代	坑爹	高帅富	微XX	你懂的	创客	一言不合就XX	流量	江南style	待我长发及腰	保证不打死你	明明可以靠脸吃饭,却偏偏要靠才华	葛优躺	皮皮虾,我们走
纠结	拼爹	卖萌	中国式	大V	断舍离	脑洞大开	友谊的小船,说翻就翻	可能XXX假XXX	躺着也中枪	喜大普奔	萌萌哒	我想静静	辣眼睛	扎心了,老铁
钓鱼	控	吐槽	压力山大	女汉子	失联	任性	供给侧	油腻	屌丝,逆袭	女汉子	时间都去哪儿了	吓死宝宝了	全是套路	还有这种操作
秒杀	帝	气场	赞	土豪	神器	剁手党	葛优躺		舌尖上的中国	土豪金	我读书少你别骗我	内心几乎是崩溃的	蓝瘦香菇	惹
蜗居	达人	悲催	最美	奇葩	高大上	网红	套路	惹	最炫民族风	摊上大事了	画面太美我不敢看	我妈是我妈	老司机	你有freestyle吗
蚁族	穿越	忐忑	接地气	点赞	萌萌哒	主要看气质	蓝瘦,香菇	打call	给跪了	涨姿势	且行且珍惜	主要看气质	厉害了我的哥	油腻

**Tableau 26 : Les nouveautés langagières en commun dans les sous-corpus A et B**

Nous observons tout d'abord que le surgissement de nouveautés langagières est régulier d'année en année. Pourtant, nous ne pouvons pas déterminer si leurs usages sont éphémères ou non. L'observation de leur circulation sur *Sina Weibo* nous aidera peut-être à le faire. Deuxièmement, nous avons repéré deux cas spécifiques (surlignés en bleu). Ils bénéficient dans une certaine mesure d'une popularité continue par rapport aux autres nouveautés langagières parce qu'ils sont sélectionnés successivement dans les deux listes de palmarès. Nous avons constaté aussi que certains termes sont relevés sous des formes différentes par les deux institutions, comme 高富帅 (gāo fù shuài) vs 高富帅, 白富美 (gāo fù shuài, bái fù měi), 高大上(gāo dà shàng) vs 高端大气上档次 (gāo duān dà qì shàng dàng cì)<sup>212</sup>. Il s'agit de formes raccourcies par rapport à leurs formes originales. Ces variations nous paraissent témoigner d'une évolution d'utilisation mais nous considérons que ce sont des équivalents et nous les incluons dans l'échantillon de mots en commun. Nous avons aussi remarqué que la quantité de résultats en commun pour chaque année est différente (par ex. : 2 résultats pour l'année 2015, 7 pour l'année 2016), sans pouvoir avancer que cela dépend des critères de sélection appliqués ou d'autres facteurs.

### 6.2.2. Une étape intermédiaire de pré-analyse contextualisée

Nous allons maintenant examiner la liste des 21 éléments communs. Ce sera une étape

<sup>212</sup> Voir les traductions dans le Tableau n°27 *infra* des 21 éléments communs : ce sont les termes n°4 et n°8.

intermédiaire qui devra déboucher sur le choix d'un nombre plus limité d'entrées pour la sélection proprement dite des tweets. Nous avons travaillé dans le chapitre V dans une perspective essentiellement linguistique. Nous adoptons à présent une optique d'analyse discursive et communicationnelle. Cela nous conduit nécessairement à chercher à connaître préalablement le contexte des créations et les motivations sociales et culturelles de celles-ci. Pour éclairer cet aspect, nous avons consulté les travaux existants : nous allons reprendre, discuter et compléter notamment deux travaux chinois dont les corpus sont en partie communs avec notre corpus de liste analysé dans le chapitre V et les 21 termes sélectionnés que nous présentons dans le Tableau 27.

Ce tableau comporte trois colonnes. De gauche à droite les colonnes apportent les informations suivantes : la nouveauté langagière en caractères et en pinyin est suivie du rappel des procédés néologiques exposés et recensés dans le chapitre V ; une traduction en français du mot ou de l'expression est proposée, d'abord littérale (unité par unité) puis globale afin de bien dégager sa valeur sémantique ; une présentation du contexte social, culturel et médiatique précise les circonstances d'apparition. En effet nous supposons que ces nouveautés langagières sont liées aux phénomènes sociaux et/ou aux pratiques individuelles des usagers de microblog et/ou aux environnements discursifs et technologiques.

Il est à noter que la traduction et la présentation du contexte sont destinées à la transmission des aspects néologiques chinois en français. Ces termes chinois sont des nouveautés langagières sur le plan morphologique, syntaxique ou sémantique. Nous faisons ici quelques rappels. Premièrement, par rapport à la langue française, le chinois est remarquable par la compacité de ses caractères (un caractère chinois ne correspond forcément à un mot français), lesquels sont morphologiquement inchangeables. Deuxièmement, il faut souligner que même pour un locuteur natif du chinois, il n'est pas toujours facile de saisir ces termes néologiques sans interroger leurs contextes et sources. Cela peut provoquer plus de difficultés de compréhension pour les lecteurs qui ignorent la langue et la culture chinoises.

Donc, la transposition de ces éléments nous demande de mobiliser des connaissances linguistiques, sociales et culturelles. Notre statut de bilingue est avantageux pour l'activité de traduction mais la méconnaissance de certaines réalités dans les deux pays et les deux langues nous empêche de trouver des termes équivalents en français. Il est aussi difficile de distinguer s'il existe des équivalents ou non comme il s'agit précisément de créations langagières.

Précisons enfin que nous n'avons pas pour but de proposer une traduction correspondante dans la langue française ou de fabriquer des nouveautés langagières françaises. Nous allons adopter une façon descriptive pour présenter les 21 termes chinois afin de faciliter notre analyse comparative.

Pour illustrer les aspects néologiques et les contrastes entre les deux langues, nous avons effectué une traduction littérale des termes chinois par unité de sens puis globale. Nous avons ensuite présenté le sens contextualisé de chaque terme chinois en nous référant principalement aux explications proposées par les deux institutions de sélection. Nous avons aussi consulté des sites encyclopédiques<sup>213</sup> pour obtenir des informations relatives au contexte.

---

<sup>213</sup> Site de référence :

- Blog du magazine 《咬文嚼字》 (yǎo wén jiáo zì): <http://blog.sina.com.cn/u/1275434742>, consulté le mars 2018 ;
- <https://www.cnkeywords.net/>, consulté le 8 avril 2018 ;
- Baidu Baike : <https://baike.baidu.com/>, consulté le 9 avril 2018.

N°	Nouveautés langagières (caractères et pinyin)  +  [procédé création]	Traduction littérale française selon unité de sens + traduction littérale globale <sup>214</sup>	Sens contextualisé
1	元芳, 你怎么看  yuán fāng, nǐ zěnmē kàn  [MS]	元芳/你/怎么/看  yuán fāng / nǐ / zěnmē / kàn  Yuan Fang [nom d'un personnage] / tu / comment / voir  <i>Yuan Fang, qu'en penses-tu?</i>	Cette phrase est issue d'un feuilleton télévisé policier de 2004. Elle est souvent utilisée pendant la conversation entre le procureur et son assistant Yuan Fang pour discuter les détails d'un procès. La réponse correspondante est : « Il y a quelque chose de louche ».  Elle est largement employée par les internautes en 2012 au sujet du décès mystérieux d'une fille qui est considéré finalement par la police comme un suicide. Ensuite, cette phrase devient populaire pour appeler à commenter les actualités. Ainsi, des imitations de cette structure surgissent en changeant des éléments, comme le prénom. Tous ces usages ont favorisé la formation de «la structure Yuan Fang » et son nouveau sens ironique.
2	舌尖上 (corpus A)  shé jiān shàng  [MS]  舌尖上的中国 (corpus B)  shé jiān shàng de zhōng guó  [MS]	舌尖/上 (corpus A)  shé jiān / shàng  le bout de la langue / sur  <i>sur le bout de la langue</i>  舌尖/上/的/中国 (corpus B)  shé jiān / shàng / de / zhōng guó  le bout de la langue / sur / préposition / Chine  La Chine sur le bout de la langue	《舌尖上的中国》 est le titre d'un documentaire chinois destiné à présenter les spécialités culinaires chinoises. Il signifie « Goûter la cuisine chinoise ». La version française du titre qui circule en ligne est « La Chine au bout de la langue »  Ce documentaire a connu un très grand succès. Ensuite, cette expression est devenue populaire pour exprimer des choses liées à la culture culinaire.  On constate aussi des variations comme « 舌尖上的快乐 » (shé jiān shàng de kuài lè), le plaisir de goûter les plats ; « 舌尖上的爱情 » (shé jiān shàng de ài qíng), relation amoureuse nouée par la gastronomie. « 舌尖上 » est sa forme raccourcie.
3	躺着也中枪  tǎng zhe yě zhòng qiāng  [SS]	躺着/也/中枪  tǎng zhe / yě / zhòng qiāng  s'allonger / aussi / se prendre une balle    <i>se prendre une balle malgré la position allongée.</i>	Généralement, la position allongée est une position de protection face aux menaces de coups de feu, comme cela, la probabilité d'être abattu est faible.  Initialement employée dans le dialogue d'un film comique, cette expression signifie: « Être mêlé à une affaire sans aucune raison ». En utilisant cette expression, on manifeste son étonnement en créant un effet humoristique.  Les formes raccourcies comme 躺中、躺枪 expriment le même sens. Une nouvelle expression figée et son nouveau sens se sont formés .

<sup>214</sup> Les barres obliques sont utilisées comme marqueur de séparation.

4	<p>高帅富 (corpus A)</p> <p>gāo shuài fù</p> <p>[PH]</p> <p>高富帅, 白富美 (corpus B)</p> <p>gāo fù shuài, bái fù měi</p> <p>[PH]</p>	<p>高/帅/富 (corpus A)</p> <p>gāo / shuài / fù</p> <p>grand / beau / riche</p> <p><i>garçon ou homme en possession d'avantages d'apparence physique et financière.</i></p> <p>高/富/帅, /白/富/美 (corpus B)</p> <p>gāo / fù / shuài, / bái / fù / měi</p> <p>grand / beau / riche / blanc / riche / belle</p> <p><i>garçon ou homme ou fille ou femme en possession d'avantages d'apparence physique et financière</i></p>	<p>« 高帅富 » est une nouvelle expression figée, combinant trois adjectifs positifs qui sont simples et directs, destinés à décrire les conditions majeures à remplir par un époux parfait selon des valeurs traditionnelles.</p> <p>La version féminine de cette expression est « 白富美 ». Elle est originalement employée par les usagers des forums pour décrire les caractères des héros dans les séries pour les ados ou dans les mangas, par exemple, un prince ou une princesse. L'utilisation de 3 adjectifs successifs crée un effet de redoublement. Selon le contexte, ces termes peuvent être péjoratifs ou positifs.</p>
5	<p>逆袭 (corpus A)</p> <p>nì xí</p> <p>[Emprunt]</p> <p>屌丝, 逆袭 (corpus B)</p> <p>diǎo sī, nì xí</p> <p>[Emprunt]</p>	<p>逆/袭 (corpus A)</p> <p>nì / xí</p> <p>inverse / attaquer</p> <p><i>contre-attaquer</i></p> <p>屌丝/, 逆/袭 (corpus B)</p> <p>diǎo sī / nì / xí</p> <p>perdant / inverse / attaquer</p> <p><i>la réussite d'un perdant</i></p>	<p>Cette expression d'origine japonaise signifie contre-attaquer. Après avoir fréquemment été utilisée dans des films et productions ACGN<sup>215</sup> japonais, elle est ensuite introduite en Chine dans le lexique du MMOG<sup>216</sup> pour désigner une stratégie d'attaque.</p> <p>Aujourd'hui, on constate un changement de sens et un élargissement d'application dans la société chinoise. Elle signifie actuellement qu'on réussit à faire quelque chose malgré des conditions difficiles ou défavorables, y compris la situation familiale, financière, le statut social.</p> <p>Lors de son utilisation, elle est souvent combinée avec une autre nouveauté langagière «屌丝» qui est l'antonyme de « 高帅富 », et fait référence à un jeune homme d'apparence médiocre et de statut social faible. Né dans une famille modeste, il n'a ni voiture, ni maison, ni relations. C'est plutôt une autodérision des jeunes chinois face à la pression sociale. Cette expression a aussi été appliquée dans les productions culturelles chinoises : les romans, les films, etc.</p>

<sup>215</sup> Voir note n°175.

<sup>216</sup> Sigle de l'anglais *Massively Multiplayer Online Game* qui signifie jeu en ligne massivement multijoueur.

6	<p>女汉子</p> <p>nǚ hàn zi</p> <p>[PH]</p>	<p>女/汉子</p> <p>nǚ / hàn zi</p> <p>féminin / mec</p> <p><i>femme masculine</i></p>	<p>La première utilisation de ce terme est celle d'une mannequin et animatrice LI Ai sur <i>Sina Weibo</i> pour décrire les filles ou femmes indépendantes, capables de faire quelque chose sans l'aide des hommes, particulièrement des tâches traditionnellement attribuées aux hommes, comme le bricolage.</p> <p>Ce caractère féminin est différent des stéréotypes sur les femmes, tant sur l'apparence que sur la personnalité. Il met en valeur l'image des femmes contemporaines.</p>
7	<p>土豪 (corpus A)</p> <p>tǔ háo</p> <p>[SS]</p> <p>土豪金 (corpus B)</p> <p>tǔ háo jīn</p> <p>[SS]</p>	<p>土/豪 (corpus A)</p> <p>tǔ / háo</p> <p>rustique / généreux</p> <p><i>les riches</i></p> <p>土/豪/金 (corpus B)</p> <p>tǔ / háo / jīn</p> <p>rustique / généreux / couleur d'or</p> <p><i>la couleur d'or qui représente les riches</i></p>	<p>Ce terme existe depuis longtemps dans la langue chinoise ; originellement, il désigne les riches des campagnes et particulièrement les personnes qui exercent une autorité tyrannique et excessive sur leur entourage dans les années 50 du siècle dernier.</p> <p>Son nouveau sens est né durant la conversation entre les joueurs des jeux d'Internet. On critique les autres joueurs qui gagnent facilement par l'achat d'« équipements » très chers mais non par des stratégies. Ce genre d'action détruit les loisirs et l'égalité du jeu. Aujourd'hui, il signifie : les nouveaux riches qui dépensent de l'argent sans réfléchir et juste exhiber leur fortune. Les productions achetées sont peu importantes, par ex. : acheter un café pour 1000 euros. Ces personnes sont riches mais pas nobles. En adoptant ce terme, on peut exprimer la jalousie, le mépris ou l'admiration.</p> <p>« 土豪金 » est une dérivation du terme « 土豪 ». « 金 » est la couleur d'or, un symbole de fortune en Chine. Cette expression désigne généralement tous les objets de couleur d'or, avec un teint exagéré. Les personnes qui utilisent ce genre de produit sont considérés d'avoir aucun goût d'esthétique.</p> <p>On observe aussi d'autre dérivation comme : « 土豪, 我们做朋友吧 » : nouveaux riches, je veux être ton ami/e.</p>
8	<p>高大上 (corpus A)</p> <p>gāo dà shàng</p> <p>[PH]</p> <p>高端大气上档次 (corpus B)</p> <p>gāo duān dà qì shàng dàng cì</p> <p>[PH]</p>	<p>高/大/上 (corpus A)</p> <p>gāo / dà / shàng</p> <p>supérieur / bien / meilleur</p> <p>haut de gamme, grandiose, classe</p> <p>高端/大气/上档次 (corpus B)</p> <p>gāo duān / dà qì / shàng dàng cì</p> <p>haut de gamme / grandiose / classe</p> <p><i>haut de gamme, grandiose, classe</i></p>	<p>« 高大上 » est la forme réduite de « 高端大气上档次 ». Il est destiné à décrire quelque chose de luxueux, de haut de gamme. Selon le contexte d'utilisation, on l'emploie aussi ironiquement. Il a été utilisé pour la première fois dans une série comique chinoise 《武林外传》(wǔ lín wài zhuàn). Depuis sa propagation en 2013, on constate son utilisation dans les forums, les programmes de divertissement, les médias, la publicité, etc.</p>
9	<p>萌萌哒</p> <p>méng méng da</p> <p>[PH]</p>	<p>萌萌 / 哒</p> <p>méng méng / da</p> <p>mignon / une particule modale</p>	<p>Cette expression est influencée par la culture japonaise « Moe » (萌え), « faisant référence à des sentiments de forte affection orientée principalement vers des personnages fictifs, issus des domaines de l'anime, du manga et du jeu vidéo »<sup>217</sup>. Ces personnages ont des points communs comme : grands yeux, oreilles d'animal comme celle du chat ou du renard. Ils s'expriment d'une façon mignonne. 萌 signifie à l'origine : germer. « 萌萌哒 » vise à décrire quelque chose ou quelqu'un qui a l'air mignon.</p>

<sup>217</sup> Wikipédia : [https://fr.wikipedia.org/wiki/Moe\\_\(Japon\)](https://fr.wikipedia.org/wiki/Moe_(Japon)), consulté le 9 avril 2018.

		<i>très aimable</i>	Il est difficile de retrouver sa première apparition en Chine mais il est largement utilisé en ligne sur divers réseaux sociaux ou autres plateformes de communication. On suppose que sa structure est le résultat de l'imitation d'une autre nouveauté langagière « 么么哒 » ( <i>mo mo da</i> , qui signifie bisous). L'emploi de deux mots identiques renforce le sentiment à exprimer. En effet « 哒 » est une onomatopée, fréquente dans la culture Moe. Ce genre d'expression qui essaie de se rapprocher de la langue des enfants, a pour but de créer une atmosphère de communication plus détendue et agréable.
10	宝宝 (corpus A) <i>bǎo bǎo</i> [SS]  吓死宝宝了 (corpus B) <i>xià sǐ bǎo bǎo le</i> [SS]	宝宝 (corpus A) <i>bǎo bǎo</i>  bébé  <i>un bébé innocent</i>  吓死/宝宝/了 (corpus B) <i>xià sǐ / bǎo bǎo / le</i>  mourir de peur / bébé / placé comme particule à la fin pour indiquer l'achèvement de l'action ou changement  <i>cela a effrayé le bébé</i>	Le terme « 宝宝 » est dérivé de l'expression « 吓死宝宝了 ». Selon la source encyclopédique <i>Baidu Baike</i> , pour sa création, il existe 8 possibilités qui sont liées par exemple aux productions audiovisuelles, aux programmes télévisés, à des personnes telles que les célébrités, les influenceurs, etc. Bébé est le sens original de « 宝宝 ». Actuellement, il est souvent employé par l'énonciateur pour s'appeler lui-même, c'est-à-dire « je ». C'est une manière mignonne de s'exprimer. Cela est étroitement lié à la culture Moe. Au lieu de dire, « ça me fait peur », dire « cela a effrayé le bébé » peut rapprocher les interlocuteurs.
11	主要看气质 <i>zhǔ yào kàn qì zhì</i> [PH]	主要/看/气质 <i>zhǔ yào / kàn / qì zhì</i>  principalement / voir / tempérament  <i>c'est le tempérament qui compte.</i>	Cette expression devient populaire grâce à une chanteuse qui a publié une photo sur Sina Weibo dans le but de promouvoir la vente de son album. Cette photo la montre mangeant un hamburger or certains usagers pensent qu'il n'est pas convenable de présenter ce genre de nourriture fast food dans une production culturelle et que le design de son image de communication manque de goût. Sa réponse (« On n'évalue pas quelqu'un par son apparence ou ses habits mais par son caractère et ses connaissances») a été beaucoup likée et partagée par les internautes. Sa circulation inter-plateforme est intéressante : un jeu social sur Wechat218 a été lancé alors que cette expression se propageait. Les usagers du jeu sont invités à partager dans la rubrique « moments » la photo dont ils pensent qu'elle s'accorde avec cette expression.
12	洪荒之力 <i>hóng huāng zhī lì</i> [SS]	洪荒/之/力 <i>hóng huāng / zhī / lì</i>  Temps immémorial / reliant le qualificatif et le mot qualifié / pouvoir  <i>le pouvoir du temps immémorial</i>	Selon la mythologie chinoise, au début des temps, « 洪荒之力 » est le pouvoir de déluge qui peut détruire le monde. L'expression a été connue récemment par les lecteurs d'un roman fantastique chinois et est devenue populaire suite à la mise en scène de ce roman en 2015, dans le feuilleton télévisé « huā qiān gǔ ».  Lors des jeux Olympiques de Rio en 2016, son utilisation par la championne de natation Fu Yuanhui lors d'une interview avec un journaliste après avoir battu un record asiatique, a changé le sens et créé un phénomène de partage sur les réseaux sociaux, surtout sur <i>Sina Weibo</i> , sous des formes différentes (images, émoticônes, etc.). Elle signifie actuellement : faire tout ce qu'on peut pour réaliser quelque chose.
13	吃瓜群众 <i>chī guā qún zhòng</i>	吃/瓜/群众 <i>chī guā qún zhòng</i>	Dans l'expression « 吃瓜群众 », « 瓜 » il s'agit à l'origine des graines de tournesol « 瓜子 » ( <i>guā zǐ</i> ) et non de la pastèque. Dans la vie quotidienne, quand les gens se réunissent pour bavarder, les graines de tournesol constituent l'une des nourritures préférées. Cette expression est créée pour décrire certains internautes qui réagissent à un billet en ligne mais n'ont pas pour but

<sup>218</sup> Une application de messagerie textuelle et vocale instantanée en Chine.

	[PH]	manger / pastèque / public  <i>le public ou les spectateurs qui mangent de la pastèque</i>	d'exprimer leurs opinions ou d'aider à résoudre les questions posées. Ils parlent d'autres sujets qui n'ont rien à voir avec la discussion ou ils mentionnent dès le début qu'ils sont des spectateurs et non des acteurs.  On suppose que si « 瓜子 » devient « 瓜 », c'est pour économiser du temps pour taper les caractères et réaliser une communication rapide et de ce fait les graines de tournesol deviennent pastèque. Suite au succès de cette expression, son utilisation est élargie, elle désigne aujourd'hui, les spectateurs curieux (et non seulement les internautes) du déroulement de quelque chose, soit parce qu'ils ne savent pas ce qui s'est passé, soit parce qu'ils accordent leur attention à quelque chose sans intervenir.
14	小目标 (corpus A)  xiǎo mù biāo  [SS]  定个小目标 (corpus B)  dìng gè xiǎo mù biāo  [SS]	小/目标 (corpus A)  xiǎo / mù biāo  petit / objectif  <i>petit objectif</i>  定个/小/目标(corpus B)  dìng / gè / xiǎo / mù biāo  fixer / un mot spécifique / petit / objectif  <i>fixer un petit objectif</i>	Ce terme a été employé par un homme d'affaire très connu en Chine durant un talk-show au cours duquel il était invité à partager des conseils aux jeunes Chinois qui veulent devenir riches. Il a répondu que c'était une bonne direction mais qu'il valait mieux commencer par se fixer un petit objectif, par exemple gagner 100 millions. Ses paroles ont provoqué une vive réaction du public pour qui cet objectif est impossible à réaliser.  Tout de suite, ce terme est devenu populaire en ligne et employé par ironie comme une antiphrase. C'est-à-dire que le petit objectif est en effet ambitieux et difficile à atteindre.
15	友谊的小船, 说翻就翻 (corpus A)  yǒu yì de xiǎo chuán, shuō fān jiù fān  [PH]  友谊的小船 (corpus B)  yǒu yì de xiǎo chuán  [PH]	友谊/的/小船/说翻就翻 (corpus A)  yǒu yì / de / xiǎo chuán / shuō fān jiù fān  l'amitié / précédé d'un nom exprime la relation d'appartenance / petit bateau / se renverser à tout moment  <i>le bateau d'amitié peut se renverser du jour au lendemain, ou à tout moment.</i>  友谊/的/小船(corpus B)  yǒu yì / de / xiǎo chuán  l'amitié / précédé d'un nom exprime la relation d'appartenance / petit bateau  <i>le bateau d'amitié</i>	Cette expression signifie en réalité la fragilité de la relation interpersonnelle. On peut ne plus être amis pour des raisons très simples : par exemple on ne mange pas épicé mais on est invité pour un repas de fondue chinoise très relevée.  « 友谊的小船 » au sens littéral signifie : le bateau d'amitié. Il se peut qu'il soit issu du mot anglais « friendship » (amitié) et que la traduction chinoise ait divisé le mot en deux : « friend » traduit comme l'amitié et « ship » comme bateau. On peut aussi constater ce jeu de mot dans les dialogues de la fameuse sitcom américaine <i>Friends</i> . Cette expression devient populaire en ligne grâce aux bandes dessinées de Nan Dongni qui montrent deux pingouins sur un petit bateau qui chavire à cause de différents changements de ces deux animaux (ex. : l'un a maigri).  Ensuite, les internautes ont créé diverses dérivations en imitant la structure. Par exemple : le bateau du mariage, le bateau du travail. Cette expression a tellement de succès qu'on peut être compris par les autres sans écrire la phrase complète.



16	葛优躺 gě yōu tāng [PH]	葛优/躺 gě yōu / tǎng nom d'un acteur chinois / s'allonger Geyou s'allonge	Depuis 2016, une série d'images macro <sup>219</sup> est apparue en ligne. Le héros de ces images, est un personnage du feuilleton télévisé « 我爱我家, wǒ ài wǒ jīa », qui s'allonge sur le canapé après avoir bien mangé. Comme l'acteur s'appelle 葛优 (gě yōu), les internautes ont nommé cette position paresseuse par le nom d'acteur. On suppose que la réussite de cette position et ce terme répond aux besoins psychologiques de ces usagers qui subissent la pression sociale et un rythme de vie accéléré.
17	套路 tào lù [SS]	套/路 tào / lù série / moyen <i>un ensemble de méthodes</i>	L'expression « 套路 » existe déjà dans la langue chinoise. Elle signifie une série de mouvement d'arts martiaux ou un ensemble de méthodes.  En ligne, surtout dans les jeux, cette expression a été dotée d'un nouveau sens. « 套 » est interprété comme le piège, donc « 套路 » devient « des stratégie de tromperie ». Ces tromperies sont souvent préplanifiées en suivant certaines modalités. Après avoir connu une utilisation élargie, elle est souvent employée péjorativement.
18	蓝瘦, 香菇 lán shòu, xiāng gū [MS]	蓝/瘦/香菇 lán / shòu / xiāng gū bleu / mince / champignon shiitake  <i>bleu et mince, champignon shiitake</i>	« 蓝瘦, 香菇 » est paronymie de « 难受, 想哭 » (nán shòu, xiǎng kū) qui signifie trop triste et avoir envie de pleurer. Donc, il est difficile de connaître le vrai sens sans avoir cette connaissance préalable.  L'origine de cette expression est liée à un internaute venant de la province de Guangxi. Il a posté en ligne une vidéo exprimant sa tristesse après s'être séparé de sa copine. En raison de son fort accent, les internautes n'entendent pas clairement « 难受, 想哭 » et ensuite le transcrivent et créent cette expression drôle « 蓝瘦, 香菇 ».  Nous supposons que le choix du mot « 蓝 » (bleu) n'est pas dû au hasard. On joue sur deux aspects : 1) la couleur bleu a une connotation spéciale, la tristesse. 2) « lan » à la place de « nan » renforce la présence de l'accent (les gens du Sud de la Chine ont de la difficulté à distinguer la prononciation L et N). En plus, « 瘦 » signifie maigrir. Il est raisonnable de penser que quand on est trop triste, on perd du poids. Le remplacement de « 想哭 » par « 香菇 », est aussi destiné à souligner l'influence de l'accent en ajoutant des effets humoristiques.  Cette expression est devenue populaire en très peu de temps et a circulé dans toute la Chine quelques heures après sa création. Par la suite, elle est présente sur des supports différents, en ligne et hors ligne. Dans certaines cantines, on a aussi créé et nommé des plats avec cette expression, avec bien sûr des champignons comme ingrédient. On observe même la création d'une société utilisant cette expression comme nom de marque.
19	油腻 yóu nì [SS]	油/腻 yóu / nì gras / grasieux; gras et lourd	«油腻» signifie initialement quelque chose de trop gras. En 2017, un usager de <i>Sina Weibo</i> a publié un tweet dans lequel ce terme est employé pour décrire certains caractères péjoratifs des personnes d'un âge moyen. Depuis, ce terme est porteur d'un nouveau sens. Il est destiné à montrer les apparences et les comportements qui provoquent un sentiment désagréable voire écœurant, pas seulement concernant des personnes d'âge mûr.  En même temps, on peut constater la reconnaissance et l'utilisation de ce terme dans les tweets des célébrités, des influenceurs, etc.

<sup>219</sup> Une image macro (terme d'origine anglaise) est un élément de support numérique composé par une image et des éléments textuels superposés pour créer souvent des effets humoristique ou ironique. Il s'appelle en chinois 表情包 (biǎo qíng bāo).

20	怼 duǐ  [MI] <sup>220</sup>	怼 duǐ  haïr du fond du cœur <i>haineux</i>	<p>La prononciation initiale de « 怼 » est <i>duǐ</i> qui signifie haïr (emploi comme verbe) ou cruel (adjectif). La racine de ce mot est « 心 » (xīn, le cœur). Cela montre un sentiment haineux profond, venant du fond du cœur. Dans le dialecte de la province Henan, « 怼 » peut être utilisé comme un verbe universel : offenser, critiquer selon le contexte.</p> <p>Son nouveau sens populaire en ligne, est issu d'une émission de télé-réalité, où ce terme est employé avec le sens de « réfuter ». C'est un emprunt à un dialecte mais certains estiment que, c'est peut-être une fausse utilisation du dialecte puisque qu'il existe un autre caractère « 撑 » avec la même prononciation et un sens rapproché. Suite à sa propagation, de nouvelles interprétations comme : disputer exprès, rivaliser, sont aussi ajoutées pour ce terme. Il est aussi beaucoup présent dans des articles de presse.</p>
21	打 call dǎ call  [ME]	打/call dǎ /call  taper / téléphone <i>téléphoner</i>	<p>Cette expression constitue un terme de la culture des fans, particulièrement les fans de vedettes (wota), dans le domaine des concerts Live. Ce genre de culture s'est d'abord développé en Corée du Sud et au Japon. Il englobe différents moyens de soutien des fans pour leurs idoles.</p> <p>« 打 call » est créé à partir du mot japonais « コール ». Par « 打 call », les fans crient sur place ou agitent par exemple des tiges fluorescentes, éventails, bandeaux pour montrer leur enthousiasme en créant une atmosphère chaleureuse. C'est un moyen d'interaction entre les fans et leurs idoles. Il faut souligner que ces interactions sont strictement préplanifiées selon le rythme de chanson. Nous pouvons supposer que l'emploi du caractère chinois « 打 » est pour décrire l'action ; « call » est une transcription phonétique de l'expression japonaise. Une fois introduit en Chine, son utilisation n'est pas restreinte au domaine mentionné ci-dessus. Il signifie généralement, encourager quelqu'un à faire quelque chose, comme « allez » dans la langue française. Parfois, les usagers négligeant cet aspect culturel, considère « call » comme téléphone. Cela faisant, on exprime l'encouragement aussi par « 打电话 » (电话 signifie téléphone). Or, l'expression « 打电话 » signifie originellement « téléphoner » en chinois.</p>

**Tableau 27 : Nouveautés langagières communes traduites et contextualisées**

<sup>220</sup> Dans le chapitre V, nous avons ajouté une catégorie indépendante aux matrices MS, SS, PM et PH dans les matrices internes pour souligner les influences des dialectes en chinois et en français, les influences des variants de français régional et non hexagonal. Ici nous utilisons MI (matrice interne) pour dénommer cette catégorie et pour distinguer des autres procédés de la typologie de référence.

Pour ces 21 nouveautés langagières nous nous sommes premièrement intéressée aux sources de création. D'où viennent-elles ? Des points de vue différents s'expriment selon les chercheurs. Le chercheur chinois Mo (2019)<sup>221</sup> a analysé les propositions de 2014 à 2017 de l'activité 《汉语盘点》 (hàn yǔ pán diǎn) et défini trois grandes catégories de phénomènes à l'origine des créations :

- 1) les faits d'actualité ou grands évènements qui attirent l'attention du grand public et des médias ; ex. : l'expression 洪荒之力 (n°12) qui est liée aux Jeux Olympiques de Rio en 2016;
- 2) l'influence des célébrités dont les paroles et comportements sont volontiers imités ; ex. : le fameux 小目标 (n°14) qui s'est chargé de nouveaux sens suite à son emploi par un homme d'affaire très connu en Chine ;
- 3) le contexte de l'Internet (englobant l'intercommunications en ligne, sur les réseaux sociaux ou les jeux en ligne ; les descriptions de l'environnement numérique ou les expressions de point de vue ou d'humeur en ligne) ; ex. : 吃瓜群众 (n°13) qui signifie « des spectateurs ou témoins passifs », 蓝瘦, 香菇 (n°18) qui sert à exprimer la tristesse, etc.

Dans toutes ces catégories, la création est libre, volontaire, et émane du grand public.

D'après deux autres chercheurs chinois Shen et Pu (2018), les 10 termes élus annuels proposés par 《汉语盘点》 (hàn yǔ pán diǎn) peuvent être considérés comme des mots ou expressions en vogue en prenant en compte leur fréquence d'utilisation. Après avoir étudié les propositions de 2014 à 2016 de l'activité et consulté les travaux chinois et japonais portant sur les sources de création de ces 30 mots et expressions en vogue, ils les ont classé en 9 catégories<sup>222</sup>. En effet, selon eux, les 6 catégories du chercheur japonais Akihiko Yonekawa<sup>223</sup> fondées sur un point de

---

<sup>221</sup> Dans cette recherche, à travers l'étude sur l'origine des nouveautés langagières sélectionnées, Mo s'est interrogé sur les causes principales de la création.

<sup>222</sup> Nous traduisons les catégories comme suit : 1. En lien avec des faits d'actualité ; 2. Les influences des paroles des célébrités ; 3. Pour nommer des nouvelles réalités économiques ; 4. Les influences des dialectes ; 5. Inspiré des productions culturelles ; 6. Les nouveaux mots et expressions publicitaires ; 7. Les slogans ; 8. Les nouveaux lexiques pour exprimer les sentiments ; 9. Les influences des langues et cultures étrangères. (Shen et Pu 2018).

<sup>223</sup> Notre traduction : 1. Issus des sujets de discussion communs ; 2. Pour s'exprimer sur un ton humoristique ; 3. Pour mieux désigner les nouvelles réalités ; 4. Pour impressionner les locuteurs ; 5. Imiter les paroles des célébrités ; 6. Les nouvelles expressions publicitaires. (Yonekawa 1989:11-17, cité dans Shen et Pu 2018).

vue social, ne sont pas complètes, alors que le classement en 4 catégories du chercheur chinois Qi<sup>224</sup> mettent l'accent plutôt sur les aspects linguistiques. Shen et Pu classent les sources d'un point de vue à la fois social, culturel et linguistique, ce qui nous paraît très convaincant.

Certes, les prises en position variées et l'analyse portant sur des corpus non unifiés peuvent provoquer des divergences. Nous remarquons que ces chercheurs partagent des points communs sur certaines sources de création : les faits d'actualité, les expressions des célébrités, les productions culturelles (ex. : film, série télévisée), y compris des productions étrangères<sup>225</sup>, et l'influence linguistique (emprunt ou dialecte). Revenons sur le détail de chaque étude pour mieux comprendre les différences entre les classements.

Le classement de Mo est assez général par rapport à celui de Shen et Pu, surtout pour la troisième catégorie, le contexte de l'Internet. Même si nous avons cité quelques exemples pour illustrer cette catégorie, l'influence du contexte de l'Internet nous semble un aspect très vague et dont la frontière avec les deux autres catégories est floue. Ainsi un fait d'actualité peut être étroitement lié au contexte de l'Internet voire ne se produire et ne circuler qu'en ligne ; certaines célébrités n'effectuent leurs activités que sur l'Internet. C'est le cas des « influenceurs » sur diverses plateformes numériques. Par exemple, le terme n°11 du tableau est à la fois un fait social, puisque ce sujet a provoqué de vives discussions, l'expression d'une célébrité visant à partager son point de vue, et une pratique relevant du contexte de l'Internet, puisqu'il s'agit d'un mode d'intercommunication à travers la plate-forme du microblog et sa fonctionnalité « commentaire ».

Quant à Shen et Pu, ils proposent de classer les sources de création d'une manière plus précise que Mo en prenant compte les aspects linguistique, social et culturel. Mais dans leurs propositions, il est difficile d'identifier séparément ces trois aspects. Une publicité ou un slogan sont-ils plutôt une production culturelle ou linguistique ? La catégorie du dialecte est-elle définie pour son aspect linguistique ou social ? A propos de chaque catégorie, il semble qu'il s'agit d'une énumération de toutes les sources au lieu d'un classement synthétique. Cela nous

---

<sup>224</sup> Notre traduction : 1. Les lexiques destinés à désigner les nouveaux concepts ; 2. Les nouveaux lexiques de l'économie ; 3. Les influence des dialectes ; 4. Les lexiques d'emprunt. (Qi 2006:138-140).

<sup>225</sup> Nous avons aussi relevé dans les explications de l'activité publiée par le Ministère de l'Éducation chinois ([http://www.moe.gov.cn/s78/A19/A19\\_ztzt/ztl\\_yywfzfw/shenghuoxz/](http://www.moe.gov.cn/s78/A19/A19_ztzt/ztl_yywfzfw/shenghuoxz/)), certaines sources de création comme : film, émission de télé-réalité ou plateformes de diffusion en direct.

amène à nous demander si cette « précision » est nécessaire et si leurs catégorisations peuvent s'appliquer à l'ensemble de notre sélection de 21 termes. Pour certaines nouveautés langagières qui sont populaires et circulent sur des supports variés, le critère du choix de source nous paraît peu clair.

Prenons le terme n°12. Shen et Pu considèrent que sa source de création est l'influence des célébrités. Or, l'étude contextuelle du terme nous a appris qu'il avait connu trois périodes d'évolution. Son emploi dans un roman a déjà modifié son sens d'origine. Ensuite, la mise en scène du roman a favorisé une propagation de son sens néologique. Et dans la dernière étape, l'emploi par un sportif durant les Jeux Olympiques de Rio a mis ce terme au centre de la discussion et provoqué une diffusion élargie en ligne et sur les médias. Selon nous, sa source correspond à la catégorie « production culturelle » : le rôle de la célébrité n'a aidé qu'à la diffusion mais n'a pas participé à la création sémantique.

À travers l'étude sur le sens contextualisé de cette sélection de 21 mots ou expressions, nous avons constaté une évolution dynamique de ces nouveautés langagières. Il faut remarquer que leurs premières occurrences ne relèvent pas toujours d'un emploi néologique. Ainsi pour les termes n°1 et n°3 dans notre tableau, lors de leurs utilisations dans un feuilleton et un film, leur sens n'est pas néologique, mais c'est au moment où ils sont employés pour commenter les faits d'actualité (n°1) et pour exprimer son sentiment envers quelque chose (n°3) que leur sens évolue néologiquement. Outre les supports de circulation, c'est ce processus complexe qui est négligé par les deux travaux mentionnés ci-dessus. C'est pourquoi nous avons développé particulièrement la présentation du sens contextualisé.

Nous pensons qu'il est possible d'appliquer les catégorisations proposées mais qu'il est important de distinguer les premières occurrences et les néologismes. De ce fait pour les n°1 et n°3 par exemple, au lieu de les classer sous l'origine « productions culturelles », il peut être plus convenable de relier respectivement leurs sources aux faits d'actualité et à l'expression de l'humeur. Nous n'allons pas chercher à proposer de nouvelles catégories ni un classement révisé mais nous allons reclasser les 21 nouveautés langagières dans les catégories qui nous paraissent les plus pertinentes en gardant présent à l'esprit que les aspects social, culturel et linguistique sont inséparables. Après examen des différents points de vue et retour sur les contextes des 21 termes, nous considérons que leurs sources sont les suivantes : les faits d'actualité (n°1), les paroles de célébrités (n°6, n°11, n°14), l'expression d'une opinion ou d'un sentiment (n°3, n°4,

n°7, n°10, n°13, n°17, n°19), les productions culturelles (n°2, n°8, n°12, n°15, n°16) et l'influence linguistique (n°5, n°9, n°18, n°20, n°21).

La création et la propagation de ces nouveautés langagières sont aussi inséparables de la communication interpersonnelle. À travers leurs sens contextualisés, nous pouvons apercevoir tout d'abord qu'il s'agit d'une création libre. Cette création est manifestée par les éléments suivants :

1) *Chaque locuteur est un créateur potentiel.* Même s'il est difficile de retracer les créateurs d'origine de ces 21 termes, nous pouvons intuitivement conclure que ces productions lexicales sont issues de l'emploi individuel ou d'une communauté collective (le cas des forums ou les communautés de jeux vidéo en ligne). Par exemple, les termes n°6, 11, 19 sont utilisés personnellement par des usagers de *Sina Weibo* ; les n°12, 14 ont connu une évolution de sens après avoir été employés par des influenceurs. Les n°4, 7 sont étroitement liés à des communautés en ligne. En même temps, nous avons constaté que leurs emplois se font sur des supports variés qui englobent les médias traditionnels comme la télévision, le cinéma, les journaux et les nouveaux dispositifs médiatiques comme les plateformes en ligne, les forums ; les microblogs. Cela montre que les producteurs et les diffuseurs des nouveautés langagières peuvent toucher un public large et varié. Cette action individuelle est aussi marquée par une création spontanée. En outre, les emplois sur les multiples supports mentionnée ci-dessus contribuent davantage à une diffusion mixte et rapide. Une nouveauté langagière peut être connue à l'échelle nationale en seulement quelques heures. Tout cela est différent des nouveautés langagières imposées par des institutions. Les nouveautés langagières sont créées et/ou utilisées par le grand public. Les créations et/ou diffusions sont difficiles à prévoir.

2) *Pour des objectifs variés : à chaque usager, ses propres pratiques.* Selon nos observations sur les sources de création ci-dessus, nous avons déjà constaté des éléments sociaux, culturels et linguistiques qui sont attachés à ces nouveautés langagières. Nous pouvons avancer qu'elles sont créées en lien avec les objectifs suivants :

- *Exprimer un point de vue.* Cela concerne non seulement les affaires personnelles des usagers, mais aussi des faits sociaux. Par exemple, le terme n°11 est employé par une chanteuse pour montrer son désaccord avec les commentaires de son billet de microblog critiquant sa présentation sur la couverture d'album. Les autres individus se sont fait écho immédiatement de cette expression d'idée de valeur personnelle. Ainsi, la propagation du terme n°15 indique le partage du point de vue de l'auteur sur le sujet de l'amitié. Ces opinions communes se sont ensuite élargies à d'autres sujets intimement liés à la vie quotidienne comme le travail, le mariage, à travers les dérivations. D'ailleurs, les individus d'aujourd'hui n'économisent pas leur attention à l'environnement social ou aux communautés auxquelles ils participent. L'expression d'opinions sur les faits sociaux les amène à donner une appréciation négative (voir n°1, 14, 17). Le terme n°1 peut être considéré comme un exemple typique. Son nouveau sens est fondé sur les doutes des usagers sur le jugement d'un décès par la police. Ce terme créant une nouvelle structure figée est appliqué spécialement pour appeler à commenter et exprimer un manque de confiance face à « des réalités » annoncées.
- *Exprimer une émotion.* Les nouveautés langagières sont aussi employées pour évoquer les différents sentiments personnels. Les sentiments transmis à travers les termes n°5, 16 sont liés à la pression sociale. Le n°3 exprime l'étonnement envers les affaires dues aux relations interpersonnelles. Selon le contexte d'utilisation, le n°8 peut exprimer admiration ou moquerie. Le rapprochement avec la langue des enfants du terme n°9 a pour but de manifester une tendresse aimante en créant un effet agréable et non offensif. Le n°19 traduit un sentiment de répugnance à l'égard de certains comportements.
- *Décrire les nouvelles réalités.* Les nouveautés langagières naissent aussi de la nécessité de dénomination. Dans les 21 termes, nous avons repéré plusieurs termes concernant des nouvelles pratiques de domaines variés : Le terme n°2 est associé à la culture culinaire ; les termes n°4 et 21 sont influencés respectivement par la littérature et le divertissement japonais tandis que le n°7 est le produit de nouvelles pratiques dans les communautés de jeux vidéo en ligne.

➤ *Produire des effets de sens.* La production d'effets divers durant la communication constitue un autre caractère original de ces nouveautés langagières. Dans notre sélection, nous avons remarqué que la portée de certaines nouveautés peut être humoristique ou ironique. Il faut souligner que ces deux procédés sont différents l'un et l'autre. Pour nous en tenir aux définitions d'un dictionnaire usuel : l'humour, mot anglais francisé depuis le 18<sup>e</sup> siècle, signifie « forme d'esprit qui consiste à présenter la réalité de manière à en dégager les aspects plaisants et insolites » alors que l'ironie, plus offensive, consiste à « se moquer (de qqn ou de qqch.) en disant le contraire de ce qu'on veut faire entendre » (*Le Petit Robert* 2012). L'humour est plutôt réflexif car celui qui le pratique s'inclut lui-même dans le rire alors que l'ironie rit des autres. Par exemple, le terme n°3 utilise une parole d'un film comique connu en Chine. En employant ce terme, le locuteur se compare au personnage du film et associe ses expériences à l'intrigue plaisante de celui-ci. Ceux qui partagent les mêmes connaissances peuvent immédiatement comprendre le contexte. Le terme n°5 contient une autodérision des jeunes Chinois face à la pression sociale. On se décrit soi-même comme un perdant puisqu'on ne bénéficie pas de conditions favorables. Dans ce cas-là, la réussite est considérée comme une surprise. C'est d'ailleurs une appréciation de l'effort personnel et de l'attitude consistant à ne pas abandonner face aux difficultés. Quant au terme n°18, il s'est appuyé sur la paronymie pour transmettre à la fois la dimension dialectique et le fait de ne pas pouvoir bien articuler dû à la tristesse du locuteur. Ces exemples nous montrent qu'à travers l'effet humoristique, on cherche à créer un rapprochement pour pouvoir partager des sentiments, des points de vue voire des valeurs d'une manière bienveillante.

La portée ironique joue plutôt sur l'antiphrase (termes n°4, 8, 14). Selon le contexte, les adjectifs positifs comme *grand, beau, de haut de gamme* dans les termes n°4 et 8 peuvent être considérés comme des flatteries feintes, soulignant un contraste absurde. Par l'emploi du mot *petit*, le terme n°4 cherche à souligner la difficulté d'atteindre un objectif ambitieux. L'effet d'ironie est aussi présent dans les termes en relation avec les faits d'actualité, comme c'est le cas du terme n°1. La question « Que penses-tu ? » n'a pas pour but prioritairement d'avoir une réponse ou d'inviter les autres à se questionner mais de signaler une défiance.



C'est une façon d'exprimer implicitement : « Je ne crois pas du tout ou je ne suis pas d'accord ». Ici, l'ironie montre son caractère dépréciatif et dévalorisant permettant aux usagers d'accentuer leurs critiques envers les réalités sociales, d'exprimer leurs désaccords sur certaines valeurs soit individuelles ou collectives. Pourtant, l'humour ou l'ironie, pour réussir à créer l'effet, oblige les interlocuteurs à mobiliser leurs connaissances ou à se procurer des informations. Il faut aussi souligner l'importance du contexte de communication. Comme nous avons remarqué ci-dessus, les termes n°4, 8 peuvent aussi être employés comme des compliments positifs suivant l'application concrète.

Cette pré-analyse montre qu'il est intéressant d'observer l'évolution du sens de ces nouveautés langagières dans les échanges interpersonnels, c'est-à-dire sur un plan discursif. C'est l'une des raisons qui motive la constitution d'un corpus de tweets dans l'environnement de microblog.

### **6.2.3. L'étape de sélection d'un échantillon pour l'analyse discursive des tweets**

Si la sélection des nouveautés langagières communes aux corpus A et B nous a fourni 21 éléments que nous avons examinés sous l'angle des sources de création et de l'évolution dynamique ainsi que des caractéristiques communicationnelles et rhétoriques, nous n'avons pas pour autant arrêté les éléments qui doivent donner lieu à une collecte de tweets. En effet en prenant compte la probable fréquence d'utilisation de ces nouveautés langagières sur le microblog *Sina Weibo*, nous avons supposé que les billets de tweets associés à une seule nouveauté pouvaient être très nombreux. Sachant que nous ne disposons pas d'un outil d'analyse quantitative tel que celui développé par Longhi et les participants du projet #Idéo2017, et dans l'objectif de réaliser une analyse significative, nous avons continué à réduire le nombre d'entrées en faisant une nouvelle sélection parmi les 21 termes avant de résoudre la question de la sélection des tweets eux-mêmes.

Nous avons pensé nous fonder sur l'étude contextualisée des 21 nouveautés langagières et exploiter les catégorisations en prenant un terme par catégorie par exemple. Mais nous nous sommes souvenue aussi que le corpus de tweets est constitué par les propositions annuelles des institutions et que pour les années 2012 à 2017 on trouve des éléments communs parce que l'activité 《汉语盘点》 (hàn yǔ pán diǎn) ne propose les listes de palmarès qu'à partir 2012. Dès lors c'est le choix d'un terme par année (de 2012 à 2017) qui a paru un critère de nature à

simplifier le processus. Pour autant ce processus n'est pas terminé : quel terme choisir dans une année qui en comporte plusieurs ? En essayant de tenir compte des différents éléments, nous sommes parvenue à l'échantillon suivant (Tableau 28) :

Année de sélection	N° du terme dans le Tableau 27	Nouveautés langagières
2012	n°1	元芳，你怎么看 yuán fāng, nǐ zěn me kàn
2013	n°4	女汉子 nǚ hàn zi
2014	n°8	高大上 gāo dà shàng
2015	n°11	主要看气质 zhǔ yào kàn qì zhi
2016	n°12	洪荒之力 hóng huāng zhī lì
	n°18	蓝瘦，香菇 lán shòu, xiāng gū
2017	n°21	打 call dǎ call

**Tableau 28 : Echantillon de nouveautés langagières choisi pour la collecte de tweets sur *Sina Weibo***

Nous allons expliquer les choix auxquels nous avons procédé quand plusieurs termes étaient en concurrence pour une même année. Nous avons privilégié l'expression 元芳，你怎么看 par rapport aux trois autres termes de 2012 parce que sa longue structure (qui est associé à un sens connotatif et figé) est marquante, outre sa source de création et son effet ironique. Les emprunts japonais de 2013 et 2014 ont été éliminés car nous avons gardé le 3<sup>e</sup> emprunt japonais de forme mixte pour l'année 2017. Les deux termes de l'année 2013 sont de nouvelles dénominations mais avec 女汉子, il s'agit non seulement d'un nouveau sens mais aussi d'une nouvelle création d'expression dont nous connaissons la source de création. Quant à 高大上, c'est l'un des deux termes dont nous pouvons observer la continuité de diffusion : il a été élu dans le sous-corpus A pour 2013 et dans le sous-corpus B pour 2014. Le terme de l'année 2015 est aussi spécifique sur le plan formel et communicationnel : il s'agit d'abord d'une expression longue qui se rapproche de la structure d'une phrase, et ensuite il est en lien avec un personnage connu (il a été employé par une chanteuse pour exprimer son point de vue qui a provoqué de vives discussions). Pour l'année 2016, nous avons choisi deux entrées au lieu d'une seule : 洪荒之

力 constitue un néologisme syntactico-sémantique (changement de sens), cas que nous n'avons pas spécialement pris en compte dans les autres choix ; nous savons aussi qu'il a déjà connu plusieurs étapes d'évolution dont la dernière est associée à une propagation phénoménale. Le 2<sup>e</sup> choix de l'année 2016 蓝瘦, 香菇 est sélectionné pour sa prononciation qui met en jeu un accent dialectal et son effet humoristique.

#### 6.2.4. Démarche de sélection des tweets

Pour constituer un échantillon de tweets pour l'analyse, outre le choix des entrées se pose la question de la quantité mais aussi de la période d'observation et de la méthode de collecte. Pour repérer les occurrences d'utilisation sur les microblogs, nous avons besoin de nous servir des moteurs de recherche de *Sina Weibo* comme de *Twitter*. Ceux-ci peuvent afficher tous les résultats mais les résultats ne sont pas stables puisque la plateforme actualise en temps réel. Or une trop grande masse de résultats (si la nouveauté est utilisée à une haute fréquence) ne convient pas pour le type d'analyse que nous voulons mener. Nous ne pensons pas qu'il est indispensable pour repérer les caractéristiques des nouveautés langagières en discours d'observer tous les tweets, mais plutôt de dégager des typologies de circulation par quelques exemples.

Nous souhaitons d'abord connaître la situation générale d'utilisation des nouveautés langagières sur *Sina Weibo* afin de nous permettre d'observer leur vitalité et éventuellement limiter la période de sélection du corpus de tweet. Nous allons effectuer respectivement une recherche globale et floue pour observer leurs utilisations récentes (par rapport à notre date de recherche 9 juin 2020), et une recherche filtrée pour apercevoir l'évolution diachronique. L'observation sur l'année 2020 nous permet d'un côté de vérifier si les termes les plus anciens sont toujours en circulation et, d'un autre côté, d'observer la situation de circulation des termes « relativement » récents après leurs années de sélection, surtout pour celui de l'année 2017.

Nous avons d'abord effectué la recherche globale, par mots clés dans le moteur du microblog sans appliquer aucun filtre. L'objectif de cette recherche a été de repérer les billets de tweet les plus récents publiés. Selon l'ordre de publication, nous nous intéressons plutôt aux données sur la première page d'affichage. Le résultat montre que les 7 termes de notre échantillon sont toujours en usage actuellement au moment de la recherche (le 9 juin 2020). Les billets sur la première page sont presque tous publiés le 9 juin et certains peuvent dater seulement de

quelques minutes par rapport au moment de recherche. Nous donnons ci-dessous quelques exemples en capture d'écran (Figure 37 et 38). Il faut souligner que si les billets sont produits le jour même, le système utilise l'indication « aujourd'hui » (今天) au lieu de la date. L'heure de publication est l'heure chinoise (en sachant que le décalage horaire est de 6 heures d'avance sur la France pour l'horaire d'été).



l'indication de la date de publication: « aujourd'hui »

Figure 37 : Publication du terme 元芳, 你怎么看 datée du 9 juin 2020, 20:17

Le terme 元芳, 你怎么看 a été classé parmi les mots au palmarès en 2012. Nous avons repéré son utilisation récente environ une heure avant notre recherche intervenue le 9 juin 2020.



Figure 38 : Publication du terme 洪荒之力 datée du 9 juin 2020

Le terme 洪荒之力 classé au palmarès en 2016 a été utilisé seulement quelques minutes avant notre recherche du 9 juin 2020.

Cette recherche globale que nous avons effectuée pour les 7 termes, nous a permis d’obtenir un ensemble de données soit 50 pages de résultats pour chaque terme (le nombre maximal de propositions par recherche) dont les utilisations ne remontaient pas en deçà du mois de mai 2020 pour les 50 pages concernées. Cela nous révèle qu’ils sont tous en usage en juin 2020 vu le nombre de billets retrouvés. Pourtant, la contrainte de 50 pages ne nous permet pas de remonter jusqu’à leurs années de sélection et d’avoir un aperçu diachronique sur leur circulation. Pour y parvenir nous procédons à une recherche filtrée.

Nous avons indiqué *supra* que les fonctionnalités du microblog et les pratiques des usagers ne nous permettent pas d’accéder à l’ensemble des archives sur la plateforme. En même temps, la quantité des tweets est en modification continue avec les nouvelles publications et les éventuelles suppressions. Nous décidons donc d’observer la situation d’emploi de ces termes sur une période limitée mais il est à noter que nous ne pouvons pas obtenir les mêmes résultats affichés quand nous faisons des recherches à des moments différents vu que les changements de publication dynamiques et imprévus sur la plateforme. Nous n’allons prendre en considération que la quantité d’occurrences repérées durant une période qui débute l’année où les termes sont sélectionnés dans le palmarès. Rappelons que la quantité d’occurrences est

précisément un des critères de sélection des institutions chinoises pour élire ces nouveautés. On peut s'attendre à ce que les tweets de l'année soient extrêmement nombreux et dépassent la contrainte d'affichage.

Pour pallier cette difficulté, nous avons restreint notre observation à une journée. Nous supposons que si les occurrences d'utilisation repérées pendant une journée dépassent 50 pages, cela témoigne de l'utilisation fréquente du terme sur le microblog. Nous nous intéressons également à leurs utilisations dans les deux ans qui suivent pour tenter de connaître leurs évolutions possibles. Comme il est impossible de savoir la date précise de la première occurrence, nous décidons de faire notre observation le 31 décembre de chaque année pour trois années. Nous avons donc besoin de réaliser trois recherches filtrées pour chaque terme. Par exemple, pour le terme élu de 2012 元芳，你怎么看, les filtres sont appliqués sur les périodes de publication : 31/12/2012, 31/12/2013 et 31/12/2014. Le Tableau 29 ci-dessous synthétise les résultats de cette recherche filtrée :

N°	Année de sélection	Termes choisis (en caractères et en pinyin)	Dates d'observation et résultats
1	2012	元芳, 你怎么看 yuán fāng, nǐ zěn me kàn	31/12/2012: 50 pages 31/12/2013: 50 pages 31/12/2014: 42 pages
2	2013	女汉子 nǚ hàn zi	31/12/2013: 50 pages 31/12/2014: 50 pages 31/12/2015: 50 pages
3	2014	高大上 gāo dà shàng	31/12/2014: 50 pages 31/12/2015: 50 pages 31/12/2016: 50 pages
4	2015	主要看气质 zhǔ yào kàn qì zhi	31/12/2015: 50 pages 31/12/2016: 31 pages 31/12/2017: 12 pages
5	2016	洪荒之力 hóng huāng zhī lì	31/12/2016: 50 pages 31/12/2017: 50 pages 31/12/2018: 50 pages
6	2016	蓝瘦, 香菇 lán shòu, xiāng gū	31/12/2016: 50 pages 31/12/2017: 12 pages 31/12/2018: 7 pages
7	2017	打 call dǎ call	31/12/2017: 50 pages 31/12/2018: 50 pages 31/12/2019: 50 pages

**Tableau 29 : Résultats de la situation d'emploi par recherche filtrée**

Les résultats de cette recherche filtrée nous montrent que ces nouveautés langagières ne sont pas éphémères puisque nous avons pu repérer des occurrences d'utilisation durant les deux ans qui suivent leur sélection. Pour les termes n°2, 3, 5 et 7, leurs occurrences journalières sont toujours importantes, alors que les termes n°1, 4, 6 ont connu une baisse d'emploi en 3 ans, surtout les n°4 et 6. Certes, la publication d'un tweet est le choix d'un usager donné à un moment donné et le nombre de publications et l'utilisation d'un certain terme sur une journée sont variables. Autrement dit, les occurrences ne sont pas stables et en cela les résultats que nous avons obtenus ne peuvent pas être considérés comme représentatifs. L'objectif de notre recherche filtrée étant d'observer les occurrences de ces termes sur le plan diachronique, les résultats obtenus attestent du moins nettement de leur vitalité durant une période limitée.

Nous nous heurtons cependant à des problèmes de quantité et d'imprévisibilité. Certes la recherche globale et la recherche filtrée nous ont confirmé que les 7 nouveautés langagières ont été très utilisées durant une période limitée et sont toujours en usage dans l'environnement du

microblog (en juin 2020). Pourtant, les fonctionnalités de recherche limitées par la plateforme ne nous permettent pas de connaître véritablement leurs tendances d'application dans une perspective macro. Même sur une journée, les occurrences d'emploi sont nombreuses. D'autre part, nous nous sommes rendu compte que l'évolution du sens dépend davantage des pratiques des usagers, qui sont imprévisibles, que de facteurs purement chronologiques. Cette imprévisibilité est soulignée par Paveau (2017).

Les observations obtenues et les questions rencontrées durant les recherches globale et filtrée, ainsi que les connaissances sur les fonctionnalités de deux plateformes nous a conduit finalement à observer les tweets contenant les termes hashtagués et collecter pour chaque terme élu 10 tweets afin de constituer notre corpus pour répondre aux principes méthodologiques que nous avons présenté au début du chapitre VI. Nous ne serons pas contrainte par l'ordre d'affichage et l'augmentation des publications. Ainsi, le choix d'observer les tweets populaires affichés dans la rubrique de résultats « à la une » nous a aidé à lever la difficulté du choix de période pour la sélection. Mais durant la collecte effective des tweets sur *Sina Weibo*, nous avons constaté premièrement que la quantité de tweets populaires dans la rubrique « à la une » n'est pas importante (de 2 à 10 pages)<sup>226</sup>. Après avoir éliminé les tweets selon les critères énoncés (exclusion des tweets contenant des vidéos et photos représentant des personnes), il nous reste très peu de tweets (voire moins de 5) disponibles. Face à cette situation et pour pouvoir obtenir un nombre équilibré pour chaque terme, nous avons abandonné la recherche dans la rubrique « à la une » sur *Sina Weibo* et commencé à chercher les 10 tweets parmi l'ensemble de résultats (dans la rubrique intitulée 综合) relatifs au « hashtag ». Nous pouvons ainsi disposer suffisamment de tweets entre lesquels choisir. Après cette démarche, nous avons réussi à collecter 70 tweets pour la constitution effective du corpus de tweets chinois. Les captures d'écran de ces tweets sont disponibles dans l'annexe n°4.

### **6.3. La constitution du corpus de tweets français**

#### **6.3.1 Premières observations sur le corpus de listes français et ses sources**

Le corpus de tweets chinois est constitué à partir d'une sélection de nouveautés langagières prélevée au sein des éléments en commun des deux sous-corpus de listes chinois. Nous tentons

---

<sup>226</sup> Pour la sélection sur *Twitter*, en raison de la différence technologique de mise en page, nous n'avons pas rencontré le même problème. Nous allons continuer à la sélection dans la rubrique « à la une ».



de collecter le corpus de tweets français de manière analogue. Pour cela, il est nécessaire d'entreprendre une observation comparative sur les deux sous-corpus français constitués respectivement par les nouvelles entrées de chaque dictionnaire. Rappelons que le principe d'organisation du corpus de listes français ne repose pas sur des activités institutionnelles mais sur les nouvelles éditions annuelles de deux dictionnaires largement diffusés, rapportées par plusieurs organes de presse.

Cette multi-source des nouvelles entrées citées est une caractéristique du corpus de listes français qui nous permet d'accéder à des données plus riches pour certaines années que pour d'autres. Dans l'ensemble, nous avons constaté dans un premier temps que les nouvelles entrées des dictionnaires sont présentées principalement de deux manières par les organes de presse :

- 1) Les nouveaux mots ou expressions sont pré-classés par ces organes de presse et ensuite introduits thématiquement par exemple en vocabulaire culinaire, technologique, médical (*yuzu, Flashcode, algologie*) ; en lien avec l'économie, l'éducation ou l'Internet (*le made in, MOOC, viralité*). Ou bien encore ce sont les origines qui sont mises en exergue : les anglicismes, les emprunts aux pays de la francophonie, les mots ou expressions familiers (*lol, chialeux, gloups*). Ce pré-classement proposé varie d'une année à l'autre et d'un article à l'autre même si on retrouve toujours plus ou moins les mêmes rubriques thématiques.
- 2) Certaines nouvelles entrées sont intégrées dans un contexte fabriqué. Nous en voyons ci-dessous un exemple avec la présentation de la nouvelle édition 2014 du *Petit Larousse* par *20 Minutes*<sup>227</sup> :

*Bernard Pivot*<sup>228</sup> fait son entrée et s'il n'est pas «botoxé», ni sans doute le fruit d'un «speed dating», le journaliste sait «monter au filet» pour défendre la langue française. Toujours beaucoup d'intronisés issus des nouvelles technologies, comme «googliser» ou «googler», «hashtag», «post», «extoter», «télévision connectée», «cyberdéfense» ou «même».

Ces deux types de présentation sont complètement différents de celles des institutions chinoises

---

<sup>227</sup> Source : <https://www.20minutes.fr/culture/1164155-20130529-20130529-bernard-pivot-tous-mots-entrent-dico>

<sup>228</sup> Il est journaliste et écrivain français, twitteur et figure de la télévision du temps de l'émission littéraire *Apostrophes*. Il préside l'Académie Goncourt de 2014 à 2019. Depuis 2011, son compte *Twitter* a attiré de nombreux abonnés. Une compilation de ses tweets est sorti sous forme d'ouvrage (*Les Tweets sont des chats*, Albin Michel) en 2013. Il explique dans cet ouvrage : « J'aime les tweets parce qu'ils partent en silence, circulent en silence et arrivent en silence. Les tweets sont des chats. »

qui annoncent un palmarès et proposent simplement une liste en énumérant le top 10 des mots et expressions sélectionnés.

Par ailleurs, nous avons également constaté que certaines nouvelles entrées d'une même édition sont citées en commun par plusieurs articles de presse. Il ne faut pas négliger le fait qu'un article peut être la source de citation d'un autre article. Ainsi en introduisant la nouvelle édition du *Petit Larousse* 2015, *20 Minutes* fait référence aux nouvelles entrées annoncées par *Le Figaro*. D'un autre côté, nous nous demandons si cette répétition peut impliquer la popularité et la fréquence d'utilisation de ces nouvelles entrées.

Selon le récapitulatif du corpus français dans le chapitre V, parmi les articles publiés, ceux du *Figaro* occupent un pourcentage important, suivis par les publications de *20 Minutes*. On remarque que *Le Figaro* consacre systématiquement un article aux nouvelles éditions du dictionnaire *Larousse* (2011, 2012, 2014, 2015, 2016, 2017) et qu'il cite dans ces articles une grande quantité de nouvelles entrées. Quant au *Petit Robert*, même si ses nouvelles éditions annuelles sont aussi traitées par les mêmes organes de presse, les articles qui y sont relatifs, d'après la recherche que nous avons effectuée, sont plus dispersés, comptant en moyenne deux ou trois publications sur 7 ans pour chaque organe de presse.

### 6.3.2 Une démarche adaptée de sélection des entrées

En cherchant à adopter le même principe d'intersection de nouveautés en commun, nous avons d'abord comparé les nouvelles entrées apparues sur les mêmes années de publications, c'est-à-dire 2011, 2015, 2016 et 2017 (étant donné que nous n'avons pas trouvé d'articles consacrés à citer les nouvelles entrées de l'un ou l'autre dictionnaire pour certaines années<sup>229</sup>) et nous avons trouvé seulement un mot en commun, le *big data* qui est annoncé comme intégré dans les deux dictionnaires. Par la suite, à travers une comparaison sur l'ensemble des nouvelles entrées rapportées constituant notre corpus, nous n'avons obtenu que 8 résultats (*selfie*, *big data*, *zadiste*, *covoiturer*, *scud*, *nomophobe*<sup>230</sup>, *psychoter*, *écoblachiment*), soit beaucoup moins que le résultat chinois de 21 nouveautés langagières en commun repérées.

---

<sup>229</sup> Nous manquons d'articles qui citent les nouvelles entrées du *Petit Robert* en 2012 et 2014, et pour le *Petit Larousse* en 2012.

<sup>230</sup> Le *Petit Larousse* a sélectionné le terme mais Le *Petit Robert* a choisi *nomophobia*.

En outre, nous avons constaté que les organes de presse présentent parfois dans un même article les nouvelles entrées des deux dictionnaires vu leur relation compétitive dans la promotion de vente. Par exemple, les articles du *Figaro* de l'année 2011, 2016, l'article 2016 du *20 Minutes* et celui 2016 du *RFI*. Il est probable que ces organes de presse choisissent alors de rapporter des nouvelles entrées différentes de ces deux dictionnaires pour éviter les répétitions et introduire plus de nouveauté. De ce fait les éléments communs sont minorés. Nos observations exposées *supra* nous ont appris que le choix des journaux de rapporter de nouvelles entrées est influencé par plusieurs : les types de présentation, les thèmes sélectionnés, la structure de l'article, etc. Même s'il est important pour nous de restreindre le nombre, il faut aussi que nous puissions opérer un choix dans une gamme de possibilités suffisantes.

Nous sommes parvenue à la conclusion que les démarches mises en œuvre pour le corpus chinois ne convenaient pas pour la situation française. Rappelons que notre sélection pour collecter des tweets sur *Sina Weibo* est constituée de 7 termes choisis selon deux critères : la limitation de la production langagière pour chaque année parmi les répétitions relevées ; les caractéristiques marquantes sur le plan formel et communicationnel suite à la comparaison avec l'ensemble des 21 termes. La décision de constituer un échantillon à partir des termes en commun entre les sous-corpus émanait d'une hypothèse selon laquelle des choix en commun à partir de critères et de positionnements variés sont plus significatifs.

La comparaison des sous-corpus français a abouti directement à un nombre de choix équivalent de la liste finale chinoise (respectivement 8 et 7). Mais nous pouvons constater que les 8 résultats français obtenus ne nous permettent pas de faire un choix par année comme pour le corpus chinois puisque la plupart d'entre eux ne sont pas intégrés dans les deux dictionnaires la même année. Par exemple : *nomophobe* est déjà dans *Le Petit Larousse* en 2014 mais il faut attendre jusqu'en 2017 son intégration dans *Le Petit Robert*. En outre, la comparaison entre ces termes nous a montré qu'ils présentent des caractéristiques communes : *selfie*, *big data* peuvent être considérés comme des emprunts à l'anglais (relevant de la matrice externe) ; *zadiste*, *covoiturer*, *psychoter* sont des mots nouvellement fabriqués qui peuvent être analysés selon les matrices morphosémantiques. Si ces traits sont significatifs d'un point de vue social, culturel ou communicationnel, il semble que certaines caractéristiques soient absentes de cette liste limitée. Notons que dans le reste du corpus de listes français, des locutions sont également présentes : *prendre cher*, *partir en cacahuète*, *tendu comme un string*. Est-il pertinent de les laisser de côté ? En reprenant l'échantillon chinois finalement établi pour collecter les tweets,

nous constatons des caractéristiques linguistiques diversifiées comme la forme mixte classée dans la matrice externe (composition intégrant à la fois un mot chinois et anglais) : 打 call (n°7) ; une création à partir d'un accent dialectal : 蓝瘦,香菇 (n°6) ; une expression figée qui se rapproche d'une structure d'une phrase : 元芳...(n°1).

Cela nous montre que la recherche des termes communs aux sous-corpus français ne nous permet pas de prendre en considération des types de nouveautés langagières bien représentées pourtant dans le corpus de listes français et nous ferait aboutir à un échantillon trop simple sur le plan néologique par rapport à l'échantillon chinois. Donc la démarche de sélection pour collecter le corpus de tweets chinois ne peut être appliquée à l'identique pour la constitution du corpus de tweets français.

Afin de sélectionner les termes français pour la constitution du corpus de tweets, nous avons décidé de revenir sur les observations d'analyse linguistique présentées dans le chapitre V pour dégager certains critères. Les situations proportionnelles des matrices montrent une diversité de procédés de création des nouveautés langagières. Aussi, ces créations concernent toutes les matrices de grands niveaux, c'est-à-dire les matrices internes MS, SS, PM et PH et la matrice externe. Cela nous conduit à choisir un terme par matrice pour constituer la liste des termes au lieu d'un terme par année. Sachant que nous n'avons pas pu repérer année par année les articles des organes qui rapportent les nouvelles éditions des dictionnaires. Il est aussi à noter que nous allons faire une observation sur l'ensemble du corpus français et ne pas faire une distinction entre les sous-corpus.

Vient alors la question de comment choisir une petite quantité de termes parmi environ 300 créations ? En nous basant sur les résultats d'analyse des procédés français, nous avons éliminé principalement les cas suivants : les emprunts dans la matrice externe et les noms propres de personne en supposant que leurs évolutions sur le plan morphologique et/ou sémantique sont moins fluctuants. Rappelons que Sablayrolles ne considère que les influences directes des langues étrangères comme emprunt. Il est intéressant d'étudier l'intégration des lexiques étrangers dans la langue mais il ne s'agit pas de notre objectif ici. Quant aux noms propres qui servent à désigner des entités en particulier, ils assurent une continuité de référence et leurs sens sont relativement invariants. Or, l'objectif de l'observation du corpus de tweets vise surtout plutôt à une éventuelle évolution du sens. Comme les deux cas (emprunts et noms propres) occupent un grand pourcentage dans le corpus de listes et leurs éliminations peuvent beaucoup

diminuer les termes candidats et faciliter la sélection finale. En prenant compte des éléments sociaux, politiques et culturels saillants ainsi que nos expériences outre les caractéristiques linguistiques diversifiées, nous sommes parvenue à un échantillon de 7 termes que nous présentons sommairement dans le tableau suivant (Tableau 30) en les présumant beaucoup plus aisés à comprendre que les termes chinois (dans une thèse d'université française). Ce tableau comporte trois colonnes. De droite à gauche on trouve le numéro du terme ; le terme lui-même suivi de la matrice dont relève son procédé de création entre crochets et de la date de dictionnarisation ; une définition succincte émanant d'une ressource en ligne<sup>231</sup>. Nous n'avons pas précisé dans ce tableau la contextualisation mais nous aurons l'occasion d'y revenir dans les analyses consacrées au corpus de tweets dans le chapitre suivant (chapitre VII).

---

<sup>231</sup> Nous avons choisi l'édition en ligne : <https://www.larousse.fr/>, accessible en 2021. Il s'agit dans un premier temps d'indiquer le sens et non de faire une comparaison lexicographique pour elle-même.

N°	Termes +[procédé de création]	Définition sommaire
1	<i>tuerie</i> (n.f.) [SS]  2015 Petit Larousse	Action de <u>tuer</u> en masse ; carnage. Action de tuer les animaux de boucherie. Lieu où sont abattus les animaux de boucherie. <i>Familier</i> <b>1.</b> Mets ou breuvage délectable ; <u>délice</u> : Cet éclair au chocolat est une tuerie. <b>2.</b> Ce qu'on juge magnifique, remarquable, et qui suscite l'engouement : Ces chaussures en vitrine sont une tuerie ! Le concert d'hier soir ? Une tuerie !
2	<i>marche blanche</i> (expression) [MS-CC]  2014 Petit Larousse	Marche blanche [blanc] marche collective silencieuse organisée à la suite d'un fait divers tragique.
3	<i>nomophobe</i> (n. et adj.)  2014 Petit Larousse  <i>nomophobie</i> (n.f.) 2017 Petit Robert [PM]	Se dit de quelqu'un qui ne peut se passer de son téléphone portable et éprouve une peur excessive à l'idée d'en être séparé ou de ne pouvoir s'en servir.
4	<i>zadiste</i> (n. et adj.) [MA-CA]  2016 Petit Robert 2017 Petit Larousse	Militant qui s'oppose à l'aménagement d'une <u>ZAD</u> (zone à défendre). ZAD : Espace, le plus souvent rural, occupé par des militants s'opposant à un projet d'aménagement qu'ils estiment inutile, coûteux et susceptible de porter atteinte à l'environnement et à l'intérêt des populations locales. (Cet emploi du terme constitue un détournement de sens de l'acronyme ZAD [zone d'aménagement différé].
5	<i>tendu comme un string</i> (expression) [PH]  2016 Petit Robert	Métaphore du registre familier pour exprimer l'angoisse, le stress (le string étant une sorte de sous-vêtement avec un cordon) (humoristique)
6	<i>vivre ensemble</i> (n.m. inv.) [SS]  2017 Petit Robert	Cohabitation harmonieuse entre individus ou entre communautés : La fête des voisins se veut une initiative propre à favoriser le vivre-ensemble.
7	<i>loup solitaire</i> (expression) [MS-CC]  2017 Petit Larousse	Personne qui, par tempérament ou par goût, travaille, agit ou vit seule ; en particulier, terroriste qui, au nom d'une idéologie, perpétue seul des actes violents (attentat, prise d'otage, par exemple). [La notion est controversée dans la mesure où il s'avère, le plus souvent, que l'auteur de ces actes avait un lien avec une organisation terroriste.]

**Tableau 30 : Echantillon de nouveautés langagières choisi pour la collecte de tweets sur *Twitter***

Précisons rapidement l'application de nos critères. Tout d'abord, dans les matrices MS, nous avons sélectionné le terme *zadiste* formé par affixation ; les termes *marche blanche* et *loup solitaire* construits par composition. Dans les matrices SS, nous avons choisi les termes *tuerie* et *vivre ensemble*. Nous nous intéressons ensuite au terme *nomophobe* dans les matrices PM et la création d'expression *tendu comme un string* dans les matrices PH. Le choix des termes *zadiste*, *marche blanche*, *loup solitaire* et *vivre ensemble* est également motivé par leurs contextes sociaux d'émergence en privilégiant la dimension socio-politique. Cela s'accorde avec le niveau élevé de politisation qui caractérise les usagers de *Twitter* comme nous l'avons indiqué *supra*. Nous avons privilégié *tuerie* pour les matrices SS, notamment la matrice « changement de sens » parce que le saut de signification entre son sens d'origine et son sens néologique est remarquable. En tant que locutrice étrangère, nous sommes immédiatement sensibilisée à ce sens figuré (« c'est formidable ») qui nous paraît étrange par rapport à notre connaissance du sens de massacre. Le terme *nomophobe* est spécifique pour nous parce qu'il s'agit d'une pratique étroitement liée au contexte numérique, ce qui coïncide avec nos intérêts de recherche sur les changements langagiers et leur circulation sous l'influence du numérique. Outre ces nouveaux mots simples et composés, les nouvelles créations d'expression ne restent pas inaperçues dans les procédés de création. Donc, nous avons aussi choisi de prendre en considération ce procédé et sélectionné le terme *tendu comme un string*. En plus, cette expression est la plus utilisée avec hashtag parmi les expressions regroupées dans les matrices PH selon nos recherches sur *Twitter*. La collecte d'une quantité de tweets suffisante constitue également une des raisons importantes du choix de ce terme.

Maintenant que nous avons déterminé l'échantillon qui nous permettra d'opérer la collecte de tweets, venons-en à cette collecte proprement dite en rappelant nos principes appliqués pour la faire sur *Sina Weibo*. Pour constituer le corpus de tweets chinois, nous avons choisi d'observer les nouveautés langagières sous forme de hashtags après avoir mené une observation de la circulation des nouveautés langagières sur le dispositif *Sina Weibo*. Cette observation nous a permis de confirmer dans un premier temps notre hypothèse sur leur vitalité et de constater de nombreuses utilisations durant une période indéfinie (en raison de la contrainte technique d'accès aux données intégrales à partir du moteur de recherche). Nous avons ainsi mesuré combien la détermination d'une quantité et d'une période de sélection reste un grand défi.

À ces difficultés s'ajoutent aussi les spécificités de l'environnement techno-relationnel du microblog et les opérations possibles pour un observateur-usager. Nous nous sommes

finalement appuyée sur la fonction de groupement de hashtags et son fil du temps pour collecter notre échantillon de tweets car cela nous permet d'accéder à une ressource relativement restreinte mais ouverte à tous les usagers du microblog. En plus, le choix d'observer les hashtags associés aux nouveautés langagières peut également faciliter la collecte du corpus. C'est-à-dire, en nous basant sur la fonctionnalité relationnelle d'un hashtag, nous pouvons déterminer et présenter un affichage simple d'un tweet comme dans la Figure 32. Cela nous permet de nous concentrer sur le corps du message d'un tweet (où les éléments textuels et contextuels sont présents), et ne pas prendre en considération les autres interactions qui nous posent problème pour décider de leur portée et leur présentation par capture d'écran.

Par la suite, nous avons continué à filtrer le champ de collecte, c'est-à-dire que nous nous sommes intéressée aux résultats de tweets affichés sous la rubrique « à la une ». La popularité de la rubrique « à la une » implique que le taux d'engagement des tweets est élevé. Donc, nous considérons que ces tweets pré-filtrés par popularité sont plus visualisés sur la plateforme et touchent un public élargi. Nous supposons en outre que les utilisations des nouveautés langagières sont dans certaine mesure reconnues par le public touché. En même temps, les tweets populaires ne sont pas classés dans un ordre chronologique. De ce fait, nous n'avons plus besoin d'établir un critère de temps pour la sélection. Il est difficile de fixer une période de sélection, vu la dynamique de production et les pratiques imprévues des usagers, mais nous avons essayé de sélectionner des tweets produits dans des années différentes pour prendre en considération leur évolution dans le temps. Comme *Sina Weibo* et *Twitter* disposent tous les deux de cette rubrique de popularité, le choix d'un lieu de collecte similaire vise à favoriser une analyse comparative ultérieure dans les deux langues et les deux contextes socio-culturels.

Dans l'objectif de réaliser une analyse significative et non exhaustive, nous avons fixé aléatoirement la collecte à 10 tweets pour chaque nouveauté langagière, exactement comme pour le corpus chinois. La démarche de collecte des tweets proprement dite est donc identique à un détail près, qui tient aux différences d'affichage de résultats entre les deux plateformes, et nous amène à effectuer des démarches complémentaires pour la collecte sur *Twitter*.

Ces filtres sont établis principalement autour de la date de publication. Nous cherchons à collecter dans la rubrique caractérisée par la popularité (« à la une ») les publications de tweets d'années différentes. Nous avons limité notre période d'observation de 2010 à 2020. D'une part, cette période englobe les années d'intégration des nouveautés langagières élues. D'autre part,



l'année 2010 marque le plein essor des réseaux sociaux numériques y compris les microblogs. Dès lors il est intéressant d'observer les emplois des nouveautés langagières avec une croissance importante des utilisateurs. Le critère du temps d'observation jusqu'en 2020 permet d'évaluer la vitalité de ces nouveautés langagières sur *Twitter*, surtout pour observer les utilisations des nouvelles entrées dans les éditions relativement récente comme les éditions 2016 et 2017. Dans ce cadre temporel, nous avons besoin d'appliquer respectivement un filtre de recherche pour chaque année<sup>232</sup>. Ce critère du temps a pour but de faciliter l'accès à un résultat de recherche. Il ne modifie toutefois pas l'importance du critère de popularité.

Suivant les démarches ci-mentionnées, nous avons collecté pour chaque nouveauté langagière française 10 tweets et au total 70 tweets (voir l'annexe n°5) pour constituer le corpus français.

## **Conclusion**

Les caractéristiques techno-relationnelles du microblog et la dimension discursive de l'analyse visée nous ont amenée à décider de présenter les tweets de nos corpus par capture d'écran et de prendre en considération la portée des affichages et des types de tweet dans la constitution effective du corpus. En même temps, la quantité importante de publication, l'imprévisibilité de la pratique des usagers de microblog ainsi que les points communs et les nuances techniques des deux plateformes nous ont conduite à entreprendre une analyse qualitative et à observer les tweets populaires contenant les termes issus de notre corpus de liste sous la forme hashtaguée.

Cette décision d'analyse qualitative nous a également menée à établir préalablement un échantillon des nouveautés langagières significatives, et à collecter pour chaque nouveauté langagière sélectionnée 10 tweets à analyser. Les observations sur les corpus de liste dans les deux langues ont montré que bien que les mêmes critères (popularité et fréquence d'utilisation) soient appliqués lors de la constitution de ces corpus, les démarches mises en œuvre pour sélectionner les termes chinois ne convenaient pas pour la situation française. Il en est de même pour la constitution du corpus de tweets. Au lieu de chercher les termes en commun des sous-corpus comme pour le corpus chinois, nous nous sommes plutôt appuyée sur les principaux procédés de création pour choisir les nouveautés langagières françaises, tout en prenant compte les spécificités socio-politique et culturel. Nous avons ainsi choisi respectivement 7 termes en

---

<sup>232</sup> Durant l'application des filtres, nous avons également observé des phénomènes d'utilisation spécifiques pour le corpus de tweets français. Nous allons détailler ces points dans la partie d'analyse suivante au chapitre VII.

chinois et en français et collecté 70 tweets dans chaque langue pour l'analyse discursive. Celle-ci fait l'objet du chapitre suivant.



## CHAPITRE VII

### Analyse discursive de deux corpus de tweets

#### Introduction

Nous avons détaillé dans le chapitre précédent les critères et les étapes qui ont abouti à la constitution de deux corpus de tweets. Ceux-ci sont composés de hashtags associés<sup>233</sup> aux nouveautés langagières sélectionnées. Ces corpus devraient nous permettre d'observer sur échantillon les circulations et les évolutions de sens des nouveautés langagières dans un contexte techno-relationnel. Nous nous inscrivons pour ce faire dans une approche écologique, discursive et communicationnelle.

Nous allons nous intéresser dans un premier temps aux rapports qu'entretient le hashtag avec les nouveautés langagières associées d'une part et avec les éléments non hashtagués d'un tweet d'autre part (corps du message et métadonnées d'un tweet). Nous allons aussi aborder les pratiques des énonciateurs en nous basant sur leurs identités sur les plateformes de microblog et leurs objectifs de production.

Comme précédemment nous traiterons d'abord le corpus chinois puis le corpus français avant de proposer une comparaison pour conclure tout en faisant des mises en relations entre les deux corpus quand nous le penserons utile.

#### 7.1. Analyse discursive d'un corpus de tweets chinois

Nous rappelons que nous avons collecté 10 tweets pour chacun des 7 termes retenus. Pour faciliter la démarche d'analyse, nous avons numéroté les 7 termes selon leurs années de création. Le terme n°1 est celui créé en 2012 元芳，你怎么看 (yuán fāng, nǐ zěn me kàn), et ainsi de suite jusqu'au 7<sup>e</sup> terme dont la création date de 2017 打 call (dǎ call). Les 10 tweets pour chaque terme sont numérotés comme suit : n°1.1 pour le 1<sup>er</sup> tweet du terme n°1, n°1.2 pour le

---

<sup>233</sup> Sur *Sina Weibo*, les résultats d'une recherche sont marqués en couleur rouge tandis que sur *Twitter*, les résultats sont marqués en gras avec la couleur originelle du hashtag en bleu.

2° et ainsi de suite... n°2.1 pour le 1<sup>er</sup> tweet du terme n°2 et ainsi de suite.

### 7.1.1. Discussion sur la catégorisation des hashtags selon leur emplacement

La distinction entre hashtags « externe » et « interne » est fréquemment utilisée. Dans le carnet de recherche destiné à définir la notion de « hashtag », Paveau (2013) indique que si un hashtag précède ou suit un texte de tweet, il peut être considéré comme « externe », alors que, s'il est intégré syntaxiquement dans l'énoncé, souvent comme substantif ou adjectif, c'est un hashtag interne. En nous basant sur notre corpus, nous avons constaté que la frontière entre les hashtags externes et internes est floue si nous adoptons le principe de distinction ci-dessus mentionné, surtout pour repérer le hashtag externe. Nous ne pouvons pas classer un hashtag comme externe s'il est placé au début ou à la fin du contenu verbal d'un tweet.

Par exemple, dans les tweets n°3.5<sup>234</sup> (Figure 39) et n°5.3<sup>235</sup>, les hashtags sont placés en position initiale mais ils constituent également une partie indispensable du corps de message. Il en va de même pour les hashtags qui se situent en position finale dans les tweets n° 2.3 qui parle de « femme masculine » (Figure 40), n° 2.6, n° 7.6 et n° 7.8.



Figure 39 : Capture d'écran du tweet n°3.5 avec hashtag en position initiale

<sup>234</sup> La nouveauté langagière n°3 signifie « grandiose, classe ».

<sup>235</sup> La nouveauté langagière n°5 signifie « faire tout ce qu'on peut pour réaliser quelque chose ».



Figure 40 : Capture d'écran du tweet n°2.3 avec hashtag en position finale

L'analyse des cas concrets du corpus nous a convaincue que cette distinction n'était pas vraiment opératoire. Voyons deux autres cas particuliers :

Le terme n°1 est une longue expression figée créée à partir d'une phrase indépendante. Cette phrase interrogative est originellement mise à la fin d'une conversation pour exprimer un doute ou inviter les auditeurs à exprimer leurs opinions. Autrement dit, elle fait partie intégrante de la syntaxe et a même une complétude syntaxique. Compte tenu de ses spécificités formelle et sémantique, quand le hashtag associé à ce terme n°1 est placé à la fin du texte du tweet, il est difficile de savoir s'il est intégré dans le corps du message ou non. Les tweets n°1.3, n°1.5 et n°1.6 en sont des exemples.

Le contenu du tweet n°1.5 ci-dessous (Figure 41) rapporte l'affaire suivante : une agence immobilière ne verse pas les loyers déjà payés par les locataires aux propriétaires et cède ensuite son droit de gestion, niant toute responsabilité ; faute de moyens et pour diminuer leurs pertes, les propriétaires ont changé les serrures, laissant les locataires sans logis. L'auteur du tweet qui est le rédacteur en chef d'un journal, demande, après avoir relaté cette affaire si l'agence doit assumer une responsabilité légale. Le terme n°1 figure à la fin du tweet, posant la question « Yuan fang, qu'en penses-tu ? ». Syntaxiquement, ce peut être une question complémentaire demandant l'avis des lecteurs sur le jugement de l'auteur du tweet. Mais la question peut aussi être considérée comme indépendante pour demander l'opinion des internautes sur le contenu global, sur l'ensemble de la situation, auquel cas il s'agirait d'un hashtag « externe ».

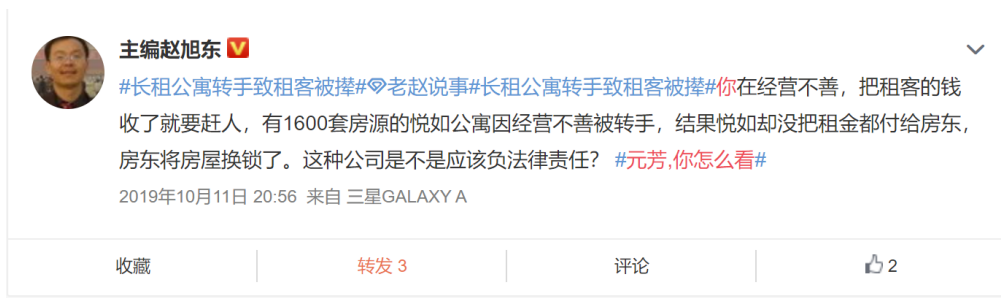


Figure 41 : Capture d'écran du tweet n°1.5 avec hashtag en position finale

Un autre cas concerne le nombre de hashtags dans le tweet. Le corps du message peut être constitué seulement par un ou plusieurs hashtags. En effet, la technologie du tweet permet d'inclure tous les éléments dans un hashtag dans la limite de 140 caractères (avant 2016). Dans le texte des tweets n°1.2, n°1.9, il n'existe qu'un seul hashtag. Un hashtag ne peut pas être considéré comme interne ou externe quand il n'y a aucun élément non-hashtagué dans le tweet.



Figure 42 : Capture d'écran du tweet n°1.2 avec le texte du tweet constitué par un hashtag

En nous basant sur ces premières observations, nous nous sommes rendu compte qu'il ne convenait pas d'analyser les rapports entre le hashtag et les autres éléments du tweet par la (seule) distinction de son emplacement. Nous allons à l'étape suivante, nous focaliser sur les fonctions des hashtags dans leurs occurrences d'utilisation.

### 7.1.2 Fonctions des hashtags

Les fonctions des hashtags que nous avons relevées sont diverses. La plateforme du microblog offre dans un premier temps *une fonction basique de nature techno-relationnelle*. Il s'agit de la possibilité de regrouper d'une manière automatique tous les tweets contenant les mêmes hashtags et de les rendre ainsi investigables (Paveau 2017:203). C'est aussi à travers cette fonction que nous pouvons collecter notre corpus. L'aspect relationnel est manifesté par l'offre d'une interaction ouverte aux usagers sans besoin de construire préalablement des relations à travers l'abonnement<sup>236</sup>. Par l'insertion d'un hashtag, l'auteur du tweet le peut joindre une audience indéfinie et plus large que ses relations construites. Le hashtag crée de cette façon un réseau relationnel dynamique. Dès lors nous pouvons dire que l'insertion du hashtag dans un tweet est une action intentionnelle du producteur se basant sur les ressources techniques du microblog.

Selon nos observations et sur des bases à la fois sémantiques, pragmatiques et communicationnelles, nous proposons de distinguer trois catégories de hashtags : le hashtag principal par rapport au contenu thématique, le hashtag auxiliaire par rapport à ce même contenu, le hashtag promotionnel. Nous donnerons un ou des exemples de chaque catégorie.

#### *Hashtag principal*

Il s'agit souvent de hashtags clés quant aux thématiques de tweets dont le contenu est étroitement lié à chacune des nouveautés langagières sur le plan sémantique et/ou contextuel. Dans ces cas-là, le hashtag sert principalement à résumer ou mettre en évidence le reste du contenu. Le hashtag principal peut être le seul hashtag dans un tweet ou coexister avec les autres hashtags.

Par exemple, dans le tweet n°2.4 (Figure 43), le hashtag « femme masculine » (女汉子 nǚ hàn zi, *litt.* « féminin mec») a pour objectif de récapituler le contenu textuel qui suit et qui valorise l'idée que les femmes sont capables d'accomplir des actions qu'on considèrerait comme réservées aux hommes. Ce tweet est destiné à présenter les réussites d'une sportive de Nouvelle Zélande qui a établi un record du monde dans le domaine du lancer du poids ainsi que des records de

---

<sup>236</sup> Fonction sur microblog permettant de suivre les autres comptes du microblog.



compétition dans des pays différents. La photo mise en attachement est le portrait de cette sportive. Elle peut d'une part fournir des informations supplémentaires et d'autre part, présenter l'image d'une « femme masculine » d'une façon directe et intuitive. Le texte et la photo constituent un éloge de femmes qui peuvent aussi être excellentes que les hommes dans une activité physique demandant de la force.



Figure 43 : Capture d'écran du tweet n°2.4

Le tweet n°4.1 (Figure 44) que nous allons voir maintenant présente un article hébergé par une plateforme tierce portant sur le sujet des « critères pour juger la beauté ». Le tweet commence par le titre de l'article. On peut accéder à l'article entier en cliquant sur l'hyperlien qui suit. Le reste du contenu est le début de l'article. Le hashtag « c'est le tempérament qui compte » (主要看气质 zhǔ yào kàn qì zhi) en position finale est employé pour résumer l'idée essentielle de cet article.



**Figure 44 : Capture d'écran du tweet n°4.1**

Dans le tweet n°3.5 (Figure 39 présentée précédemment), le hashtag #高大上# (haut de gamme, grandiose, classe) est souligné pour être en contraste avec la photo qui l'accompagne. Ce hashtag a une valeur adjectivale renvoyant au luxe. Au sens littéral, le tweet parle donc d'un mode de vie de haute qualité. Or le thème principal n'est pas celui du prestige mais il s'agit de plaisanter sur l'emblème d'une marque de luxe. La marque de voiture Mercedes-Benz est en effet peinte à la main sur un vieux tricycle très banal qui est souvent destiné à un usage quotidien de collecte et recyclage de déchets par des individus. Avec cette marque, ce tricycle est transformé en « voiture de luxe ». Le propriétaire du tricycle bénéficie dans ce cas-là d'une vie « haut de gamme ». Le contenu du hashtag a pour but de créer un effet humoristique par contraste, sur la base d'une ironie antiphrastique.

### ***Hashtag auxiliaire***

Un hashtag auxiliaire est lié à certains éléments constituant un tweet. Donc, il est attaché à un contenu partiel d'un tweet, ce qui est différent du hashtag principal consacré au sujet central d'un tweet. Il est souvent utilisé pour une fonction descriptive. Voici quelques exemples :

Dans le tweet n°3.2 (Figure 45), il y a deux hashtags : #葡萄酒# « le vin » et le terme n°3 #高大上# (« haut de gamme, grandiose, classe ») qui signifie « luxueux » ici. Le contenu du tweet concerne la promotion d'un itinéraire d'œnotourisme. Les informations de localisation sont fournies au début du tweet. Les photos montrent différents paysages, y compris le vignoble et la cave, ainsi que plusieurs productions de vin. Le tweet est donc dédié au thème du vin et du

tourisme. Cependant, le hashtag contenant la nouveauté langagière est utilisé pour qualifier la destination et plus généralement ce genre d'activité touristique.



Figure 45 : Capture d'écran du tweet n°3.2

Il est à noter que le rôle auxiliaire du hashtag est déterminé par sa relation avec les autres éléments du tweet et non par la quantité des hashtags. Prenons le cas du tweet n°3.4. Il s'agit d'un retweet avec commentaire. Il est composé d'un hashtag, élément textuel, d'émoticônes, des mentions de trois autres comptes et du contenu retweeté. L'ensemble est destiné à promouvoir le Forum des Jeunes de l'UNESCO et à faire appel à la participation. Le seul hashtag présent dans le commentaire est employé comme adjectif pour décrire et qualifier cette activité et non pour remplir une fonction de résumé thématique.



Figure 46 : Capture d'écran du tweet n°3.4

Un autre tweet, le n°4.2 est constitué par un message textuel suivi de trois hashtags juxtaposés à la fin. L'auteur vise à exprimer que le temps peut transformer une fille en une femme, c'est-à-dire que le temps cause le changement de l'apparence physique, alors que la lecture peut permettre à une femme d'accroître ses connaissances et de devenir de ce fait très riche spirituellement. À première vue, il nous est difficile de connaître le contexte du tweet. C'est à travers le 3<sup>e</sup> hashtag #婚姻的故事# « l'histoire de mariage » que nous parvenons à le situer. Les deux autres hashtags, 1<sup>er</sup> #读书# « lecture » et le 2<sup>e</sup> #主要看气质# « c'est le tempérament qui compte », qui est la nouveauté langagière dans le tweet, composent deux éléments importants pour les femmes, selon l'auteur, dans le mariage. Ces deux hashtags jouent donc un rôle auxiliaire dans ce tweet.



Figure 47 : Capture d'écran du tweet n°4.2

Un autre exemple particulier est le tweet n°1.10. Cette particularité est due à deux raisons. Premièrement, il s'agit d'un tweet qui rapporte les résultats de la sélection des nouveautés langagières du magazine « yǎo wén jiào zì » en 2012. Cette liste de résultats fait partie de notre corpus de liste A. Donc, la présence du terme n°1 hashtagué est un élément de discours rapporté puisque le résultat est cité et non un réemploi. Deuxièmement, la quantité des hashtags est de 11 hashtags qui peuvent être classés en deux catégories : le premier hashtag « le top 10 des expressions populaires de 2012 » est le hashtag principal du tweet qui donne un aperçu rapide du contenu, tandis que les 10 expressions de l'année 2012 sont englobées par les signes # et constituent les hashtags suivants. Nous supposons que ces hashtags servent à souligner chaque nouveauté langagière et leur permettent en même temps de joindre leurs propres fils de hashtag. Nous avons également constaté le même type de tweet et l'application de hashtags pour les autres nouveautés langagières.



Figure 48 : Capture d'écran du tweet n°1.10

### *Hashtag de promotion*

Une troisième catégorie de hashtag destinée à augmenter la visibilité d'un tweet peut être dite « de promotion » ou « algorithmique ». D'une part, comme nous l'avons mentionné lorsque nous avons expliqué les démarches de constitution de notre corpus de tweet, l'insertion de hashtags et d'autres éléments multi-sémiotiques dans un tweet est une méthode pour augmenter la visibilité requise par le système. D'autre part, nous avons constaté que les twitteurs s'appuient également sur les tendances populaires pour pouvoir exposer leurs tweets à une audience importante à un moment donné. Cela est souvent constatés dans la pratique publicitaire. Nous avons relevé plusieurs cas qui illustrent ce phénomène.

Par exemple, le tweet n°2.10 est publié par un compte destiné à promouvoir le film chinois *The Rise of a Tomboy*<sup>237</sup> et qui porte le nom du film. Sa publication intervient au moment de l'annonce des résultats de Gaokao<sup>238</sup>, un événement national important en Chine. Le contenu consiste en un message de félicitation (au nom de l'héroïne du film) pour les candidats. Nous pouvons trouver deux hashtags accompagnant le texte : l'un en position initiale est le terme n°2 « femme masculine», l'autre qui suit le texte est le hashtag suivant : « les résultats de Gaokao 2015 ». Le personnage sur la photo est l'héroïne du film. Les quatre caractères en bas de la photo correspondent à un message de souhait court portant sur le même contenu. Apparemment, la nouveauté langagière a peu de relation avec le contenu du message. Mais, à travers le film, nous pouvons savoir que l'héroïne est une étudiante brillante qui croit tout pouvoir résoudre par des formules mathématiques, y compris trouver son amour. La nouveauté langagière « femme masculine » renvoie dans ce film au personnage d'une femme intelligente mais excessivement rationnelle, ce qui bouleverse les stéréotypes sur les femmes. De ce fait, l'auteur essaie de se rapprocher des candidats en soulignant le caractère fort du personnage de film et en profitant de l'attention accordée à la nouvelle expression et à l'événement pour augmenter la visibilité du tweet. Les candidats au concours sont aussi le public ciblé du film.

---

<sup>237</sup> *Tomboy* est la traduction anglaise du terme n°2 « femme masculine ». Ce film est sorti en 2016. Il raconte la poursuite d'amour d'une étudiante par une formule mathématique issue d'une recherche scientifique.

<sup>238</sup> Le Concours national d'entrée de l'éducation supérieure est un concours sur plusieurs jours, à la fin de l'enseignement secondaire du lycée, qui sert d'examen d'entrée pour les études supérieures en Chine.



Figure 49 : Capture d'écran du tweet n°2.10

Au cours de la constitution du corpus, nous avons aussi constaté des phénomènes de publicité virale s'appuyant sur le hashtag pour augmenter la visibilité des tweets. Prenons l'exemple du tweet n°5.4. Nous le considérons comme une publicité virale sur la base des observations suivantes :

1) Sur 10 pages de résultat, ce tweet est produit par des comptes variés dont les dates de productions sont très proches, ce qui ressemble au phénomène de publication par des comptes de robot.

2) Le contenu du tweet comporte des éléments favorisant l'augmentation du taux d'engagement, ce qui constitue un critère important pour la visibilité d'un tweet. A part le hashtag et l'image, nous avons également constaté des interactions fabriquées entre les comptes de robots sous forme de mention @. Le contenu de ce tweet consiste à introduire une application de livre audio dont l'icône est ajoutée en attachement. Pour profiter de la popularité de ce hashtag, le texte de présentation est spécialement créé pour rapprocher le contexte de création de la nouveauté langagière 洪荒之力 (hóng huāng zhī lì, *litt.* «temps immémorial pouvoir») qui signifie initialement « le pouvoir du temps immémorial » et a acquis le sens de « faire tout ce qu'on peut pour réaliser quelque chose » après avoir été utilisé par une sportive de natation lors de Jeux Olympiques de Rio en 2016.

Dans le texte de ce tweet, on interprète 洪荒之力 (hóng huāng zhī lì) comme une compétence à progresser quotidiennement. Pour cela, on compare cette compétence à l'apprentissage de la

natation qui se développe à travers des exercices réguliers. Par cette analogie, l'emploi continu de cette application de livre audio peut aider le public des internautes à enrichir ses connaissances, ce qui constituerait la compétence dont il est question dans ce tweet.



**Figure 50 : Capture d'écran du tweet n°5.4**

Dans les deux tweets promotionnels que nous venons de voir, on essaie toujours de créer un lien entre la nouveauté langagière et la production à promouvoir.

Il existe aussi une utilisation du hashtag purement algorithmique dans les actes de promotion. Ainsi le tweet n°4.10 contient la nouveauté langagière 主要看气质 (zhǔ yào kàn qì zhi, « c'est le tempérament qui compte »). Celle-ci critique la pratique stéréotypée qui consiste à juger quelque chose ou quelqu'un d'après l'apparence. Le contenu du tweet est la présentation d'un produit financier, une plateforme de prêt entre particuliers. On parle de ses caractéristiques (une haute rentabilité annuelle) et des avantages offerts aux usagers qui veulent investir (l'opportunité est ouverte pour un temps limité à trois mois). Un hyperlien ajouté à la fin sert à accéder au site extérieur du produit. L'image qui accompagne le message a pour objectif d'afficher le slogan de ce produit : plus rapide sera la participation, plus nombreux ceux qui s'investissent, plus avantageux sera le bénéfice, il ne faut pas hésiter ! L'image est donc directement liée au contenu du tweet. La nouveauté langagière figure comme hashtag en tête et est ensuite implantée dans la première phrase du tweet : il vaut mieux prendre en considération le « tempérament » d'une plateforme de prêt pour le choix d'investissement. Nous pouvons donc considérer les fonctions du produit comme son caractère mais le contraste d'ici entre



l'apparence et le tempérament est forcé. Le tweet ne nous fournit pas d'élément explicite ou implicite faisant référence à l'apparence. Donc, nous considérons que le contenu multi-sémiotique du tweet n'a rien à voir avec la nouveauté langagière. Il s'agit seulement de profiter de la nouveauté et de la popularité du segment hashtagué pour une promotion.

Nous avons aussi constaté une autre forme de pratique pour le même objectif ; le tweet n°3.10 (Figure 51, position inférieure) en est un exemple. Une des spécificités de ce tweet est la quantité (au nombre de 23) de hashtags comme si on pouvait dire : plus de hashtag, plus de visibilité. Ce tweet cherche à promouvoir un article portant sur les relations entre les habits et les caractères des femmes. Après avoir consulté cet article en cliquant sur l'hyperlien inséré, nous ne pouvons toujours pas relier sémantiquement le texte du tweet avec la nouveauté langagière décrivant quelque chose de luxueux.



Figure 51 : Capture d'écran des tweets n°4.10 et n°3.10

### 7.1.3. Les relations entre hashtags et nouveautés langagières

#### *Mise en relief des nouveautés langagières*

Les catégories de hashtags que nous avons relevés dépendent des pratiques des producteurs. Qu'il s'agisse d'une expression personnelle ou d'une communication commerciale, nous supposons que les hashtags associés à chaque nouveauté langagière dans notre corpus de tweets ont également une valeur réflexive marquant le caractère neuf du segment hashtagué. Cela ne signifie pas forcément que tous les producteurs ont une bonne connaissance des nouveautés langagières. Pour certaines pratiques publicitaires, la maîtrise du segment hashtagué n'est pas obligatoire (Voir tweets n° 4.10 et 3.10).

Nous estimons que le choix de créer un hashtag associé à une nouveauté langagière dans un tweet montre que les producteurs ont conscience de l'apparition et de l'application des nouveaux phénomènes langagiers. Les pratiques de définition en sont des exemples : à travers le tweet n°2.8, l'auteur donne son point de vue sur la définition de « femme masculine » : ce n'est pas la capacité des femmes à faire des tâches réservées habituellement aux hommes mais un caractère distingué qui rend les femmes indépendantes quelles que soient les situations rencontrées. Dans cet exemple, nous pouvons constater que l'auteur a pris connaissance de cette nouveauté langagière avant de publier le tweet.

Le début du tweet n°6.5 indique un phénomène observé : il y a pas mal de gens qui cherchent à comprendre l'expression 蓝瘦, 香菇 (lán shòu, xiāng gū, être trop triste et avoir envie de pleurer). La suite du tweet concerne l'explication de cette nouveauté langagière en combinant la promotion d'un produit attaché à la compagnie du propriétaire du compte de tweet. Ce produit est une carte fidélité offrant divers avantages (nourriture, loisirs, transports, logement, etc.) aux membres : ceux qui n'ont pas la carte, ne peuvent pas bénéficier ces avantages, et, pour cela, seraient tristes voire pleureraient. La publication de ce tweet et la création du hashtag sont basées sur l'observation du phénomène langagier et la connaissance du sens de l'expression.

Nous avons également rencontré des tweets qui s'interrogent sur le sens et la pertinence des nouveautés langagières. Voici un exemple en dehors de notre corpus d'étude<sup>239</sup>. Dans un tweet

---

<sup>239</sup> Pour constituer l'échantillon de notre corpus d'étude, nous avons travaillé sur un corpus de référence plus vaste dont fait partie cet exemple.

publié en 2014, l’auteur énumère trois nouveautés langagières<sup>240</sup> (dont le terme n°1 de notre corpus) dont le sens n’est pas clair pour lui/elle. Nous avons constaté que même si l’auteur n’a pas bien saisi le sens, on suit toujours la situation de l’évolution de certaines expressions. L’auteur souligne dans le tweet que le premier terme énuméré 雷 (léi, *litt.* « tonnerre») n’est plus à la mode et que l’expression populaire actuelle est le 3<sup>e</sup> terme 我也是醉了 (wǒ yě shì zuì le, *litt.* « je suis ivre »).

Pourtant, parmi les tweets hashtagués de notre corpus restreint, nous n’avons pas rencontré des questions sur le sens de nouveautés langagières. Cela peut impliquer que les producteurs (à part les comptes de robots évidemment) ont déjà intégré une connaissance des nouveautés langagières employées.

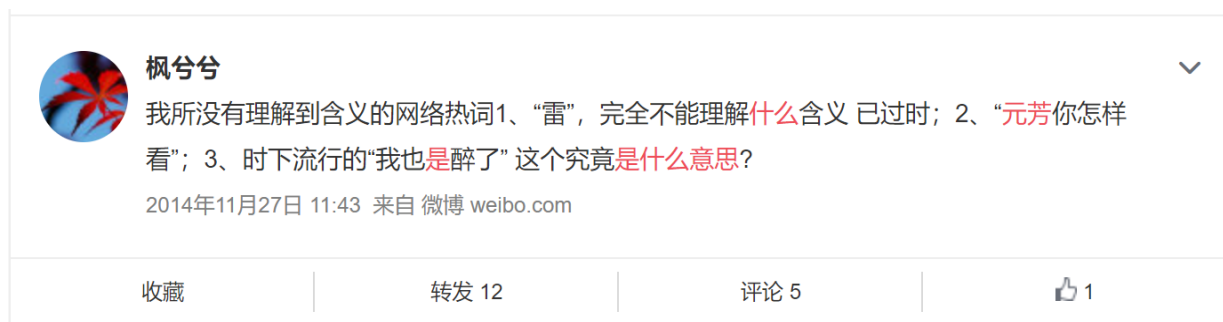


Figure 52 : Capture d’écran d’un tweet de 2014 qui s’interroge sur le sens des nouveautés langagières

Or les usagers de *Sina Weibo* ne sont pas tous en faveur de la création et/ou de l’utilisation des nouveautés langagières. Prenons le tweet n°6.9. L’auteur constate la popularité augmentée de l’expression 蓝瘦，香菇 (lán shòu, xiāng gū, être trop triste et avoir envie de pleurer) mais exprime par la suite sa désapprobation à l’égard de cette expression ainsi que des autres phénomènes langagiers qui sont éphémères. Il trouve que ce genre de « création » ne constitue pas un changement sémantique<sup>241</sup> et ne comprend pas pourquoi le public aime suivre ces tendances langagières. Il conseille à la fin du tweet de ne pas participer à l’application et la

<sup>240</sup>

1.雷 (léi, *litt.* « tonnerre»), Il sert à décrire l’étonnement dû à l’aspect négatif de quelque chose. Il devient populaire autour de l’année 2009.

2.元芳，你怎么看 (yuán fāng, nǐ zěnmeyàng kàn) : Yuan fang, qu’en penses-tu ?

3.我也是醉了 (wǒ yě shì zuì le, *litt.* « je suis ivre ») : sentiment de ne rien y pouvoir.

<sup>241</sup> Nous supposons que l’auteur ne considère pas la paronymie mise en jeu ici comme un procédé de création qui fait naître une nouvelle expression.

circulation de ces nouveautés langagières.

Le groupement de tweets sous un même hashtag favorise d'ailleurs la constitution d'une communauté autour de la pratique d'une nouveauté langagière donnée. Ainsi, la création d'un hashtag permet aux usagers de se faire remarquer et de développer un sentiment d'appartenance au groupe réuni autour d'une certaine nouveauté langagière. Par exemple, nous avons constaté que plusieurs comptes publics<sup>242</sup> (tweets n°1.5, 1.6, 1.8) ont choisi, lorsqu'il s'agit de rapporter un fait social de créer un hashtag associé au terme n°1 元芳, 你怎么看 (yuán fāng, nǐ zěn me kàn, « Yuan fang, qu'en penses-tu ? ») : qui exprime ironiquement un doute sur un résultat ou qui invite à commenter. Ces comptes concernent la sécurité routière, un organe de presse, la Cour de justice. Nous supposons que l'emploi de ce hashtag peut permettre dans un premier temps de se rapprocher du public. Comme il est probable que cette expression est aussi utilisée par ce public d'internautes, cet usage commun de l'expression crée le sentiment de proximité. D'autre part, cette expression est aussi une invitation implicite à interagir et cela peut créer une relation entre les comptes des institutions et les internautes. L'emploi de ce hashtag peut aussi aider les comptes en question à établir un image objective et professionnelle et ainsi les faire remarquer parmi les autres comptes publics.

Du point de vue des producteurs, la création d'un hashtag associé à une nouveauté langagière peut être considérée comme une pratique consciente qui met en relief le caractère neuf du segment hashtagué et s'accompagne parfois d'un métadiscours. Du point de vue des récepteurs, la création d'une nouveauté langagière hashtaguée dans un tweet peut sensibiliser les lecteurs potentiels aux phénomènes langagiers et à leurs différentes utilisations (en sachant que producteur et récepteur sont deux rôles qui ne cessent de s'échanger) :

- Ceux qui sont familiers avec le contexte de création mobilisent en premier leurs connaissances. Les hashtags ont une fonction métalinguistique et éventuellement métadiscursive qui renvoie au contexte de création des nouveautés langagières.
- Ceux qui ne connaissent pas le segment hashtagué doivent recourir au reste du message du

---

<sup>242</sup> Sur *Sina Weibo*, plusieurs types de comptes sont proposés. Les comptes certifiés sont marqués par le signe « V ». Des V en couleurs différentes sont employées pour distinguer les catégories de compte certifié. Par exemple, la couleur jaune est pour marquer les comptes des individus et la couleur bleu pour des comptes publics, y compris la presse, le gouvernement, les entreprises.

tweet, qui explicite le hashtag. Dans ce cas-là, les hashtags composent des métadonnées qui résument le contenu associé.

### ***Forte appropriation de la part des usagers***

La constitution d'un échantillon de tweet avait notamment pour but d'observer les évolutions de sens des nouveautés langagières à travers leurs utilisations concrètes dans la publication des tweets. L'étude de leurs sens contextualisés<sup>243</sup> nous a permis de constater déjà une évolution dynamique après leurs premières occurrences néologiques. Nous nous sommes aperçu aussi que l'appropriation de nouveautés langagières dans les tweets se caractérise par une forte dimension personnelle. Cela se manifeste souvent par une interprétation enrichie du sens néologique et un usage élargi.

Intéressons-nous d'abord au sens néologique (ré)interprété par les producteurs des tweets. À titre d'exemple, le sens initial néologique du terme n°2 « nǚ hàn zi » (femme masculine) concerne des femmes indépendantes. Le terme est parfois utilisé au sens péjoratif pour décrire celles ayant des comportements et apparences soi-disant « masculins », comme ne pas porter de jupe ou avoir des cheveux courts. Les 10 tweets de notre échantillon pour ce terme nous ont présenté des images de « femme masculine » plus précises mais totalement hétérogènes. Ainsi le terme (hashtagué) peut renvoyer :

- à une sportive excellente dans le domaine du lancer du poids (tweet n°2.4). L'expression est employée dans un tweet encyclopédique<sup>244</sup> ayant l'objectif de présenter cette sportive. Ici, l'aspect « masculin » est associé à la performance. Les résultats obtenus par cette sportive dans une activité physique demandant la force est mise en valeur puisque les femmes sont souvent considérées comme inférieures aux hommes sur cet aspect.
- à une héroïne qui combat courageusement un malfaiteur pour sauver la vie d'un passager (tweet n°2.7). Ce tweet est un résumé de reportage dont la version complète est intégrée dans le contenu de tweet sous forme d'hyperlien. Il est évident que c'est le courage qui est souligné pour valoriser une image de femme désirant protéger les autres face au

---

<sup>243</sup> Voir Tableau 26 du Chapitre VI.

<sup>244</sup> Comme décrit dans le profil, ce compte est destiné à présenter des informations et connaissances encyclopédiques sur la Nouvelle-Zélande.

danger, au lieu d'être protégée.

- à une dame qui n'arrive pas à trouver l'amour à cause de sa grande taille (tweet n°2.9). Dans ce tweet, l'auteur fait part de ses observations : la plupart des hommes préfèrent les filles de petite taille comme compagnes. On obtient de ce fait l'inférence que la grande taille qui fait « masculin » est un désavantage pour une femme. Il s'agit d'une image de femme centrée sur l'apparence physique. Cette expression est donc employée dans un sens défavorable mais peut aussi être considérée comme une critique portée sur les critères des hommes jugeant l'apparence des femmes.
- à une fille qui porte un élastique à cheveux au poignet au lieu d'un bracelet de luxe (tweet n°2.2). L'auteur du tweet fait un contraste entre les femmes exquises d'un point de vue stéréotypé (selon lequel la féminité est déterminée par les habits ou les accessoires raffinés comme des bijoux chers) et les femmes négligées dont les accessoires sont des objets pratiques. Nous supposons que l'auteur veut présenter une image ordinaire des femmes à travers l'autodérision.
- à un chat femelle qui s'endort avec des positions bizarres (tweet n°2.6). Dans ce cas-là, même si l'expression ne désigne plus des humains, les caractéristiques traditionnelles faisant contraster femmes et hommes sont toujours appliquées. Dans ce contexte du tweet explicité par les photos d'un chat dormant, ce sont des comportements très généraux qui réfèrent à l'image d'une « femme masculine ».

Ces exemples nous permettent de constater que chaque tweet s'appuie sur tel aspect essentiel conditionné par telle situation pour interpréter l'expression. Les différentes interprétations personnelles permettent de rendre la signification de l'expression dynamique sans sortir du cadre d'opposition masculin et féminin.

Nous avons également remarqué que certaines interprétations font construire un sens qui a peu de lien avec le contexte initial. Prenons les termes n°4 主要看气质 (zhǔ yào kàn qì zhì, « c'est le tempérament qui compte ») et n°5 洪荒之力 (hóng huāng zhī lì, « faire tout ce qu'on peut pour réaliser quelque chose ») comme exemple.

Dans le terme n° 4 , le tempérament en s'opposant à l'apparence physique désigne les caractères

inhérents ou les connaissances de quelqu'un. Il est également employé dans le tweet n°4.5 pour décrire les spécificités peu visibles d'une plante aidant à l'identification hors saison de fleurs. Dans le tweet n°4.6, il s'agit de la créativité artistique à développer lorsqu'on s'exprime par le dessin, autrement dit, le style personnel du dessinateur.

Regardons maintenant le terme n°5. Son sens nouveau est associé à un pouvoir fantastique et ultime dans le monde imaginaire d'un roman. Dans le tweet n°5.2, ce pouvoir mystérieux s'est transformé en les efforts qu'un père fait pour maigrir et aussi pour représenter un exemple pour son fils. Le tweet n° 5.3 fournit une autre interprétation de ce pouvoir. Il s'agit dans ce contexte d'une envie forte de manger des produits carnés après avoir été forcé à manger seulement des légumes pour des raisons de santé.

Par ailleurs, pour certains termes utilisés dans des contextes différents, le sens reste pourtant constant. Bien évidemment, c'est aussi à travers les contextes particuliers que nous pouvons repérer les usages divers. Prenons les termes n°1 et n°7. Dans l'ensemble de tweets collectés contenant le terme n°1 元芳，你怎么看 (yuán fāng, nǐ zěn me kàn, « Yuan fang, qu'en penses-tu ? »), son sens néologique reste figé par rapport aux autres termes, mais il est appliqué dans de nouvelles situations différentes de celle initiale consistant à commenter des actualités controversées. Par exemple, le tweet n° 1.1 est publié par un compte de météo de la ville de Xiamen. L'objectif de ce tweet est de partager une connaissance sur l'année en cours, l'année du rat, à travers une activité d'interaction : l'enquête. La spécificité de l'année 2020 est qu'il s'agit d'une année bissextile comptant au total 384 jours selon le calendrier lunaire chinois. On plaisante sur le fait qu'il faudrait travailler un mois de plus par rapport aux autres années pour pouvoir bénéficier des vacances du Nouvel An chinois<sup>245</sup>. L'enquête cherche à collecter les opinions pour et contre cette proposition humoristique. L'emploi du terme vise à faire appel aux lecteurs du tweet afin de les inciter à participer à cette enquête. Il a également été utilisé pour les critiques des productions audiovisuelles, comme dans le tweet n°1.7 qui demande d'évaluer un feuilleton télévisé japonais.

Le terme n° 7 打 call (dǎ call *litt.* «téléphoner») est initialement attaché à la culture des fans de vedettes dans le domaine de concerts live. Les tweets collectés montrent que son utilisation

---

<sup>245</sup> La date du Nouvel An chinois varie d'une année sur l'autre dans le calendrier grégorien parce qu'elle est calculée selon le calendrier lunaire chinois.

s'est beaucoup élargie. En effet, il est employé dans le domaine du fitness pour encourager les membres du club à continuer à faire des exercices (tweet n° 7.1) ; à valoriser des connaissances scientifiques dans le domaine de l'astronomie (tweet n°7.5) ; dans le domaine touristique pour voter pour une ville (tweet n°7.6).

#### **7.1.4. Observations sur la circulation des nouveautés langagières**

Les occurrences des nouveautés langagières que nous avons repérées sur *Sina Weibo* (et sur *Twitter*) sont des hashtags : nous avons expliqué ce choix dans le chapitre VI. Si la liste d'amis d'un compte permet de construire un réseau relationnel fermé et unique à chaque usager, le hashtag dans un tweet favorise quant à lui l'établissement d'un réseau ouvert et dynamique. Comme nous l'avons mentionné précédemment la consultation des tweets regroupés sous un même hashtag ne nécessite pas être abonné à un compte ; non seulement les producteurs peuvent consulter leurs tweets à tout moment sur le réseau, mais aussi les visites des lecteurs sont sans contraintes et imprévisibles. C'est dans ce réseau construit autour du partage d'un même hashtag que nous avons constaté une circulation complexe des nouveautés langagières que nous avons retenues.

La consultation des profils des comptes émetteurs de tweets sur *Sina Weibo* nous a permis de les classer en trois catégories et de constater aussi bien un usage personnel qu'un emploi public :

- 1) *Compte privé portant un pseudonyme, souvent sans certification de Sina Weibo.* Les nouveautés langagières sont dans ce cas utilisées dans des situations personnelles. Par exemple, dans le tweet n°6.1, 藍瘦,香菇 (lán shòu, xiāng gū , trop triste et avoir envie de pleurer) décrit les souffrances liées au fait d'être malade.
- 2) *Compte public, distingué par la marque de certification « V » en bleu.* Il peut concerner des organes gouvernementaux comme la Cour de justice (tweet n°1.8), le Bureau de la sécurité publique (tweet n°1.10) ; des groupes institutionnels comme l'université (tweet n°3.6) ; des entreprises (tweet n°4.10) et des médias (tweet n°1.6, n°3.7), etc. Les objectifs sont déterminés par les rôles du compte. Par exemple, le tweet n°5.5 qui concerne le trafic sur l'autoroute de Hubei est publié par la Police de la route de la province de Hubei.



3) *Compte d'influenceur dont le pseudonyme est suivi par la marque de certification « V » en jaune.* La couleur de certification est identique à celle des comptes d'individus mais ces derniers ont souvent une quantité d'abonnés peu importante alors que les abonnés des comptes d'influenceur sont très nombreux. La publication de tweets de la part de ce type de comptes peut avoir un objectif personnel ou être relative à une réalité publique. Le tweet n°3.10 se moque de la traduction anglaise dans les lieux publics<sup>246</sup> alors que le tweet n°1.9 concerne le partage d'une découverte personnelle.

Le fait que ces trois types de comptes soient impliqués montre que les nouveautés langagières sont largement reconnues. Par ailleurs, les occurrences liées à ces comptes relèvent de domaines de discours très variés :

- marketing : promotion d'un produit (tweet n°5.3), d'un lieu touristique (tweet n° 6.5) ou d'une œuvres audiovisuelle (tweet n°7.2)
- information médiatique concernant: les réalités sociales (tweet n°3.3), les nouvelles scientifiques (tweet n°7.4)
- culture et savoirs : critique portant sur les productions culturelles (tweet n°1.7) ou partage de connaissances encyclopédiques (tweet n°4.6).

Pour résumer, les identités des comptes producteurs et leurs pratiques attestent d'usages étendus des nouveautés langagières. En outre, le partage d'une même nouveauté langagière hashtaguée permet de regrouper des tweets indépendants (sur plan de la relation construite par la liste d'amis) et de former ainsi une communauté propre. Il est à noter que pour certaines occurrences, plusieurs hashtags portant sur des sujets variés sont présents. Dans le même ordre d'idées, il se peut que les nouveautés langagières hashtaguées apparaissent également dans les autres fils de regroupement associés aux hashtags coexistants. De cette manière, une circulation inter-communautés est réalisée, exposant les nouveautés langagières à un groupe de lecteurs élargi.

Les tweets témoignent aussi d'une circulation externe des nouveautés langagières. L'analyse des occurrences de ces dernières dans les tweets nous a permis de constater un phénomène

---

<sup>246</sup> Dans le tweet, l'auteur a énuméré deux exemples : dans le premier cas, le nom anglais d'un magasin dénommé « huá dū » qui vend des aliments et des appareils électriques est en réalité une traduction littérale de son nom chinois : 华都食品机械商店 = *chinese have food machinery store*. Le 2<sup>e</sup> cas est celui d'un panneau de signalisation routière qui indique la voie de bus. Il est étonnant que le message 公交车专用 (voie de bus) soit partiellement transcrit en pinyin dans la traduction anglaise qui devient finalement « bus zhuan yong ».

intéressant. Certains tweets constituent un métadiscours qui rapporte les utilisations de nouveautés langagières dans les situations hors microblog et/ou hors ligne. Prenons le tweet n°6.4 comme exemple. Ce tweet parle de l'utilisation du terme n°6 comme nom d'un plat dans une cantine. Or, le sens néologique du terme concerne en réalité l'expression d'un sentiment de tristesse. En même temps, le terme est présent aussi dans un hyperlien en position finale du tweet. En cliquant sur l'hyperlien, nous avons accès à un article d'une plateforme tierce et dans le titre et le corps de l'article, le terme a été repris. Par un même tweet, nous pouvons donc constater à la fois l'utilisation du terme sur *Sina Weibo*, sur une autre plateforme, et aussi hors ligne.

Après avoir analysé l'échantillon de tweets chinois, nous allons maintenant nous intéresser aux messages et aux situations d'utilisation des nouveautés langagières françaises sur la plateforme *Twitter*.

## **7.2. Analyse discursive d'un corpus de tweets français**

Nous rappelons que comme pour le corpus chinois nous avons collecté sur *Twitter* en français 10 tweets pour chacun des 7 termes retenus. Pour faciliter la démarche d'analyse, nous avons numéroté les 7 termes selon leurs années de création. Dans ce sous-chapitre les numéros affectés aux tweets ne concernent que le corpus français.

Lors de la constitution du corpus de tweets, nous avons constaté que les emplois des nouveautés langagières sont étroitement liés aux contextes linguistiques et communicationnels, aux pratiques des producteurs de tweets et à l'environnement caractéristique du microblog. Nous allons donc reprendre ces trois angles en donnant à ce sous-chapitre une certaine autonomie d'organisation par rapport au précédent tout en nous appuyant sur l'étude déjà réalisée et en procédant sur certains points à une comparaison.

### **7.2.1. Caractéristiques linguistiques sémantiques**

#### *Polysémie entre ancien et nouveau sens*

Les tweets chinois relevés sont basés sur les aspects néologiques des mots ou expressions et nous présumons de ce fait que la création des sujets de la rubrique « Tendances » sur *Sina Weibo* est aussi relative à ces aspects, alors qu'il n'en va pas de même pour les hashtags des tweets

relevés sur *Twitter*. Nous avons en effet remarqué au cours de la collecte que pour certaines nouveautés langagières, les hashtags sont associés à la fois aux sens ancien et néologique. Ce phénomène de mélange entre ancien et nouveau est présent dans les utilisations du terme n°1 *tuerie* et du terme n° 7 *loup solitaire*.

D'un point de vue linguistique, ces deux termes sont marqués par un renouvellement sémantique. En lançant la recherche par mots clés sur *#tuerie*, nous avons constaté que la plupart des hashtags obtenus sont associés au sens de « massacre » comme dans le tweet ci-dessous (Figure 53) qui rapporte une attaque au Japon. Néanmoins, ce à quoi nous nous intéressons, visant la nouveauté langagière c'est le sens de « formidable, sensationnel » (généralement utilisé dans le domaine culinaire) ciblé par les dictionnaires. Pour cela, nous avons spécialement cherché par la suite des hashtags comportant la construction avec présentatif qui consiste en une expression figée *c'est une tuerie* pour la constitution de notre échantillon de tweets.



**Figure 53 : Capture d'écran d'un tweet employant « tuerie » dans son sens ancien**

Quant à l'expression *loup solitaire*, des hashtags avec le signifié « solitude » sont présents avec une nouvelle connotation relative au terrorisme mais beaucoup moins nettement que dans le cas de *tuerie*. Par exemple, dans le tweet n°7.1, l'auteur évoque le fait de ne pas recevoir de notification (probablement concernant les interactions sur les forums ou les réseaux sociaux numériques) et l'interprète comme l'absence de liens affectifs. Le hashtag est donc employé

pour résumer la situation personnelle de l'auteur et souligner son sentiment de solitude. Ou encore dans le tweet n°7. 5, *loup solitaire* sert à décrire le fait de vivre volontairement à l'écart des autres (Figure 55).



Figure 54 : Capture d'écran du tweet n°7.1

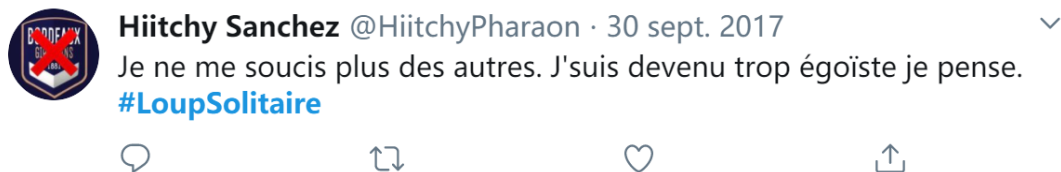


Figure 55 : Capture d'écran du tweet n°7.5

Nous attribuons dans un premier temps ce phénomène d'emploi mixte entre ancien et nouveau sens à une caractéristique linguistique de polysémie. L'utilisation dynamique de leur sens « ancien » peut constituer aussi une explication. Un *tuerie* au sens de massacre est un choc social qui se trouve au centre de la diffusion d'informations. Un *loup solitaire* pour décrire métaphoriquement un vécu personnel est moins explosif dans son utilisation, mais répond peut être à la pratique individuelle en ligne, au besoin de s'exprimer dans un réseau de relation construit. Nous nous demandons si c'est la visée des mots et expressions qui conduit à cette différence observée entre tweets chinois et français ou encore si la contrainte technique de *Twitter* ne nous permet pas de retrouver le même phénomène dans le corpus chinois.

### ***Spécialisation du sens***

Rappelons tout d'abord une particularité de notre échantillon de liste pour le français. Par rapport à la liste chinoise, celle en français inclut un grand pourcentage de noms au sens strictement « spécialisé » qui est reconnaissable alors même que les contextes sont variés. Cette spécialisation dépend de leurs sources de création. Les nouveautés langagières de la liste sont issues d'un mouvement social (n° 2 *marche blanche* et n°4 *zadiste*), d'une nouvelle pratique

(n°3 *nomophobe*), d'un phénomène émergent précis (n°7 *loup solitaire*) ou d'une notion politique (n°6 *vivre ensemble*). Dans la liste chinoise comparable du point de vue des domaines concernés, nous avons découvert des interprétations personnelles qui enrichissent le sens néologique comme par exemple le terme « femme masculine » ou encore un usage élargi. Nous nous demandons si une évolution sémantique des termes au sens spécialisé peut être également constatée sur *Twitter*.

Nous avons premièrement remarqué que dans les tweets français, les emplois des nouveautés langagières hashtaguées concernent souvent une reprise de leur sens néologique mais associée à des événements divers. En voici des exemples :

- L'expression n°2 *marche blanche* est originellement le nom donné à une marche silencieuse de protestation organisée en Belgique en 1996 à la suite de l'affaire Dutroux, un assassin pédophile<sup>247</sup>. Depuis, elle est employée pour désigner une marche organisée autour de la mort d'une victime de fait divers en hommage à cette dernière. Nous avons constaté dans certains tweets collectés que l'expression indiquait toujours les mouvements de soutien et de protestation ou de dénonciation pour des sujets différents. Ainsi dans le tweet n°2.2, il s'agit d'une marche en mémoire de Sylvia Auchter, morte de violences conjugales en 2020. Le tweet n°2.4 parle d'une marche à Etouvans en hommage à Maxime Roussel, jeune victime d'un meurtre en 2012. Nous avons également trouvé une référence à la marche organisée en 2013 en mémoire des victimes de Mohamed Merah, terroriste qui a assassiné plusieurs personnes dont des enfants dans la région de Toulouse l'année précédente (tweet n°2.5), à une marche pour dénoncer des actes pédophiles au Maroc (tweet n°2.6).

Le terme *zadiste* est originellement employé pour nommer les opposants au projet d'aéroport (finalement abandonné en 2018) qui occupent la *ZAD* de Notre-Dame-des-Landes dans l'Ouest de la France à partir des années 2010. Le sigle *ZAD* qu'on utilise aujourd'hui en France pour « zone à défendre » signifie à l'origine « zone d'aménagement différé » et le détournement de sens du sigle en gardant la même forme a lui-même une valeur de protestation. Nous avons constaté sur *Twitter* qu'à part les emplois relatifs au contexte de création de Notre-Dame-des-Landes, « zadiste » est employé pour décrire dans le tweet n°4.1 les bloqueurs des universités qui protestent contre les nouvelles modalités d'accès à l'enseignement supérieur (la loi ORE

---

<sup>247</sup> [https://fr.wikipedia.org/wiki/Marche\\_blanche](https://fr.wikipedia.org/wiki/Marche_blanche).

*Orientation et réussite des étudiants* ou loi Vidal du nom de la ministre) ; pour désigner les militants qui se battent contre le projet de barrage de Sivens dans le Tarn (tweet n°4.10) ; voire à renommer ironiquement le président de la République à la Une du journal satirique *Charlie Hebdo* du 11 avril 2018 (le tweet n°4.4). Pour ce dernier exemple (Figure 56) la caricature reprend les éléments de décor caractérisant la ZAD de Notre-Dame-des-Landes et les événements qui s’y sont déroulés (cabanes construites par des *zadistes*, forces de l’ordre évacuant les occupants) pour critiquer le président Macron réfugié dans une cabane et pointant un fusil, le *zadiste* de l’Elysée »<sup>248</sup>. En plus, *zadiste* est aussi utilisé dans un contexte général pour désigner les protecteurs de l’environnement (tweet n°4.7) ; cette protection est selon l’auteur de ce tweet le vrai rôle des *zadistes* de NDDL.



Figure 56 : Capture d’écran du tweet n°4.4 et page complète de la Une, image rapportée

Les exemples que nous venons de voir confirment que les emplois des hashtags que nous étudions sur *Twitter* soulignent le caractère néologique des nouveautés langagières comme nous l’avons vu sur *Sina Weibo*. Dans les situations évoquées ci-dessus, les producteurs ont une connaissance préalable des nouveautés langagières. La création des hashtags est une action consciente qui demande implicitement aux lecteurs de mobiliser leurs connaissances acquises

<sup>248</sup> Son choix exclusif du média TF1 pour une grande interview télévisée près d’un an après son élection est moqué car il dit « je ne parlerai qu’à Jean-Pierre Pernaut ». Ce présentateur vedette du journal de 13h est jugé peu critique. Il n’est pas spécialiste de politique et valorise les terroirs français. Or l’interview se déroule dans une salle d’école de petit village (c’est une hypothèse que nous faisons après avoir cherché les informations relatives à Macron autour de la date de publication de la Une qui est proche de l’abandon du projet de Notre-Dame-des-Landes mais aussi des évacuations violentes du site).

pour pouvoir appréhender le reste du contenu du tweet.

Nous remarquons aussi que ce sont les contextes associés aux nouveautés qui se multiplient dans les tweets. Nous sommes amenées à faire la comparaison avec le terme n°1 chinois « Yuan Fang, qu'en penses-tu ? » dont le sens néologique reste relativement figé dans ses différentes utilisations. Pour les tweets du corpus français il nous semble le contenu qui implique la nouveauté langagière varie mais reste toujours fortement liés au contexte initial. Ainsi différents mouvements et manifestations sont considérés comme « marche blanche » mais c'est en raison de leur objectif d'apporter du soutien aux victimes. Les militants qui s'opposent à un projet par l'acte d'occupation d'un espace sont appelés des « zadistes » même s'ils n'occupent pas une ZAD. Nous pouvons donc dire que les hashtags ont une fonction de rappel du contexte de création.

### ***Métadiscours***

La conscience ou la connaissance d'une nouveauté langagière sur *Twitter* ne se manifeste pas seulement par les utilisations. Nous avons aussi rencontré des tweets qui parlent des nouveautés langagières elles-mêmes et constituent des séquences métalinguistiques et/ou métadiscursives. Nous avons distingué des cas différents.

- Les tweets s'emploient à rapporter les nouveaux phénomènes langagiers ou à présenter une nouveauté langagière :

Dans le tweet n°4.2 par exemple, l'auteur parle de sa découverte en 2020 de l'intégration des termes *zadiste* et *Zad* dans l'édition 2016 du dictionnaire *Le Petit Robert*. Pour ce même terme *Zadiste*, le tweet n°4.3 datant de 2013 se demande si le terme fera son entrée dans un dictionnaire puisqu'il est en train de faire florès en France suite aux mouvements partout dans le pays qui se réfèrent à Notre-Dame-Des-Landes. Prenons un autre exemple : le tweet n°3.4 est consacré premièrement à présenter l'étymologie du terme *nomophobe* qui est « un mot-valise construit par contraction de l'expression anglaise *no mobile-phone phobia* ». La suite du tweet évoque le sens du terme : « Personne extrêmement dépendante de son téléphone portable ». Et pour une autre expression, *loup solitaire*, le tweet n°7.6 partage l'information, à savoir les explications d'un linguiste sur le sens du terme.

- Les tweets s'interrogent sur la pertinence des nouveautés langagières. Les tweets publiés autour du terme n°7 *loup solitaire* sont d'après nous les plus remarquables :

Selon l'auteur du tweet n°7.3, il ne convient pas à comparer un terroriste qui commet individuellement des actes violents à un *loup solitaire* puisque les loups vivent et chassent en meute. Cela rejoint en partie les idées exprimées dans la Une de *Charlie Hebdo* le 4 octobre 2017, rapportée dans le tweet n°7.4 (Figure 57). Nous pouvons remarquer que l'image de la Une du journal est remplie des *loups terroristes* avec chacun une ceinture explosive. Ce tweet partage donc l'opinion que le *loup solitaire* appartient en réalité à un groupe organisé ou un réseau même s'il agit individuellement. Dans le tweet n°7.8, les interrogations sur le terme sont aussi liées à la nature du loup qui vit en meute. Si les terroristes sont associés aux loups pour leurs actes violents et cruels, l'auteur du tweet conteste la signification néologique de *loup solitaire* à travers un angle différent. Il indique que la figure du loup est dotée d'une impression positive voire noble. De ce fait, il propose de nommer le terroriste individuel *chien solitaire*. Dans le même ordre d'idée, l'auteur du tweet n°7.10 pense que le sens néologique accordé à *loup solitaire* par les journalistes est une insulte aux loups.



Figure 57 : Capture d'écran du tweet n°7.4 et page complète de la Une, image rapportée

Les métadiscours qui prennent pour objet les nouveautés langagières formulent donc soit un commentaire linguistique soit un commentaire sur les phénomènes sociaux à partir d'une réflexion sur les mots employés pour les décrire.



### 7.2.2. Un discours déterminé par les visées des producteurs

À part une forte appropriation personnelle dans des contextes divers, nous avons remarqué que les nouveautés langagières hashtaguées en chinois ont des emplois élargis dans des domaines autres que le contexte initial. Ce qui nous frappe concernant les hashtags du corpus français, c'est que leurs usages semblent plutôt marqués par leurs visées pragmatiques. Autrement dit, les hashtags des twitteurs dans les cas que nous avons relevés se caractérisent par le fait qu'ils sont mis au service de visées variées ayant une portée informative mais aussi argumentative. Nous allons nous intéresser à quelques exemples significatifs dans notre corpus.

#### *Expression de points de vue polémiques*

L'expression *vivre ensemble* est un néologisme politique qui fait appel à « la cohabitation harmonieuse sur un même territoire de peuples et d'ethnies différents »<sup>249</sup>. Mais dans les tweets collectés, cette expression n'est pas comprise de la même manière, non pas pour des raisons sémantiques mais parce que les points de vue ne sont pas identiques. L'expression est ainsi un point de départ au débat quand l'avis sur son contenu est défavorable :

- Le tweet n°6.1 partage à travers un hyperlien des informations relatives à un fait divers : un étudiant israélien a subi une agression raciste dans un espace public après avoir laissé un message téléphonique en hébreu à son père (Figure 58). Le corps du tweet est composé de deux parties : dans un premier temps, l'agression est commentée en la décrivant comme *les miracles du #VivreEnsemble* et ensuite un résumé de cette affaire est proposé permettant aux lecteurs du tweet d'en avoir un aperçu avant d'être dirigé vers un hyperlien. Le racisme est bien évidemment une idéologie qui va à l'encontre de la notion de *vivre ensemble* où les différences et l'harmonie sont mises en valeur. Alors dans ce tweet, avec l'emploi du *vivre ensemble*, le twitteur a en réalité l'intention d'exprimer le contraire : la cohabitation d'ethnies différentes provoque des conflits dans la vie réelle, débouche sur le racisme et l'harmonie n'est pas vraie. Ici, l'expression est donc utilisée ironiquement.

---

<sup>249</sup> Définition du Wikipédia : [https://fr.wiktionary.org/wiki/vivre\\_ensemble](https://fr.wiktionary.org/wiki/vivre_ensemble), consulté le 2 octobre 2020.



Figure 58 : Capture d'écran du tweet n°6.1

Le tweet n°6.2 suit la même présentation que le 6.1 : un partage d'information via un hyperlien avec un commentaire constitué du point de vue du twitteur et un résumé du fait concerné (Figure 59). Pour ce twitteur, le choix du prénom *Jihad* pour un petit garçon prête vraiment à controverse dans le contexte d'attentats terroristes que connaît la France. Ce prénom qui peut faire penser à « la guerre sainte menée pour propager l'islam » est considéré comme une provocation dans le monde du #VivreEnsemble selon l'auteur du tweet. Ici, on peut se demander si l'expression *vivre ensemble* est employée dans sa véritable signification ou si elle est mise à distance sans aller jusqu'à l'ironie du tweet précédent. Les deux faits concernés n'ont pas le même degré de gravité mais renvoient tous deux à un contexte négatif et pas du tout harmonieux.



Figure 59 : Capture d'écran du tweet n°6.2

Nous avons constaté les mêmes phénomènes tendant à mettre en question l'acception des

termes pour aller vers une critique et un débat social dans d'autres cas. Par exemple, l'auteur du tweet n°2.1 considère que la *marche blanche* qui se déroule à Argenteuil suite à la mort d'un jeune (non casqué) dans une collision à moto n'est plus un mouvement contre une injustice mais plutôt en faveur de la délinquance. Selon une recherche supplémentaire sur cette affaire, nous avons découvert que la mort de cette victime a suscité des nuits de révolte successives avant la *marche blanche*. Cette colère violente est en lien avec la présence policière à proximité de la collision mortelle et aux conflits qui opposent les policiers et les habitants des quartiers populaires.



Figure 60 : Capture d'écran du tweet n°2.1

Le tweet n°4.1 fournit un autre exemple de polémique. Ce tweet est publié dans le contexte de la contestation de la réforme de l'accès à l'université (ORE). A partir des informations retweetées qui indiquent que la décision de la reprise des cours est largement approuvée au sein de la FAGE<sup>250</sup>, nous pouvons comprendre que l'auteur du tweet n'est pas favorable au blocage des universités (une forme d'action des opposants à la loi). Dans son commentaire, les bloqueurs sont considéré comme des *zadistes* à la poursuite d'une démocratie « alternative ». L'adjectif « alternative » mis entre guillemets signale la mise en doute ironique de cette pratique de la démocratie.

---

<sup>250</sup> La Fédération des associations générales étudiantes.



Figure 61 : Capture d'écran du tweet n°4.1

### *Expression d'opinion en lien avec l'identité du compte*

Si le contexte est important pour élucider les points de vue, les identités des comptes de *Twitter* sont aussi un élément d'influence. Nous avons trouvé l'expression d'« identité numérique » chez Cardon (2019) pour décrire les phénomènes des identités en ligne. Il la définit comme « une coproduction où se rencontrent les stratégies des plateformes et les tactiques des utilisateurs » (Cardon 2008:97). Fanny Georges (2009) a repris cette définition et distingué trois dimensions de « l'identité numérique », l'identité déclarative, l'identité agissante, l'identité calculée. L'« identité agissante » est constituée des messages répertoriés par le Système, concernant les activités de l'utilisateur (exemple : « x et y sont désormais amis ») ; L'« identité calculée » se compose de chiffres, produits du calcul du Système, qui sont dispersés sur le profil de l'utilisateur (comme : le nombre d'amis, le nombre de groupes). L'identité du compte à laquelle nous nous intéressons ici rejoint la dimension de l'« identité déclarative » qui « se compose de données saisies directement par l'utilisateur, notamment au cours de la procédure d'inscription au service (exemple : nom, centres d'intérêt, amis) » (Georges 2009:179).

Il faut souligner que les identités créées par les usagers pour se représenter sur la plateforme sont des identités propres à *Twitter*. Il se peut que l'identité numérique sur *Twitter* soit différente que celles des autres plateformes ou de la vie réelle. Selon les caractéristiques techniques, la reconnaissance d'identité des twitteurs peut être réalisée à partir des éléments suivants :

- *Le nom d'utilisateur et/ou le nom d'affichage*. Un nom d'utilisateur est le nom demandé lors de la création du compte sur Twitter. Il est unique et non modifiable après avoir validé

l'inscription. Sa forme de présentation sur la plateforme est la suivante : @+nom d'utilisateur. Le nom d'affichage ou le nom du profil est le pseudonyme du compte qui se place en tête du nom d'utilisateur. Il peut être modifié à tout moment. Les deux noms associés à un compte peuvent être identiques, comme le montre la Figure 62 ci-dessous : le Président de la République française a appliqué son nom personnel à son compte *Twitter* et au nom du profil.



**Figure 62 : Eléments d'identité déclarative sur le compte d'Emmanuel Macron**

- *La description du profil.* Le profil d'un compte *Twitter* est composé de plusieurs éléments<sup>251</sup>. Nous choisissons de nous intéresser à la partie biographie du profil car nous pouvons avoir un accès relativement direct à la description verbale et/ou aux signes qui s'y trouvent. Quant aux images, nous avons besoin de les agrandir pour observer les détails et ces images sont dans la plupart des cas des photos de portraits qui nous posent des difficultés. Il nous est impossible de savoir s'il s'agit du portrait du twitteur ou d'autres personnes. Donc, nous trouvons que ce n'est pas un élément qui convient pour repérer l'identité numérique d'un compte. Mais, il est à noter que ce champ n'est pas obligatoire à remplir. Tout dépend du choix de chaque twitteur. Ainsi, la description du profil ne s'affiche pas automatiquement dans le corps du message de tweet. C'est-à-dire que dans notre corpus collecté, nous ne pouvons pas acquérir cette information. Il nous est nécessaire de faire des démarches supplémentaires pour connaître la description du profil.
- *Les types de compte:* Sur *Twitter*, les catégories des comptes se divisent en deux : les comptes certifiés et les comptes non certifiés. Les comptes certifiés se distinguent des autres par un badge bleu placé après le nom sur le profil du compte. Selon l'explication de la plateforme<sup>252</sup>, « un compte peut être certifié s'il présente un intérêt public ». Ce système

<sup>251</sup> Y compris : photo de profil, le nom, la biographie, la localisation, le site web, la date naissance, etc.

<sup>252</sup> <https://help.twitter.com/fr/managing-your-account/about-twitter-verified-accounts>, consulté le 20 octobre 2020.

de classification n'est pas aussi détaillé que sur *Sina Weibo* qui divise également les comptes certifiés en personnel et public en attribuant des couleurs différentes. Nous pensons qu'un compte certifié jouit souvent d'une grande visibilité et peut être reconnu facilement.

Nous présumons que les identités des comptes encadrent préalablement le positionnement de chaque twitteur. Notre attention est d'abord attirée par les comptes certifiés puisque qu'il est facile de les repérer. À titre d'exemple, les tweets n°6.3 et n°6.6 sont publiés respectivement par Robert Ménard, un homme politique d'extrême droite et Christophe Clergeau, adhérent du Parti Socialiste. Leurs identités peuvent être immédiatement saisies à travers leurs noms pour des lecteurs qui ont une bonne connaissance du domaine politique. Les deux comptes disposent l'un et l'autre d'une courte description qui peut nous permettre d'avoir plus d'informations sur leurs identités et positionnements politiques. Pour ces deux personnes, *vivre ensemble* ne porte pas les mêmes valeurs. Ménard (tweet n°6.3) est accusateur : « le #vivre ensemble est la plus grande escroquerie du monde » (Figure 63) et renvoie aux sujets de l'islam et de l'immigration. Si nous nous référons à son identité politique, nous pouvons comprendre que l'idée exprimée dans ce tweet rejoint les fondements idéologiques de l'extrême droit qui sont marqués par leur hostilité générale à l'immigration et par des positions ouvertement racistes. Le *vivre ensemble* employé par Ménard est visiblement pris dans un sens défavorable. Alors que Clergeau lui emploie l'expression pour faire l'éloge de la loi 1905 sur la laïcité (Figure 64). Depuis son profil sur *Twitter*, nous pouvons comprendre que le socialisme est le point de départ de ses expressions : on recherche une organisation sociale et économique plus juste. Nous avons également effectué une recherche d'information supplémentaire sur la personne et nous savons qu'il est enseignant à l'université. Tout cela explique son soutien à la loi de 1905 et le fait que *vivre ensemble* n'est plus mal vu comme dans le tweet de Ménard mais représente une valeur positive.



Figure 63 : Capture d'écran du tweet n°6.3



Figure 64 : Capture d'écran du tweet n°6.6

Le tweet n°6.10 constitue aussi un exemple intéressant qui porte sur un sujet politique. Le twitteur est un individu dont le compte n'est pas certifié. Mais à partir de son pseudonyme (SniperDeDroite) et la description du profil<sup>253</sup>, nous pouvons connaître la prise de position de ce twitteur qui est un supporter de la droite. Dans son tweet, l'auteur joue d'une manière ironique sur les différences des opinions politiques entre les partis français pour exprimer sa défiance et son mépris à l'égard de l'interprétation de la sécurité des socialistes (Figure 65). Pour s'en moquer, il compare ce point de vue socialiste de la sécurité au soutien au *vivre ensemble* par Marine Le Pen. En associant un parti d'extrême-droite à une idée politique à laquelle celui-ci s'oppose, cette comparaison crée un contraste entre idées exprimées et positions politiques, ce qui révèle l'ironie. Comme il est impossible que Marine le Pen approuve le « vivre ensemble », on ne peut pas croire à la « sécurité » socialiste. Dans ce tweet qui est destiné à se moquer de la crédibilité des socialistes, *vivre ensemble* version Marine le Pen constitue à nouveau un autre emploi de la nouveauté langagière qui diffère son sens discursif initial.



Figure 65 : Capture d'écran du tweet n°6.10

### *Transmission d'informations*

Nous avons rencontré plusieurs tweets de type informatif dans notre corpus. Ils possèdent les

---

<sup>253</sup> « Sarkozyste historique, pour une #DroiteDesValeurs plus que jamais dans l'opposition ».

points communs suivants: le tweet est souvent composé d'un message synthétique et/ou un hyperlien. Cet hyperlien peut diriger les lecteurs vers la version complète des informations (il s'agit des articles de presse ou des productions encyclopédiques) hébergées par des plateformes extérieures. Le contenu transmis dans le corps du message du tweet est un résumé ou une citation des éléments de l'article. Dans ce cas-là, les nouveautés langagières hashtaguées sont employées comme des mots clés. En nous référant aux fonctions des hashtags constatées dans le corpus de tweet chinois, nous avons découvert que dans ces tweets informatifs, les hashtags ont une fonction principale ou auxiliaire.

Tout d'abord, nous avons remarqué plusieurs tweets similaires qui partagent des articles au sujet de *nomophobe*. Il s'agit des tweets n°3.5, n°3.6, n°3.7. Prenons le tweet N°3.5 comme exemple (Figure 66). Ce tweet sert à présenter un article qui est consacré à évoquer le phénomène de la « nomophobie », y compris la manifestation de cette nouvelle « maladie », certains chiffres d'enquêtes et des suggestions. Le message du tweet commence par une question « Etes-vous #nomophobe » qui est aussi le titre de cet article. Cette question est suivie par une explication du terme *nomophobe* (« souffrir de ne pas avoir son téléphone portable à portée de main ») émanant également de l'article. La source de publication (le site de *Slate*) et un hyperlien renvoyant à l'article sont ajoutés à la fin du tweet. Nous présumons qu'en posant une question à ses futurs lecteurs, l'auteur a pour objectif de les inviter à cliquer sur le lien et à explorer les détails pour en chercher une réponse. Selon le contenu partagé, nous pouvons aussi déduire que ce tweet vise un public qui ne connaît peut être pas le terme mais a en commun ou témoigne des même « symptômes ». Le tweet comporte un seul hashtag #nomophobe qui joue une fonction principale. C'est-à-dire que le hashtag souligne le thème du tweet et sert à attirer les lecteurs qui s'intéressent à ce sujet (une fois le néologisme élucidé).



Figure 66 : Capture d'écran du tweet n°3.5

Cet aspect informatif est particulièrement évident dans les tweets médiatiques. Le tweet n°2.2 est publié par un ancien journaliste d'*Europe 1*. Il rapporte une *marche blanche* (Figure 67).



Nous apprenons tout d'abord le nombre de participants à la marche blanche et ensuite les informations relatives à la victime, Sylvia Auchter qui a subi des violences conjugales. En même temps, le journaliste invite aussi les lecteurs à suivre une émission sur *Europe 1* à 12h30 qui fera le récit du mouvement de rassemblement. Ce tweet partage en même temps deux photos prises sur place qui montrent le déroulement de la marche. A la fin du corps du message, le hashtag *marche blanche* est précédé des trois autres hashtags qui concernent respectivement la cause de mortalité #féminicide, la commune #Oberhoffensurmoder et la région #Alsace où se déroule cette activité de soutien. En général, le tweet a pour objectif de diffuser les informations relatives à un mouvement. La création du hashtag *marche blanche* permet de donner un aperçu rapide du sujet de contenu et de classer en même temps ce tweet dans les files des tweets regroupés sous le même hashtag. Par rapport aux autres hashtag qui figurent dans le message, *marche blanche* est un hashtag clé puisque les autres se concentrent chacun sur un aspect d'information comme nous l'avons évoqué ci-dessus.



Figure 67 : Capture d'écran du tweet n°2.2

La publication du tweet n°2.5 a le même objectif de partage d'information, mais le contenu du tweet est plus bref (Figure 68). Le compte appartient à un site d'information sur Toulouse et sa périphérie. Selon le partage, nous pouvons savoir qu'il y avait un rassemblement de 1000 personnes environ à Toulouse en mémoire des victimes de Mohamed Merah. Un hyperlien est ajouté pour accéder à la page d'information sur le site. Nous avons trouvé également un hashtag

principal *marche blanche* et d'autres hashtags auxiliaires qui indiquent du lieu de déroulement du mouvement #Toulouse et du meurtre #Merah.

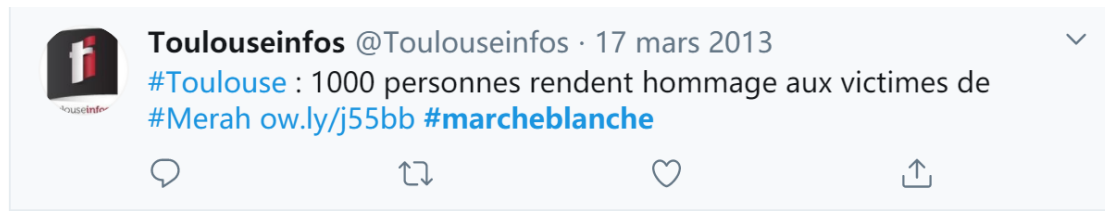


Figure 68 : Capture d'écran du tweet n°2.5

Les exemples ci-dessus montrent que la création d'un hashtag associé à une nouveauté langagière a pour but d'inviter les lecteurs à consulter le contenu relatif à ces nouveaux termes.

### *Des objectifs d'exhortation*

Nous avons ainsi remarqué que dans les tweets français collectés, certains sont publiés pour inciter d'une manière implicite ou explicite leurs lecteurs à réaliser quelque chose. En voici deux exemples.

L'expression *c'est une tuerie* est beaucoup utilisée aujourd'hui au sens figuré « c'est formidable ; c'est super ». Ce sens néologique est originellement rattaché au domaine culinaire pour qualifier les plats et *tuerie* n'est plus associée à l'action de massacre dans cette métaphore hyperbolique qui insiste sur l'intensité et inverse le négatif en positif dans une sorte d'antiphrase. Faut-il comprendre que c'est tellement bon (intéressant, excitant...) qu'on en meurt (de plaisir) ?

Dans les tweets relatifs à l'expression, nous avons premièrement constaté que les domaines d'application sont élargis. L'expression est associée à un jeu vidéo (tweet n°1.1), à un album (tweet n°1.3), à une fonction photographique de téléphone (tweet n°1.6), à toutes les possibilités qui peuvent provoquer une passion. Parmi ces utilisations diverses, nous avons ensuite découvert que le partage d'un enthousiasme envers une chose donnée peut être interprété en même temps comme une recommandation ou une invitation d'essai. Par exemple, en ajoutant l'expression hashtaguée à la fin de son message, l'auteur du tweet n°1.5 a pour but de souligner son appréciation pour l'émission télévisée « The Voice » édition 2013 (Figure 69). Le choix de montrer son intérêt vif envers une émission peut être perçu comme une incitation à regarder cette émission pour ceux qui ne connaissent pas encore, ou pour ceux qui en ont déjà

connaissance une incitation à échanger leurs opinions. On constate qu'il n'en va donc pas de même dans les tweets relatifs à l'expression *marche blanche* dans lesquels les actes d'incitation sont expressément manifestés.



Figure 69 : Capture d'écran du tweet n°1.5

Pour donner un autre exemple, le tweet n°2.6 est un message publié sur la plateforme pour appeler les habitants de Casablanca à participer à un mouvement pour lutter contre la pédophilie (Figure 71). L'auteur a tout d'abord précisé l'heure et l'endroit de rassemblement. Il a indiqué dans le message suivant qu'il ferait spécialement le déplacement de Marrakech à Casablanca pour apporter son soutien et il a lancé à nouveau cet appel pour inviter les habitants locaux à y participer. Le hashtag *marche blanche* est placé dans le message avec les informations de rassemblement corrélées. Nous trouvons deux autres hashtags à la fin du message qui indiquent le sujet #pedophilie et l'endroit #maroc de la *marche blanche*. La *marche blanche* de ce message est une activité à venir et non pas comme dans les cas vus précédemment un événement révolu. L'objectif de l'auteur est d'inciter les lecteurs du tweet à aller manifester.



Figure 70 : Capture d'écran du tweet n°2.6

Le même type d'information est aussi constaté dans le tweet n°2.4 publié par le compte certifié d'un quotidien régional, *l'Est Républicain*. Selon le partage, nous pouvons avoir accès à des informations clés : la date (demain, soit le 14 janvier 2012), le lieu (à #Etouvans) de la « marche blanche » ainsi que la victime (#MeurtredeMaxime) à laquelle il est rendu hommage. Ce message fournit en plus un hyperlien permettant aux lecteurs de connaître plus de détails sur la victime d'une part et le déroulement de l'événement à venir d'autre part (Figure 72). A travers

ce tweet, nous pouvons constater une double visée : une information est rapportée et un appel est lancé. L'identité du compte nous fait associer dans un premier temps ce tweet à l'objectif de rapporter l'information. Or, le contenu précis, surtout le rappel d'une date à venir (demain) ainsi que l'intégration du tweet dans le fil *marche blanche* en créant le hashtag nous permet d'associer ce tweet à un objectif semblable à celui du tweet n°2.6.



Figure 71 : Capture d'écran du tweet n°2.4

En fait, il se peut que les différentes visées (informative, polémique, exhortative) soient présentes dans un même tweet et se combinent. L'identification des visées de publication dépend non seulement des éléments langagiers du corps message du tweet mais aussi de l'ensemble discursif ou plutôt techno-discursif d'un tweet.

### 7.2.3. Caractéristiques technodiscursives sur *Twitter*

Nous avons choisi d'observer la forme composite (Paveau 2017:65) des nouveautés langagières sur la plateforme de *Twitter*. C'est-à-dire que la forme n'est pas purement constituée par des éléments langagiers mais permet en même temps de réaliser une fonction technique. Leur circulation n'est pas conditionnée par le réseau amical construit autour d'un compte mais intervient dans un réseau de relation plus ouvert et dynamique. Comme nous l'avons évoqué dans la partie d'analyse du corpus de tweet chinois, la création d'un même hashtag dans les tweets favorise la formation d'un réseau techno-relationnel où les pratiques sont multiples. Des phénomènes de circulation similaires sont constatés tant sur *Sina Weibo* que sur le *Twitter* version française. Mais la circulation des nouveautés langagières françaises est selon nous beaucoup marquée par la dimension hypertextuelle du tweet, c'est-à-dire que la présence des hyperliens dans le corps du message est remarquable.

## *Intérêts individuels et publics*

La distinction des types de compte<sup>254</sup> sur *Twitter* est basée sur l'intérêt tel que l'apprécie la plateforme. Nous avons indiqué précédemment que la plateforme accorde une certification à des comptes s'ils présentent un intérêt public. Cela crée une opposition entre les comptes certifiés et non certifiés. Dans le corpus de tweet français, nous avons des publications émanant de ces deux types de compte et nous avons constaté que la situation était nuancée selon les nouveautés langagières.

Nous avons rencontré davantage de comptes individuels et non certifiés dans les emplois des hashtags *C'est une tuerie*, *nomophobe*, *tendu comme un string*. Les autres termes de notre sélection, *marche blanche*, *zadiste*, *vivre ensemble*, *loup solitaire* sont employés soit par des comptes personnels, soit par des comptes « publics ». Il faut préciser que les comptes publics dont nous parlons ne sont pas que « certifiés ». Ils comprennent à la fois des comptes certifiés selon les critères de *Twitter* et des comptes d'institutions comme par exemple *La Médiatrice* (qui publie le tweet n°7.6), compte du service de la médiation de Radio France. Ce n'est pas un compte personnel mais il n'est pas non plus certifié.

Nos constats nous ont amenée à nous interroger dans un premier temps sur le fait de savoir si les caractéristiques sémantiques des nouveautés langagières étaient à l'origine des nuances. Les trois premières expressions sont plus associées à une émotion ou une pratique personnelle qu'à des événements sociaux comme le sont les quatre autres. Les termes relatifs à l'actualité seraient-ils plus utilisés par des comptes publics ? Dans un deuxième temps, nous nous demandons si c'est le petit volume du corpus et le choix aléatoire de celui-ci qui nous conduisent à une observation relativement limitée sans que nous puissions conclure sur ces tendances.

Quoi qu'il en soit des intérêts personnels ou privés, nous avons bien constaté une utilisation élargie des nouveautés langagières. Dans les termes à tendance d'usage personnel, c'est l'expansion sémantique qui nous retient le plus. Nous avons déjà mentionné que dans les utilisations du terme *c'est une tuerie* l'enthousiasme personnel peut être suscité par des objets divers. De même, *tendu comme un string* ne signifie pas que « stressé ». La comparaison

---

<sup>254</sup> Nous n'avons pas pris en compte les identités de compte lors de la sélection du corpus ; nous avons privilégié dans un premier temps le contenu du tweet.

humoristique a été employée pour décrire des sentiments d'excitation (tweets n°5.2 et n°5.8), d'angoisse (tweets n°5.3, n°5.4), d'impatience (tweet n°5.5), etc. Les usages nous paraissent donc liés aux identités des comptes et aux visées des auteurs.

### ***Forte hypertextualité***

Le partage d'une nouveauté langagière dans les tweets collectés peut être réalisé sur plusieurs dimensions. La première dimension dépend de la création d'un hashtag associé qui permet de regrouper les tweets concernés dans un même réseau thématique (le thème étant déterminée par le hashtag). La deuxième dimension est marquée par la relation entre le hashtag et les autres éléments du tweet. Dans le corpus de tweets français, notre attention est attiré par le rapport entre les nouveautés langagières hashtaguées et les autres formes techno-langagières (les autres hashtags, les mentions précédées de @ et les hyperliens insérés).

Nous avons constaté qu'une part importante des tweets français est fortement hypertextuelle. C'est-à-dire que ces tweets sont souvent composés de plusieurs hashtags et/ou un hyperlien. Ces éléments cliquables rendent l'ensemble du tweet délinéarisé (Paveau 2017). Un tweet peut être dirigé vers plusieurs fils internes de *Twitter* mais différents ou vers une plateforme externe comme des sites web. Nous avons également constaté les circulations interne et externe des nouveautés langagières sur *Sina Weibo*, mais cette dimension hypertextuelle est plus marquée sur *Twitter* à travers le visuel et les relations qu'entretiennent les éléments. Il s'agit de formes brèves enrichie d'éléments techno-discursifs.

À première vue, les tweets français semblent moins riches en contenu que les tweets chinois du moins dans notre corpus. Cela peut être expliqué par la compacité de la langue chinoise par rapport au français (autrement dit, l'expression d'une même idée dans la langue chinoise peut engendrer une forme plus courte que dans la langue française) et la différence du comptage de signes sur les deux plateformes (un caractère chinois occupe un espace sur *Sina Weibo* alors que sur *Twitter*, une lettre correspond à un espace). Dans ces messages brefs, il peut arriver que les hashtags et/ou hyperliens occupent une place prédominante. Par exemple, reprenons le tweet n°2.4 (Figure 72). Nous y trouvons trois hashtags et un hyperlien affichés en couleur bleue. Les informations clés du message (relatives à un mouvement de soutien) sont transmises à travers ces trois hashtags. L'hyperlien sert à donner des informations complémentaires et détaillées visant à élucider le contexte.



Figure 72 : Capture d'écran du tweet n°2.4

Dans les autres tweets du corpus français, les éléments verbaux ne sont pas si brefs mais la densité hypertextuelle reste une des caractéristiques non négligeables. Il faut aussi les distinguer des tweets chinois destinés à promouvoir certains produits et qui à cette fin ajoutent des hashtags en quantité. Mais il se peut que ces hashtags aient peu de relation avec le contenu du tweet. Or, ce n'est pas le cas dans nos observations sur le corpus français.

La présence des hyperliens connaît deux présentations : 1) l'hyperlien se trouve dans le contenu de retweet ; 2) l'hyperlien est inséré manuellement dans le corps message du tweet. Cette deuxième présentation domine dans le corpus que nous avons collecté. La circulation de la nouveauté langagière est aussi réalisée à travers ces hyperliens.

Dans le premier cas, à part les possibles interactions qui peuvent être suscitées par la publication de retweet (c'est-à-dire, les interactions dans le réseau relationnel construit autour du compte qui retweet et commente un tweet), la circulation entre également dans le réseau généré par le tweet original. Les lecteurs du tweet original peuvent être sensibilisés à l'utilisation des nouveautés langagières s'ils consultent les interactions qui se trouvent en bas du message de tweet. La circulation est ainsi réalisée simultanément dans les deux sens mais aussi indépendamment. Le retweet d'un hyperlien permet d'élargir le réseau de circulation à l'intérieur de *Twitter*.

Dans le deuxième cas, l'insertion manuelle d'un hyperlien peut générer également deux formes différentes. L'une est présent directement dans le message comme un lien cliquable de couleur bleue, comme dans le tweet n°2.4. L'autre est affiché comme un contenu joint en bas du message, avec souvent un résumé d'information et une image de couverture, comme dans le tweet n°6.1 plus haut. Cet hyperlien offre généralement un accès au contenu extra-plateforme où les nouveautés langagières sont également employés. Dans ce cas-là, l'hyperlien sert comme témoin d'une circulation externe des nouveautés langagières. Nous l'avons aussi constaté dans le corpus de tweet chinois.

### *De nouvelles relations entre hashtags et corps du message*

Les relations qu'entretiennent les hashtags dans un même tweet semblent plus complexes dans le corpus français que dans le corpus chinois et nous ne pouvons pas caractériser les nouveautés langagières hashtaguées comme fonction principale, auxiliaire ou de promotion comme nous l'avons fait pour l'échantillon des tweets chinois. Nous avons pu distinguer deux catégories de hashtags du point de vue de leurs relations

- *Des hashtags explicatifs* : Ce sont des hashtags qui apparaissent comme des synonymes de la nouveauté langagière hashtaguée qui réalisent une expansion sémantique de celle-ci. Il faut pourtant les distinguer des hashtags qui portent sur les informations contextuelles. Par exemple, dans le tweet n° 3.8, nous pouvons trouver deux hashtags : #nomophobe, #accroàTwitter (Figure 73). Le message parle du sentiment de l'auteur face au dysfonctionnement du Wi-Fi. Nous pensons que *accroàTwitter* est une réinterprétation du terme *nomophobe*. Cette nouveauté langagière est employée ici pour souligner la passion de l'auteur envers *Twitter*. Elle ne limite plus à décrire les personnes dépendantes de leurs téléphones portables. Selon ce hashtag explicatif, nous pouvons comprendre qu'il est insupportable pour l'auteur de se déconnecter du réseau d'Internet puisque cela l'empêche d'utiliser les service de *Twitter*.



**Figure 73 : Capture d'écran du tweet n°3.8**

Pour distinguer un hashtag explicatif d'un hashtag qui donne des informations contextuelles, prenons encore le tweet n°6.9 publié par le compte des Allocations Familiales. Dans ce message, #cohésionsociale constitue le hashtag explicatif de la nouveauté langagière hashtaguée #vivreensemble qui est utilisée dans un sens positif. Tout cela a pour objectif de présenter le Senacs, un système d'échanges national des centres sociaux. Le hashtag associé



à ce système offre à son tour un contexte pour interpréter l'usage de la nouveauté langagière.



Figure 74 : Capture d'écran du tweet n°6.9

- *Des hashtags juxtaposés/inter-thématiques.* Il s'agit de hashtags qui sont relatifs originellement à un autre sujet, avec un contexte prédiscursif<sup>255</sup>. L'interprétation de ces hashtags demande aux lecteurs d'obtenir des connaissances sur ces sujets. Et c'est plutôt à travers le hashtag inter-thématique qu'une circulation inter-communauté de la nouveauté langagière se réalise. Dans le tweet n° 4.10 ci-dessous (Figure 75), six hashtags ont été créés.

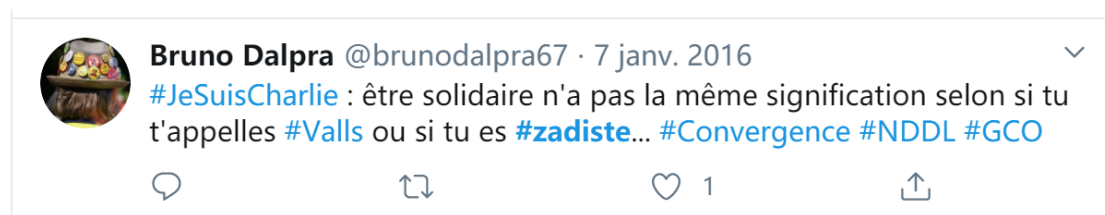


Figure 75 : Capture d'écran du tweet n°4.10

Selon les contextes, le slogan #JeSuisCharlie qui est utilisé pour défendre la liberté d'expression se différencie des autres hashtags qui sont liés à la protection d'une zone à défendre. Par cette différenciation, nous pouvons classer #JeSuisCharlie comme un hashtag inter-thématique. Nous présumons à travers la création de ce slogan dans le tweet, que l'auteur a pour objectif d'apporter son soutien à une forme de militantisme solidaire, les actions des zadistes, en osant exprimer un point de vue différent de celui des autorités. Cela s'inscrit dans les valeurs de #JeSuisCharlie. Revenons maintenant sur le contenu détaillé du message. Selon la date de publication (le 7 janvier 2016), ce tweet est probablement une réaction à la décision du gouvernement Valls de ne pas renoncer au projet d'aéroport de Notre-Dame-des-Landes ou de désigner le concessionnaire du projet autoroutier GCO<sup>256</sup>.

<sup>255</sup> Nous appliquons ici l'interprétation du « prédiscours » selon Paveau (2006:18), qui le définit comme « ce qui relève de ce qui se passe avant le discours ».

<sup>256</sup> Voir : <https://www.rue89strasbourg.com/sept-questions-pour-tout-comprendre-sur-le-gco-142802>.

Cela peut être compris comme une critique contre le gouvernement qui n'écoute pas les voix différentes. Or, dans le même cadre de mouvements zadistes, nous trouvons deux hashtags associés respectivement à deux affaires différentes, #NDDL et #GCO. Ils sont ici un hashtag inter-thématique l'un et l'autre si l'attention est accordée à l'affaire précise. Cet exemple montre que le terme *zadiste* peut être connu par les lecteurs qui s'intéressent à des tendances #JeSuisCharlie, #NDDL et #GCO sur *Twitter* puisque ce tweet est aussi regroupé dans les fils d'affiche de ces trois hashtags.

Les observations sur l'échantillon de tweet français montre dans un premier temps que les utilisations des nouveautés langagières sur *Twitter* sont beaucoup plus marquées par un usage multiple que par une création sur le seul plan sémantique. Leurs emplois sont associés à des contextes différents, aux visées des twitteurs parmi lesquelles nous pouvons distinguer : exprimer un point de vue, servir un objectif informatif ou appeler à agir. Pourtant, cela n'implique pas que les créations sémantiques ne sont pas importantes dans le corpus français. Le recours à l'aspect linguistique et contextuel de ces nouveautés langagières nous permet de formuler une hypothèse selon laquelle l'évolution des nouveauté langagière est influencée dans une grande mesure par les termes eux-mêmes, autrement dit, ce sont leurs caractéristiques linguistiques qui déterminent les usages sémantiques et contextuels. En même temps, nous avons constaté que l'identité du twitteur joue aussi un rôle déterminant dans l'interprétation du sens des nouveautés langagières. Ensuite nous avons remarqué que les nouveautés langagières sélectionnées sont employées par des comptes individuels aussi bien que par des comptes publics, indiquant un usage large. En plus, les tweets français sont souvent caractérisés par un contenu bref comportant des éléments hypertextuels, comme par exemple, les hashtags et hyperliens.

## **Conclusion**

En nous interrogeant sur les relations qu'entretiennent les nouveautés langagières hashtaguées et les autres éléments technodiscursifs, nous avons pu observer une circulation interne des nouveautés langagières favorisée par l'écosystème de la plateforme. Nous distinguons dans cet environnement écologique un réseau relationnel construit autour d'un compte et un réseau techno-relationnel formé autour de la création d'un hashtag. En même temps, la dimension hypertextuelle à travers les hyperliens nous amène à constater une circulation extra-plateforme des nouveautés langagières concernées.

Au cours de l'analyse, nous avons réalisé en parallèle une comparaison entre les corpus chinois et français. Tout d'abord, les nouveautés langagières hashtaguées sur les deux plateformes montrent une évolution de sens et un usage élargi dans les deux langues. Pourtant, la création sémantique est plus significative dans le corpus chinois alors que les emplois associés à des contextes différents sont très remarquables dans le corpus français.

Nos observations sur le corpus français ont confirmé l'hypothèse élaborée lors de l'analyse sur le corpus chinois. Le choix d'observer les nouveautés langagières hashtaguées, une forme fixée sous-jacente, ne nous permet pas d'observer les potentialités d'évolution formelle. Quant à la circulation, nous avons également constaté des points communs dans les deux corpus. Les nouveautés langagières chinoises et françaises bénéficient d'une circulation intra et extra plateforme à travers les fonctions technodiscursives des microblogs. Leur popularité est ainsi prouvée.

Nous pouvons également avancer que les différences linguistiques entre les deux langues et les nuances fonctionnelles des plateformes affectent la composition des tweets. Un tweet français est moins riche dans son contenu et fortement hypertextuel. Cette influence de la structure nous a amenée à nous concentrer spécialement sur les rapports entre les éléments hypertextuels et à repérer les fonctions explicatives et inter-thématique des hashtags du corpus français alors que dans le corpus chinois, notre attention a surtout porté sur la relations qu'entretiennent les hashtags avec les nouveautés langagières associées et les éléments non hashtagués. Pour nous résumer, à travers les angles différents, nous avons confirmé que les emplois des nouveautés langagières sur les microblogs chinois et français dépendent à la fois de l'environnement techno-discursif et des pratiques des twitteurs.

## CONCLUSION GÉNÉRALE

Notre recherche a eu pour objectif de comparer les caractéristiques linguistiques des nouveautés langagières en Chine et en France, ainsi que leurs circulations et leurs évolutions de sens dans un environnement techno-relationnel, celui du microblog chinois *Sina Weibo* et de *Twitter* version française. Nous avons ainsi recueilli un corpus de listes des nouveautés langagières et un corpus de tweets contenant ces nouveautés langagières.

Les premières démarches ont été consacrées à aborder les caractéristiques contextuelles de la recherche qui concernent les aspects linguistiques, socioculturels et communicationnels qu'elle implique.

La présentation rapide du chinois contemporain nous a permis de mettre en relief : les deux systèmes d'écriture qui sont le système alphabétique pinyin et l'écriture en caractères simplifiés ; un aperçu synthétique de la situation sociolinguistique en Chine qui se caractérise par une diversité de dialectes régionaux utilisés en parallèle de la langue chinoise, le *mandarin*, dans la vie quotidienne ; les grandes différences linguistiques entre le chinois et le français résidant sur le caractère monosyllabique et morphologiquement inchangeable du chinois.

Les incidences de ces caractéristiques relevées sur les nouveautés langagières ont ensuite été constatées dans les analyses linguistiques de notre double corpus de listes (chinois et français). Cette démarche d'analyse que nous avons mise en place en nous appuyant sur la typologie des néologismes proposée par Pruvost et Sablayrolles puis Sablayrolles a eu pour but d'observer et de comparer les principaux procédés de création utilisés dans la formation des nouveautés langagières collectées dans les deux langues. Le corpus chinois pour cette analyse a été constitué à partir des propositions issues d'activités de sélection institutionnelles. Mais faute de pratique institutionnelle et médiatique régulière comparable, nous avons dû collecter les nouveautés langagières en français à travers une démarche différente tout en partant des mêmes principes de recueil (nouveau, fréquence, popularité).

Un bilan des procédés de création a été dressé et une assez forte disparité est apparue dans les proportions respectives des matrices concernées pour le chinois et le français. Certaines nouveautés langagières chinoises caractérisées par une forme mixte ou une mise en jeu de paronyme sont issues du recours alternatif à deux écritures chinoises. Les influences des

dialectes chinois sont manifestées tant sur le plan sémantique que sur le plan phonétique dans les processus néologiques chinois. En prenant compte à la fois des caractéristiques du caractère chinois (qui est invariable sur le plan morphologique) et des spécificités des nouveautés langagières elles-mêmes, nous avons choisi de privilégier l'aspect sémantique dans les procédés de création, et une grande part des réalisations chinoises ont été placées ainsi dans les matrices phraséologiques. Les principaux procédés repérés dans la formation des nouveautés langagières françaises sont l'affixation et la composition comme le laissent attendre les propriétés linguistiques du français.

En dépit de ces écarts, nous avons bien identifié de grandes tendances de création communes dans les nouveautés langagières chinoises et françaises. Les influences des langues étrangères sont constatées dans les deux langues avec des manifestations et degrés différentes. Une grande majorité des créations est issue des matrices internes, c'est-à-dire ce sont des créations propres à chaque langue. Les différences de proportion constatées quant aux matrices internes infirment toutefois en bonne partie notre hypothèse sur les grandes tendances de formation constatées.

À partir des éclaircissements que nous avons essayé d'apporter sur les notions relatives à l'Internet et à ses divers services et dispositifs, nous avons saisi les caractéristiques communicationnelles et sociotechniques du microblog. Cela nous a conduit à mettre en avant la nécessaire prise en compte à la fois de l'environnement technorelationnel, des rôles du cotexte et du contexte, des identités des usagers, ainsi que de l'engagement du chercheur dans la constitution et l'analyse d'un corpus discursif en contexte numérique.

Le nombre des nouveautés langagières collectées, leurs situations d'utilisation diversifiées, ainsi que la production dynamique et imprévisible sur les plateformes de microblog nous ont confronté à la question de la quantité et de la période d'observation à déterminer dans les premières démarches de la constitution des corpus de tweets. Nous sommes arrivés à une solution que nous avons expliquée et justifiée en entreprenant une analyse qualitative d'un échantillon significatif.

Les contraintes et les nuances d'affichage sur les deux microblogs nous ont amenées à nous appuyer sur le critère de popularité (classée selon chaque microblog) pour la sélection. D'autre part, nous avons porté notre attention sur les tweets contenant les nouveautés langagières hashtaguées, trouvés sur le fil des résultats et non sur les autres fils d'affichage. Cela nous

permet d'observer la circulation tout en restant sur le message du tweet et de lever les contraintes d'affichage. Pour recueillir ces tweets, en nous appuyant sur l'approche écologique proposée par Paveau dans le cadre de son analyse du discours numérique, nous avons opté pour la capture d'écran pour pouvoir observer et montrer les tweets dans leur environnement.

Nous avons établi une association privilégiée entre la fonction des hashtags et les nouveautés langagières qui nous a paru très productive et a supposé une mise en relation par contraste avec les éléments non hashtagués. Nous avons obtenu plusieurs résultats appréciables.

Nous considérons que les hashtags associés aux nouveautés langagières ont tout d'abord une fonction qui rejoint celle de « redocumentation » définie par Paveau (2017). Selon elle, ils ont une valeur de traces permettant de rechercher et classer les tweets contenant le même hashtag. Ces traces sont générées automatiquement suite à l'interaction de l'utilisateur avec le système. En même temps, à travers l'observation des relations entre les hashtags et les autres éléments du tweet, nous arrivons à distinguer des rôles différents du hashtag déterminés par les pratiques des usagers. Pour reprendre les catégories d'analyse du corpus chinois : un hashtag principal a plutôt une vocation thématique globale ; un hashtag auxiliaire n'est associé qu'à un aspect du contenu du tweet ; un hashtag de promotion est employé pour répondre à l'algorithme du système destiné à augmenter la visibilité d'un tweet. Pour le corpus français, les hashtags sont reliés à des objectifs pragmatiques variés : exprimer un point de vue, transmettre des informations, exhorter à des actions. Par ailleurs, nous avons également remarqué que la création d'un hashtag associé à une nouveauté implique que l'auteur du tweet a pris en compte à divers degrés ce fait langagier. Le hashtag peut être employé par l'auteur du tweet pour se joindre à la communauté construite autour des utilisations de nouveautés langagières, et aussi pour sensibiliser les futurs lecteurs à l'aspect néologique du segment hashtagué.

L'analyse contextualisée des occurrences authentiques que nous avons recueillies nous a menée à la conclusion que le microblog entraîne davantage une créativité sur la sémantique que sur la morphologie. Cette créativité sémantique est marquée par une forte interprétation personnelle dans les situations chinoise et par les emplois dans des contextes différents pour les cas français. En outre, nous avons constaté que les nouveautés langagières sont utilisées globalement dans leurs formes néologiques initiales. La nouvelle modalité d'expression caractérisée par la contrainte de longueur du message semble avoir peu d'influence sur l'évolution formelle de ces nouveautés langagières. Nous nous demandons si les changements sur l'aspect formel sont

entraînés par les pratiques des usagers et si c'est le choix d'adopter la forme hashtaguée des nouveautés langagières, autrement dit une forme strictement conditionnée, qui ne nous permet pas d'y repérer des créations formelles.

Les occurrences relevées montrent en outre que microblog constitue une chaîne importante qui favorise la circulation des nouveautés langagières et en témoigne. Au sein de la communauté établie autour d'un même hashtag, une utilisation répandue est manifestée par les identités variées des comptes et les types de discours pratiqués. Ainsi, la fonction techno-relationnelle du hashtag permet de faire circuler ces nouveautés langagières d'une communauté à une autre, ouvrant la possibilité d'atteindre une audience plus élargie. Par ailleurs, les hyperliens insérés et les messages rapportés dans les tweets indiquent qu'une diffusion massive au-delà de la plateforme de microblog s'est produite parallèlement.

À cet égard, l'environnement techno-relationnel du microblog et les pratiques des usagers nous semblent caractérisés par une circulation de masse des nouveautés langagières, accompagnée par une évolution sémantique. Donc, notre hypothèse portant sur les corpus de tweets est ainsi confirmée.

Notre travail s'est inscrit dans une perspective comparative. Or, nous n'avons pas pu collecter un corpus homogène issu d'une démarche comparable en dépit des mêmes principes appliqués lors de la constitution. Les démarches de collecte que nous avons résolues de manière empirique et adaptée n'échappent pas à la subjectivité. Nous nous demandons si cette subjectivité a eu une incidence sur les résultats obtenus. Certes, la taille des corpus, surtout celle des corpus de tweets est limitée par rapport à la masse énorme de données en accroissement constant. Nous n'avons pas pu mener une analyse quantitative qui nous aurait permis d'établir la traçabilité des nouveautés et d'observer diachroniquement les circulations. Mais l'analyse qualitative sur échantillons que nous avons développée nous a paru très riche d'enseignement et de nature à dégager des éléments significatifs

En ce qui concerne les analyses linguistiques, nous avons constaté que la typologie appliquée s'est révélée très opératoire pour interroger les corpus mais insuffisante pour apporter des réponses. Tant pour le corpus chinois que pour le corpus français elle appellerait des prolongements ou des aménagements. D'une part, nous avons estimé que les effets des dialectes en chinois et l'influence des français régionaux et non hexagonaux ne peuvent pas être mis en

évidences par les matrices proposées ; nous les avons mis dans une catégorie interne à part entière. D'autre part, les frontières entre les matrices ne sont pas absolues, les solutions mobilisées en privilégiant telle matrice plutôt qu'une autre pourrait amener à des résultats différents pour un même corpus. La prise en compte fine que permet la typologie doit déboucher sur la discussion au cas par cas.

Par ailleurs, durant les observations des corpus, surtout sur les tweets français, nous avons remarqué qu'une approche écologique ne se limite pas à la collecte et à l'analyse du corpus. Des exigences sont aussi à prendre en compte par les chercheurs. Il est nécessaire pour un chercheur de disposer d'une bonne connaissance du dispositif pour effectuer des démarches d'observation. En outre, l'analyse discursive demande non seulement une connaissance linguistique mais aussi culturelle et sociale, sinon, l'exploitation du corpus est difficile. En analysant les tweets, nous devenons en même temps les lecteurs « imprévisibles » du contenu. Nous nous trouvons dans un statut semblable à celui que décrit Matt Hills : le « Aca-fan »<sup>257</sup>. Nous nous demandons comment la position d'un chercheur « aca-fan » peut être prise en compte dans une approche écologique.

En outre, il faut être conscient en tant que chercheur tiers, du fait qu'il est impossible de recréer l'environnement d'observation des auteurs de tel ou tel tweet puisque nous n'avons pas les mêmes accès d'affichage et nous ne disposons pas du réseau d'interactions déterminé par les relations construites. Une analyse est toujours postérieure. Le choix d'analyser les nouveautés langagières hashtaguées nous évite d'avoir un corpus augmentatif et nous ne sommes pas non plus contrainte par le réseau d'amis. Pourtant, nous nous interrogeons : les emplois des nouveautés langagières non hashtaguées auraient-ils plus de caractéristiques morphologiquement marquées ?

En ce qui concerne nos recherches futures, nous envisageons différentes pistes susceptibles de prolonger le présent travail. Tout d'abord, dans les analyses linguistiques, nous nous sommes intéressée aux proportions de niveaux supérieurs des matrices mais nous aimerions traiter d'une manière plus approfondie nos corpus de listes en allant jusqu'au 5<sup>e</sup> niveau de la typologie. Dans un second temps, nous voudrions revenir aux stratégies de traduction des nouveautés

---

<sup>257</sup> L'aca-fan est une nouveauté langagière d'origine individuelle : c'est la troncature de *academic fan* et le terme a été créé par Matt Hills, un spécialiste de la culture des fans en 2002. L'expression signifie que le chercheur est lui-même un fan.



langagières qui ont été un des facteurs d'intérêt à l'origine de notre sujet. Comment transposer les aspects de nouveauté d'une langue source dans une langue d'arrivée ? Un tel processus de traduction/transposition engendre-t-il de nouvelles créations lexicales ? Quels sont les enjeux de sens selon les objectifs et destinations des traductions ? Dans un troisième temps, alors que nous avons appréhendé les pratiques des usagers selon leurs publications, nous nous demandons si une enquête menée auprès des usagers des microblog pourrait apporter des compléments ou un autre éclairage sur nos résultats obtenus. Enfin nous voudrions développer une analyse des discours numériques conforme aux principes que nous avons adoptés portant sur d'autres dispositifs numériques parmi ceux qui apparaissent et prospèrent (comme *Tiktok* par exemple) et ouverte sur d'autres modalités d'expression, voire sur de nouveaux concepts.

## BIBLIOGRAPHIE

ALLAM-IDDOU S. « Processus de la néologie dans « Tranche de vie » du *Quotidien d'Oran* : l'affixation ». *Multilinguales* [En ligne]. déc. 2015. n°6. Disponible sur : < <https://doi.org/10.4000/multilinguales.987> > (consulté le 30 novembre 2020)

ALLOING C., VANDERBIEST N. « La fabrique des rumeurs numériques. Comment la fausse information circule sur Twitter ? ». *Le Temps des medias*. mars 2018. Vol. n° 30, n°1, p. 105-123.

ANIS J. « Réseaux Humains / Réseaux Technologiques - Communication électronique scripturale et formes langagières ». [s.l.] : [s.n.], 2006. Disponible sur : < <http://rhrt.edel.univ-poitiers.fr/document.php?id=547> > (consulté le 30 juillet 2019)

ARENES C. « Les réseaux sociaux numériques pour chercheurs : quelles pratiques, quels enjeux ? ». avril 2015. Disponible sur : < <https://doi.org/10.1016/j.jemep.2015.04.010> > (consulté le 11 septembre 2018)

BACHIMONT B. « L'intelligence artificielle comme écriture dynamique : de la raison graphique à la raison computationnelle ». In : PETITOT J., FABBRI P. (éds.), *Au nom du sens*. Grasset, 2000. p. 290-319.

BARNES J. A. « Class and Committees in a Norwegian Island Parish ». *Human Relations* [En ligne]. fév. 1954. Vol.7, n°1, p. 39-58. Disponible sur : < <https://doi.org/10.1177/001872675400700102> > (consulté le 12 juin 2018)

BARON N. S. « Language of the Internet ». *The Stanford Handbook for Language Engineers* [En ligne]. 2003. n°59. Disponible sur : < <http://nl.ijs.si/janes/wp-content/uploads/2014/09/baron2003.pdf> > (consulté le 28 août 2019)

BEER D. « Social network(ing) sites...revisiting the story so far : A response to danah boyd & Nicole Ellison ». *Journal of Computer-Mediated Communication* [En ligne]. 2008. Vol.13, n°2, p. 516-529. Disponible sur : < <https://doi.org/10.1111/j.1083-6101.2008.00408.x> > (consulté le 05 mai 2018)

BERNHARD D., HUTCHINS-BOISSEAU M., GERARD C.-L., GRASS T., TODIRASCUCOURTIER A. (dir.). *La néologie en contexte: cultures, situations, textes*. Limoges : Lambert-Lucas, 2018.

BIGEY M. « Twitter et l'inscription de soi dans le discours ». *Les Cahiers du numérique*. nov. 2018. Vol. 14, n°3, p. 55-75.

BOISSEAU J.-B. « Qu'est ce que le web 2.0 : Modèles de conception et d'affaires pour la prochaine génération de logiciels ». In : *InternetActu.net* [En ligne]. [s.l.] : [s.n.], 2006.

Disponible sur : < <http://www.internetactu.net/2006/04/21/quest-ce-que-le-web-20-modeles-de-conception-et-daffaires-pour-la-prochaine-generation-de-logiciels/> > (consulté le 18 décembre 2018)

BOUCHARDON S. « L'écriture numérique : objet de recherche et d'enseignement », *Les Cahiers de la SFSIC*, juin 2014. p. 225-235.

BOULANGER J.-C. « Sablayrolles, J.-F. (2000) : La néologie en français contemporain. Examen du concept et analyse de productions néologiques récentes, Paris, Honoré Champion éditeur, coll. « Lexica », n° 4. » *Meta* [En ligne]. 2003. Vol. 48, n°3, p. 466-472. Disponible sur : < <https://doi.org/https://doi.org/10.7202/007612ar> > (consulté le 05 mai 2018)

BOUZIDI B. « Néologicité et temporalité dans le processus néologique ». *Synergies Algérie* n°9, 2010. p.27-36.

BOYADJIAN J. « Les usages politiques différenciés de Twitter ». *Politiques de communication*. mai 2016. Vol. n° 6, n°1, p. 31-58.

BOYD D. M., ELLISON N. B. « Social Network Sites: Definition, History, and Scholarship ». *Journal of Computer-Mediated Communication* [En ligne]. oct. 2007. Vol. 13, n°1, p. 210-230. Disponible sur : < <https://doi.org/10.1111/j.1083-6101.2007.00393.x> > (consulté le 12 avril 2018)

BUBAŠ G. « Computer mediated communication theories and phenomena: Factors that influence collaboration over the Internet ». *3<sup>rd</sup> CARNet Users Conference Zagreb*. 2001. Disponible sur : < <https://cuc.carnet.hr/cuc2001/papers/b1.pdf> > (consulté le 12 juin 2018)

CALVET L.-J. *La guerre des langues et les politiques linguistiques*. Paris : Hachette, 1987.

CARDON D. *Culture numérique*. Paris : Les presses de Sciences Po, « Hors collection», 2019. Chap. « Identité en ligne » p. 176-188.

CARDON D. « Réseaux sociaux de l'Internet ». *Communications* n°88. janv. 2011. p. 141-148.

CARTIER E. « Emprunts en français contemporain : étude linguistique et statistique à partir de la plateforme Néoveille ». In : *L'emprunt en question(s) : conceptions, réceptions, traitements lexicographiques* [En ligne]. [s.l.] : [s.n.], 2019. Disponible sur : < <https://hal.archives-ouvertes.fr/hal-02537344> > (consulté le 28 mars 2021)

CASILLI A. « Être présent en ligne : culture et structure des réseaux sociaux d'Internet ». *Idées économiques et sociales* [En ligne]. déc. 2012. n°169, p. 16-29. Disponible sur : < <https://doi.org/10.3917/idee.169.0016> > (consulté le 3 mai 2016)

CASILLI A. *Cultures du numérique*. Paris : Seuil, 2011.

CAVAZZA F. Panorama des médias sociaux 2018. *Le Blog de CAVAZZA F.* [En ligne]. 05 mai 2018. Disponible sur : < <https://fredcavazza.net/2018/05/05/panorama-des-medias-sociaux-2018/> > (consulté le 28 mai 2019)

CHE F. « 近十余年来汉语网络新词语研究述略 ». *重庆工商大学学报(社会科学版)*. 2012. Vol. 32, n°3, p. 102-113.

CHEN C. L. « 从失范走向规范——关于网络语言影响及规范策略的思考 ». *学术界*. 2011. n°4, p. 195-202 et p.288-289.

CHEN H. Y. *微博用户的自我呈现和影响力分析 ——以新浪微博达人为例* [En ligne]. Mémoire de master. [s.l.] : 暨南大学, 2012. Disponible sur : < <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD2012&filename=1012031909.nh&v=xZGjR2b%25mmd2Fd4jQxvyBvoqTPENp0v4%25mmd2BcPJY6geFTowX718s7zKBPjywnIkQUreCVvfr> > (consulté le 2 mars 2021)

CHEN J. J. « 网络语言的互联网群体传播本质及互动 ». *当代传播*. 2017. n°4, p. 54-64.

CHEN M. Z., BAI J. H. « 汉语网络语言研究的回顾、问题与展望 ». *湖南师范大学社会科学学报*. 2012. Vol. 41, n°3, p. 130-134.

CHEN T. *网络语言的变异类型及原因探析* [En ligne]. Mémoire de master. [s.l.] : 华中师范大学, 2017. Disponible sur : < <http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201801&filename=1017266744.nh&v=MDIyNzZWcjdnVjYyNkdiRytHTmJJcTVFYlBJUjhlWDFMdxhZUzdEaDFUM3FUclldNMUZyQ1VSTE9mYnVadUZ5L2c=> > (consulté le 1 septembre 2019)

CHEN X. H. « 网络语言中的语词类型、特点及趋势 ». *南京晓庄学院学报*. 2001. n°3, p. 62-68.

CHU Y. F. « 近二十年网络语言研究综观 ». *浙江传媒学院学报*. 2016. n°5, p. 104-112.

CONSTANTINIDES E., FOUNTAIN S. J. « Web 2.0: Conceptual foundations and marketing issues ». *Journal of Direct, Data and Digital Marketing Practice* [En ligne]. janv. 2008. Vol. 9, n°3, p. 231-244. Disponible sur : < <https://doi.org/10.1057/palgrave.dddmp.4350098> > (consulté le 10 mars 2018)

COUTANT A., STENGER T. « Processus identitaire et ordre de l'interaction sur les réseaux sociaux numériques ». *Les Enjeux de l'information et de la communication*. Vol. n°1, 2010. p. 45-64.

COUTANT A., STENGER T. « Les configurations sociotechniques sur le Web et leurs usages : le cas des réseaux sociaux numériques ». In : *7ème Colloque du chapitre français de l'ISKO, Intelligence collective et organisation des connaissances* [En ligne]. Lyon, France : [s.n.], 2009. p. 27-34. Disponible sur : < <https://hal.archives-ouvertes.fr/hal-00458327> > (consulté le 10 septembre 2018)

COUTANT A., STENGER T. « L'avènement des "médias sociaux" ? ... Une histoire de participation ». *Le Temps des médias. Revue d'histoire*. 2012. n°18, p. 76-86.

COUTANT A., STENGER T. « Médias sociaux : clarification et cartographie pour une approche sociotechnique ». *Décisions Marketing* n°70, [En ligne]. juin 2013. p. 107-117. Disponible sur : < <https://doi.org/10.7193/dm.070.107.117> > (consulté le 5 avril 2018)

CRYSTAL D. *Language and the Internet* [En ligne]. [s.l.] : [s.n.], 2001. Disponible sur : < <https://doi.org/10.1017/CBO9780511487002> > (consulté le 27 septembre 2018)

CRYSTAL D. *Internet linguistics : a student guide*. Milton Park, Abingdon ; New York, N.Y : Routledge, 2011.

DAI X. J. *网络语言的性别差异研究* [En ligne]. Mémoire de master. [s.l.] : 青岛大学, 2016. Disponible sur : < <http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201701&filename=1016772144.nh&v=MjAzMTVUcldNMUZyQ1VSTE9mYnVadUZ5N21WYjNBVkYyNkdMUy9ITkRjcTVFYlBJUjhlWDFMdXhZUZdEaDFUM3E=> > (consulté le 1 septembre 2019)

DEGENNE A. « Retour à l'analyse des réseaux sociaux ». *Hermès, La Revue*. 2011. n°59, p. 39-42.

DENG Q. « 文化融合视阈下的网络语言 ». in : *首届华中地区新闻与传播学科研究生学术论坛*. 2010.

DEROY L. « Chapitre I. Un emprunt sans restitution ». in : *L'Emprunt linguistique* [En ligne]. Liège : Presses universitaires de Liège, 2013. p. 17-22. Disponible sur : < <http://books.openedition.org/pulg/674> > (consulté le 28 mars 2020)

DUBOIS J., GIACOMO-MARCELLESI M., GUESPIN, L. *Dictionnaire de linguistique*. Paris : Larousse, 2001. ISBN : 2-03-532047-X.

DURY P., DE HOYOS J. C., MAKRI-MOREL J., MANIEZ F., RENNER V., VILLAR DIAZ M. B. *La néologie en langue de spécialité: détection, implantation et circulation des nouveaux termes* [En ligne]. [s.l.] : [s.n.], 2014. Disponible sur : < <https://hal.archives-ouvertes.fr/hal-01915611> > (consulté le 22 août 2019)

ELLISON N., THIERRY A. « Réseaux sociaux, numérique et capital social ». *Hermès, La Revue*. 2011. Vol. n° 59, n°1, p. 21-23.

ELLISON N. B., BOYD D. M. « Sociality Through Social Network Sites » [En ligne]. [s.l.]. In : DUTTON W. H. (dir.). *The Oxford Handbook of Internet Studies*, Oxford University Press, 2013. Disponible sur : < <https://doi.org/10.1093/oxfordhb/9780199589074.013.0008> > (consulté le 28 mars 2017)

FILIP A., GHARBI I., GIZARD S., HONNART J., INDRECAN C. « Le Web 2.0, Rapport interne du Département Télécommunications, Services et Usages ». [s.l.] : [s.n.], 2008. Disponible sur : < <http://julienhonnart.free.fr/CGT/Le%20Web%202.0.pdf> > (consulté le 22 mai 2019)

FORSE M. « Définir et analyser les réseaux sociaux ». *Informations sociales*. juil. 2008. n°147, p. 10-19.

GAO S. *La traduction audiovisuelle du français au chinois en contexte de communication internet: le cas de « Bienvenue chez les c'htis » et le fansub*. mémoire de master en Sciences du langage option Discours texte communication. Besançon : Université de Franche-Comté, 2013.

GARDIN B., LEFEVRE G., TARDY M., MORTUREUX M.-F. « A propos du "sentiment" néologique ». *Langages* [En ligne]. 1974. Vol. 8, n°36, p. 45-52. Disponible sur : < <https://doi.org/10.3406/lgge.1974.2273> > (consulté le 8 juillet 2018)

GAUTIER L. « La "mise en formule" par les hashtags : analyse de nouvelles formes « ultra-brèves » dans les tweets de campagne des européennes (2014) ». In : *Les formes brèves de l'imprimé au numérique (XIX-XXIè siècles)*. Dijon, France : [s.n.], 2017. Disponible sur : < <https://halshs.archives-ouvertes.fr/halshs-01491690> > (consulté le 4 décembre 2019)

GEORGES F. « Représentation de soi et identité numérique ». *Reseaux*. avr. 2009. Vol. n° 154, n°2, p. 165-193.

GERVAIS J.-F. *Web 2.0 - Les internautes au pouvoir: Blogs, Réseaux sociaux, Partage de vidéos, Mashups...* [s.l.] : Dunod, 2007.

GIRARD A., FALLERY B. « Réseaux Sociaux Numériques : revue de littérature et perspectives de recherche ». in : *14ème Congrès de l'AIM*. Marrakech, Maroc, 10-12 juin 2009. Disponible sur : < [http://aim.asso.fr/upload/Colloques\\_AIM/AIM%202009/aim2009\\_submission\\_066.pdf](http://aim.asso.fr/upload/Colloques_AIM/AIM%202009/aim2009_submission_066.pdf) > (consulté le 11 septembre 2018)

GUILBERT L. « Théorie du néologisme ». *Cahiers de l'Association internationale des études francaises* [En ligne]. 1973. p. 9-29. Disponible sur : < <https://doi.org/https://doi.org/10.3406/caief.1973.1020> >

HAO Y. « 网络语言的变异现象分析 ». *语文建设*. 2017. n°24, p. 75-76.

HONG R. X. *微博环境下女性话语权的现阶段提升与再提升困局研究 —— 以新浪微博为例* [En ligne]. mémoire de master. [s.l.] : 东北财经大学, 2015. Disponible sur : < <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201501&filename=1015549966.nh&v=eqaR7YEBJiY0iehSpAE9DRKsl8%25mmd2FQ1wtKLTCpDRi0MJ7TsOHu6gj0lemWvXG4wCK%25mmd2F> > (consulté le 2 mars 2021)

HU B. Y. « 网络语言对语文教育的冲击与治理 ». *语文建设* [En ligne]. 2016. Disponible sur : < <http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CJFQ&dbname=CJFDLAST2016&filename=YWJS201614003&v=MDk2MDFEckJmYkc0SDImTnE0OUZaNF14ZVgxTHV4WVM3RGgxVDNxVHJXTTFGckNVUkxPZmJlWnVGeTduV3J6S1A=>> > (consulté le 1 septembre 2019)

HU L., LIU Y., YANG C. L. « 网络语言二十年发展综述 ». *湖南大学学报(社会科学版)*. 2014. Vol. 28, n°05, p. 136-141.

- HUANG A. M. « 网络语言中的日语外来语 ». *湖北广播电视大学学报*. 2010. p. 113-114.
- HUANG Y. « 微博的媒体特征及传统媒体的应对 ». *东南传播*. 2011. n°1, p. 84-86.
- HUMBLEY J., SABLAYROLLES J.-F. (dir.). *Neologica : revue internationale de néologie* n°1. Paris : Éditions Classiques Garnier, 2007.
- ISSAC F. « Cybernéologisme : Quelques outils informatiques pour l'identification et le traitement des néologismes sur le web ». *Langages*. 2011. Vol. n°183, n°3, p. 89-104.
- JACQUES G. « La morphologie du sino-tibétain ». in : *La linguistique comparée aujourd'hui*, EHESS, Mar 2006, Paris, France. Disponible sur : < halshs-00173948 > (consulté le 2 mars 2015)
- JENKINS H. *La culture de la convergence: des médias au transmédia*. traduit de l'anglais par JAQUET C. Paris : Armand Colin ; INA, coll. Médiacultures. 2013.
- JENKINS H. « Participatory Culture: From Co-Creating Brand Meaning to Changing the World ». *NIM Marketing Intelligence Review* [En ligne]. 2014. Vol. 6, n°2, p. 34-39. Disponible sur : < <https://doi.org/10.2478/gfkmir-2014-0096> > (consulté le 8 avril 2016)
- JEUGE-MAYNAR I., FLORENT J. (dir.). *Le petit Larousse illustré [2012] : en couleurs : 90 000 articles, 5 000 illustrations, 354 cartes, chronologie universelle*. Edition limitée. Paris : Larousse, 2011. ISBN : 978-2-03-584090-5.
- JIANG W. T. « 网络语言的分类及其特征 ». *江西省语言学会 2006 年年会论文集*. 2006.
- JULLIARD V. « Les apports de la techno-sémiotique à l'analyse des controverses sur Twitter ». *Hermès, La Revue*. déc. 2015. Vol. n° 73, n°3, p. 191-200.
- KAPLAN A. M. « Twitter ou le pouvoir de 140 caractères ». *L'Expansion Management Review*. 2011. Vol. n° 140, n°1, p. 104-113.
- KAPLAN A. M., HAENLEIN M. « Users of the world, unite! The challenges and opportunities of Social Media ». *Business Horizons* [En ligne]. janv. 2010. Vol. 53, n°1, p. 59-68. Disponible sur : < <https://doi.org/10.1016/j.bushor.2009.09.003> > (consulté le 13 mars 2014)
- KARLGREN B. *Etudes sur la phonologie chinoise*. Upsala : K.W. Appelberg, 1915.
- LAPASSADE G. *Observation participante* [En ligne]. [s.l.] : Érès, 2002. Disponible sur : < <http://www.cairn.info/vocabulaire-de-psychosociologie--9782749206851-page-375.htm> > (consulté le 22 février 2021)
- LE DEUFF O. « Le succès du web 2.0 : histoire, techniques et controverse ». fév. 2007. Disponible sur : < [https://archivesic.ccsd.cnrs.fr/sic\\_00133571](https://archivesic.ccsd.cnrs.fr/sic_00133571) > (consulté le 3 décembre 2018)

LECOLLE M. «Néologie sémantique et néologie catégorielle : quelques propositions». *Cahiers de Lexicologie*, Centre National de la Recherche Scientifique, 2012, p. 81-104. Disponible sur : < fhalshs-00785314f > (consulté le 13 mars 2014)

LI Y. C. « 从网民结构的特殊性论网络语言的发展 ». *运城学院学报* [En ligne]. 2011. Disponible sur : < <http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CJFQ&dbname=CJFD2011&filename=YCGD201101023&v=MDkxMDdmYnVadUZ5N2xWcnJOUEM3TWFyRzRIOURNcm85SFo0UjhlWDFMdXhZUZdEaDFUM3FUclldNMUZYQ1VSTE8=> > (consulté le 1 septembre 2019)

LIANG J. « 网络语言对青少年教育的影响研究 ». *语文建设* [En ligne]. 2014. Disponible sur : < <http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CJFQ&dbname=CJFDLAST2015&filename=YWJS201432061&v=MDUzNTBYUHJZOURaWVI4ZVgxTHV4WVM3RGgxVDNxVHJXTTFGckNVUkxPZmJlWnVGeTduV3JySIBeckJmYkc0SDk=> > (consulté le 1 septembre 2019)

LIU Q. R. « 大众文化语境下网络新词的传播学分析 ». *语文知识*. 2012. n°4, p. 30-32.

LIU Q. S., CHE L., LIU X., WANG X. L., WANG Y. Q., CHEN J. M. « 俄语网络语言的本体性研究 ». *文化创新比较研究*. 2018. Vol. 2, n°17, p. 64-66.

LIU S. X. *汉语描写词汇学-(重排本)* [En ligne]. [s.l.] : 商务印书馆, 2005. 415 p. Disponible sur : < <http://m.bookschina.com/6417402.htm> > (consulté le 10 mars 2021)

LIU T. M., SHI H. Y. « 网络语言研究综述 ». *晋中学院学报*. 2015. Vol. 32, n°05, p. 109-114.

LIU W. y., LI K. « 基于批评性话语分析的网络语言暴力研究框架 ». *东北师大学报*. 2017. p. 119-124.

LIU Y. « 日语模因在汉语网络语言中的复制——从“お姉さん”到“御姐” ». *重庆邮电大学学报(社会科学版)* [En ligne]. 2011. Disponible sur : < <http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CJFQ&dbname=CJFD2011&filename=CQYD201104027&v=MTU1OTJGeTdsVTcvQUppelNhcck0SDIETXE0OUhZNF14ZVgxTHV4WVM3RGgxVDNxVHJXTTFGckNVUkxPZmJlWnU=> > (consulté le 1 septembre 2019)

LIU Y. « 网络语言研究综述 ». *晋中学院学报*. 2015. Vol. 32, n°05, p. 109-114.

LOMBARD A., HUYGHE R. « Catégorisation comme néologisme et sentiment des locuteurs ». *Langue française*. oct. 2020. Vol. N° 207, n°3, p. 123-138.

LONGHI J. « Essai de caractérisation du tweet politique ». *L'Information Grammaticale* [En ligne]. 2013. Vol. 136, n°1, p. 25-32. Disponible sur : < <https://doi.org/10.3406/igram.2013.4238> > (consulté le 18 septembre 2018)

LONGHI J. « Explorer des corpus de tweets : du traitement informatique à l'analyse discursive complexe ». *Corpus* [En ligne]. janv. 2020. n°20. Disponible sur : <



<http://journals.openedition.org/corpus/4567> > (consulté le 5 février 2018)

LV S. *近代汉语指代词·序*. 商务印书馆, 1984.

MAINGUENEAU D. « Chapitre 4 - Genres de discours et web : existe-t-il des genres web ? » in BARATS C. (dir.). *Manuel d'analyse du web en Sciences Humaines et Sociales*. Paris : Armand Colin, 2013. p. 74-98.

MEJRI S. « Néologie et unité lexicale : renouvellement théorique, polylexicalité et emploi ». *Langages*. 2011. Vol. n° 183, n°3, p. 25-37.

MENGER P.-M., PAYE S. (éd.). *Big data et traçabilité numérique : Les sciences sociales face à la quantification massive des individus* [En ligne]. Paris : Collège de France, 2017. coll. Conférences. Disponible sur : < <http://books.openedition.org/cdf/4987> > (consulté le 17 juin 2019) ISBN : 978-2-7226-0467-4.

MERCANTI-GUERIN M. « Analyse des réseaux sociaux et communautés en ligne : quelles applications en marketing ?, Abstract ». *Management & Avenir* [En ligne]. 17 mai 2010. n°32, p. 132-153. Disponible sur : < <https://doi.org/10.3917/mav.032.0132> > (consulté le 20 février 2017)

MERCIER L.-S. *Néologie ou Vocabulaire de mots nouveaux : à renouveler, ou pris dans des acceptions nouvelles ([Reprod.]) / par L. S. Mercier* [En ligne]. [s.l.] : [s.n.], 1801. Disponible sur : < <https://gallica.bnf.fr/ark:/12148/bpt6k50792d> > (consulté le 28 février 2021)

MERCKLE P. « Les réseaux sociaux, les origines de l'analyse des réseaux sociaux ». 2004. Disponible sur : < [http://archive.wikiwix.com/cache/?url=http%3A%2F%2Feco.ens-lsh.fr%2Fsociales%2Freseaux\\_merckle\\_03\\_origines.pdf](http://archive.wikiwix.com/cache/?url=http%3A%2F%2Feco.ens-lsh.fr%2Fsociales%2Freseaux_merckle_03_origines.pdf) > (consulté le 11 septembre 2018)

MO X. J. « 微文化背景下网络热词话语分析 ». *当代教育实践与教学研究*. 2019. n°01, p. 233-234.

NIEL X., ROUX D. « La révolution internet ». *Que sais-je?* 2010. 2<sup>e</sup> éd., n°3809.

OTA T. *汉语史通考*. traduit du japonais par JIANG L. S., BAI W. G. [s.l.] : 重庆出版社, 1991.

PAILLE P., MUCCHIELLI A. « Chapitre 8 - Les techniques d'analyse contextualisantes, configurationnelles et systémiques ». in : *L'analyse qualitative en sciences humaines et sociales*. Paris : Armand Colin, 2012. p. 157-182.

PANNIER E. « L'analyse des réseaux sociaux : théories, concepts et méthodologies ». *Tạp chí xã hội học [Sociological Review of Vietnam]*. 1 janvier 2008. Vol. 4, p. 100-114.

PAVEAU M.-A. *L'analyse du discours numérique : Dictionnaire des formes et des pratiques*. Paris : Hermann, 2017.

PAVEAU M.-A. « Technodiscursivités natives sur Twitter. Une écologie du discours numérique ». *Épistémé (Revue internationale de sciences humaines et sociales appliquées, Séoul)*. juil. 2013. Vol. 9, p. 139-176.

PAVEAU M.-A. « Genre de discours et technologie discursive. Tweet, twittécriture et twittérature ». janv. 2012. Disponible sur : < <https://hal.archives-ouvertes.fr/hal-00824817> > (consulté le 5 mars 2020)

PAVEAU M.-A. *Hashtag [Dictionnaire]* [En ligne]. *Technologies discursives*. 2013. Disponible sur : < <https://technodiscours.hypotheses.org/488> > (consulté le 17 juillet 2020)

PAVEAU M.-A. « Mais où est donc le sens ? Pour une linguistique symétrique ». in : *Res per nomen 2* [En ligne]. Reims, France : [s.n.], 2009. p. 21-31. Disponible sur : < <https://hal.archives-ouvertes.fr/hal-00477257> > (consulté le 3 septembre 2019)

PAVEAU M.-A. *Les prédiscours: Sens, mémoire, cognition*. [s.l.] : Presses Sorbonne Nouvelle, 2006.

PRUVOST J., SABLAYROLLES J.-F. *Les néologismes*. 3<sup>e</sup> éd. Paris : « Que sais-je ? » PUF. 2016. p. 91-115.

PRUVOST J., SABLAYROLLES J.-F. *Les néologismes*. 2<sup>e</sup> éd. Paris : « Que sais-je ? » P.U.F. 2012.

PRUVOST J., SABLAYROLLES J.-F. *Les néologismes*. 1<sup>e</sup> éd. Paris : « Que sais-je ? » PUF. 2003.

QI J. T. « 现代汉语流行语生成源起考略及其英译 ». *大连大学学报*. 2007. p. 138-140.

RAUS R. « Productivité de cyber et hyper dans le lexique français d'Internet ». *La linguistique*. 2001. Vol. 37, n°2, p. 71-88.

REBOUL-TOURE S. « 2016, "Les tweets : un lieu pour la créativité lexicale?" ». in : JACQUET-PFAU C., SABLAYROLLES J.-F. (dir.). *La fabrique des mots français*. Lambert-Lucas [En ligne]. [s.l.] : [s.n.], 2016. Disponible sur : < <https://hal-univ-paris3.archives-ouvertes.fr/hal-01599333> > (consulté le 17 décembre 2019)

REBOUL-TOURE S. « 2010, "Définition terminologique et définition lexicographique : les néologismes de l'internet " ». *Publifarum* [En ligne]. 30 oct. 2010. Disponible sur : < <https://hal-univ-paris3.archives-ouvertes.fr/hal-01599553> > (consulté le 11 janvier 2021)

REY-DEBOVE J., RAY A. (dir.). *Le petit Robert : dictionnaire alphabétique et analogique de la langue française*. Nouvelle édition millésime 2012. Paris : le Robert, 2011. ISBN : 978-2-84902-898-8.

RONG J. P. *汉语网络语言英语语法现象研究* [En ligne]. mémoire de master. [s.l.] : 暨南大学, 2012. Disponible sur : < <http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201501&filena> >

me=1012031674.nh&v=MTMxNTRFYlBJUjhlWDFMdXhZUZdEaDFUM3FUclDNMUZyQ1VSTE9mYnVadUZ5L2hXNzdMVkYyNkhMTzdIOWZMcTU= > (consulté le 1 septembre 2019)

RUAN L. H., HUANG J. « 社交媒体广告营销研究——以新浪微博为例 ». *中小企业管理与科技(上旬刊)*. 2018. n°8, p. 109-112.

RUI D. R. « 网络语言对当代语言教学的负面影响 ». *时代文学(双月版)*. 2006. n°02, p. 105-107.

SABLAYROLLES J.-F. « Néologismes : une typologie des typologies ». *Cahier du CIEL*, 1996. Disponible sur : < [https://www.eila.univ-paris-diderot.fr/\\_media/recherche/clillac/ciel/cahiers/96-97/1sablayrollestexte.pdf](https://www.eila.univ-paris-diderot.fr/_media/recherche/clillac/ciel/cahiers/96-97/1sablayrollestexte.pdf) > (consulté le 15 septembre 2017)

SABLAYROLLES J.-F. *Comprendre la néologie: conceptions, analyses, emplois*. Limoges : Lambert-Lucas. 2019.

SABLAYROLLES J.-F. *La néologie en français contemporain: examen du concept et analyse de productions néologiques récentes*. Paris : H Champion. 2000.

SABLAYROLLES J.-F. « Les inflexions de trente ans de recherche néologique ». in : TALLARICO G. L., HUMBLEY J., JACQUET-PFAU C. (éd.). *Nouveaux horizons pour la néologie en français: hommage à Jean-François Sablayrolles*. Limoges : Lambert-Lucas. 2020.

SABLAYROLLES J.-F., JACQUET-PFAU C. « Les emprunts : du repérage aux analyses. Diversité des objectifs et des traitements ». *Neologica : revue internationale de la néologie* n°2, Paris : Garnier. 2008. p.19-38.

SAUTEDE É. « Les réseaux sociaux numériques en Chine : une constellation de petits mondes, Social networking sites in China : a constellation of small worlds ». *Hermès, La Revue*. 2011. n°59, p. 151-158.

SHEN Y., PU C. H. « 网络流行语的起源、词义分析及分类——以 2014 年~2016 年“网络用语”为例 ». *Modern Linguistics* [En ligne]. mars 2018. Vol. 6, p. 335. Disponible sur : < <https://doi.org/10.12677/ML.2018.62038> > (consulté le 1 septembre 2019)

SOUCHIER E. « L'écrit d'écran, pratiques d'écriture & informatique ». *Communication & Langages* [En ligne]. 1996. Vol. 107, n°1, p.105-119. Disponible sur : < <https://doi.org/10.3406/colan.1996.2662> > (consulté le 12 décembre 2020)

STENGER T. (éd.). *Ces réseaux numériques dits sociaux*. Paris : CNRS Éd, 2011.

STEUCKARDT A. « Néologie : la fabrique d'un terme ». in : JACQUET-PFAU C., SABLAYROLLES J.-F. (éd.). *La fabrique des mots français*. Limoges : Lambert Lucas, 2016. p. 23-37. Disponible sur : < <https://hal.archives-ouvertes.fr/hal-01827816> > (consulté le 28 février 2021)

SUN C. H. « 论网络语言对现代汉语发展的影响 ». *鲁东大学学报 (哲学社会科学版)*. 2007. n°02, p. 105-107.

SUN C. Y., WANG H. Y. « 英语对汉语环境下的网络语言的影响 ». 2017. Disponible sur : < <http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CJFQ&dbname=CJFDLAST2017&filename=ZYJP201715067&v=MDIzNTFOcW85RFk0UjhlWDFMdXhZUZdEaDFUM3FUcldNMUZyQ1VSTE9mYnVadUZ5L2hVTC9MUHpUQmZyRzRIOWI=> > (consulté le 1 septembre 2019)

SUN R. X. « 2013 年汉语网络新词造词法研究 ». *牡丹江大学学报*. 2015. Vol. n°24, n°1, p. 17-21. Disponible sur : < [http://www.cssn.cn/zt/zt\\_xkzt/xwcbxzt/xwxycbx/xczjytj/201501/P020150129569220249668.pdf?COLLCC=2525190460&](http://www.cssn.cn/zt/zt_xkzt/xwcbxzt/xwxycbx/xczjytj/201501/P020150129569220249668.pdf?COLLCC=2525190460&) > (consulté le 19 mai 2020)

TALLARICO G. « À propos des emprunts néologiques en français: émergence, intégration et aspects socio-pragmatiques ». in : *Nouveaux horizons pour la néologie en français : hommage à Jean-François*. Limoges : Lambert-Lucas, 2020. p. 34-51.

THELWALL M. « Social network sites: Users and uses ». *Advances in Computers*. 2009. p. 19-73.

THIAULT F. « Le produsage des hashstags sur Twitter, une pratique affiliative ». *Questions de communication* [En ligne]. déc. 2015. n°28, p. 65-79. Disponible sur : < <https://doi.org/10.4000/questionsdecommunication.10043> > (consulté le 20 septembre 2018)

TISSERON S. « Les nouveaux réseaux sociaux sur internet, The new social networks on the internet ». *Psychotropes* [En ligne]. oct. 2011. Vol. 17, n°2, p. 99-118. Disponible sur : < <https://doi.org/10.3917/psyt.172.0099> > (consulté le 15 septembre 2016)

VITALI ROSATI M. « Une philosophie du numérique ». *Sens Public* [En ligne]. nov. 2011. Disponible sur : < <http://www.sens-public.org/article882.html> > (consulté le 31 juillet 2019)

VITALI-ROSATI M. *Pratiques de l'édition numérique* [En ligne]. Montréal : Presses de l'Université de Montréal, 2014. Disponible sur : < <http://books.openedition.org/pum/306> > (consulté le 31 juillet 2019)

WANG Y. *基于话语分析的微博语言暴力研究* [En ligne]. mémoire de master. [s.l.] : 南京师范大学, 2017. Disponible sur : < <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201801&filename=1017281554.nh&v=cDLpzObNogHxOfhIixLgRg2qld3cBbWmzbhNjhpFuJ6ND2goq%25mmd2BbwqVIX6wxCVHv6> > (consulté le 2 mars 2021)

WANG L. *汉语史稿·汉语史的分期*. 2013<sup>e</sup> éd. [s.l.] : 中华书局, 1980.

WANG S. Y. « 近十年我国网络流行语研究综述 ». *重庆工商大学学报 (社会科学版)*. 2012. Vol. 29, n°05, p. 7-12.

XIA A. Q. « 新时期网络语言特点及与传统语言的对比 ». *长春教育学院学报*. 2013. Vol. 29, n°13, p. 39-40.

XIAO H. *语用顺应论视角下的中国网络语言词汇特点研究* [En ligne]. mémoire de master. [s.l.] : 广东商学院, 2012. Disponible sur : < <http://kns.cnki.net/kcms/detail/frame/list.aspx?dbcode=CMFD&filename=1012032479.nh&dbname=CMFD201301&RefType=1&vl=> > (consulté le 1 septembre 2019)

XU Z. H. *Le néologisme et ses implications sociales = Xinci yu shehui hudong guanxi yanjiu / Xu Zhenhua ; préface de Huang Jianhua*. Paris : L'Harmattan, coll. Logiques sociales, 2001. 174 p. ISBN : 978-2-7475-0465-2.

YANG W. Q. (dir.). *现代汉语*. 重庆大学出版社, 2010.

YAN Y. « 浅析网络语言反规范现象及思考——以抖音评论中的语言为例 ». *艺术科技*. 2018. p. 291-292.

YANOSHEVSKY G. « Les réseaux sociaux et l'échange entre l'homme politique et les internautes : le cas de Facebook après les élections présidentielles en France ». *Argumentation et Analyse du Discours* [En ligne]. oct. 2010. n°5. Disponible sur : < <https://doi.org/10.4000/aad.1008> > (consulté le 5 mars 2020)

YI M. *汉泰网络语言比较研究* [En ligne]. mémoire de master. [s.l.] : 天津大学, 2018. Disponible sur : < <http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CMFD&dbname=CMFD201902&filename=1019900812.nh&v=MjgwNjkrVnZGaXJnV3IzUFZGMjZGN3E0SHRuTnJaRWJQSVI4ZVgxTHV4WVM3RGgxVDNxVHJXTTFGckNVUkxPZmI=> > (consulté le 29 août 2019)

YU M. *新语言革命* [En ligne]. mémoire de master. [s.l.] : 对外经济贸易大学, 2002. Disponible sur : < <http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CMFD&dbname=CMFD9904&filename=2002060742.nh&v=MDcwOTNaRWJQSVI4ZVgxTHV4WVM3RGgxVDNxVHJXTTFGckNVUkxPZmIrUm5GQ3ZoVnJ6S1YxMjdITE8rSHRiSXI=> > (consulté le 28 août 2019)

ZAMMAR N. « Réseaux Sociaux numériques: essai de catégorisation et cartographie des controverses ». Thèse de doctorat en Sciences l'Information et de la communication, Université Rennes 2, 2012. p. 377. Disponible sur : < <https://tel.archives-ouvertes.fr/tel-00687906/document> > (consulté le 28 mai 2017)

ZHANG Y. X. *新浪微博传播机制研究 ——以“围脖女王”姚晨微博为例* [En ligne]. mémoire de master. [s.l.] : 西南大学, 2011. Disponible sur : < <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD2011&filename=1011113564.nh&v=9iVez%25mmd2FaFViDSJu8ksFvedAJAYHY7SROLt%25mmd2B492AaDWuK%25mmd2FD6QvVhvKwdqL6OIM85WQ> > (consulté le 2 mars 2021)

ZHANG C. « 浅谈新时代网络语言对汉语言文学发展的影响 ». *佳木斯职业学院学报*

[En ligne]. 2017. Disponible sur : <  
<http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CJFQ&dbname=CJFDLAST2017&filename=JMSJ201709090&v=MDUwMTdXTTFGckNVUkxPZmlrVnZGaXpsVmJ6S0x5RFlaTEc0SDliTXBvOU1aSVI4ZVgxTHV4WVM3RGgxVDNxVHI=>> (consulté le 29 août 2019)

ZHANG Y. H. « 网络语言的词汇语法特征 ». *中国语文*. 2007. n°06, p. 531-535.

ZHAO L. *人类学视野下的微博互动 ——以新浪微博为例* [En ligne]. mémoire de master. [s.l.] : 陕西师范大学, 2012. Disponible sur : <  
<https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&dbname=CMFD2012&filename=1012424263.nh&v=DY9ueqTnQNQapDiQiRX7LjEyN3o%25mmd2FIaz0JKM%25mmd2FQek8rkn5EKJf%25mmd2BK3f5W%25mmd2FzfrtpM1rr>> (consulté le 2 mars 2021)

ZHOU H. Z. « 一门崭新的语言学科——网络语言学 ». *科学*. 2000. n°9, p. 62-62.

ZHOU M. « 语言文字规范的新领域与新对策——也谈网络语言的规范问题 ». in : *第四届全国语言文字应用学术研讨会论文集*. 2005. Disponible sur : <  
<http://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CPFD&dbname=CPFD9908&filename=JBYW200512001045&v=MTY4NDllc0xDUk5LdWhkaG5qOThUbmpxcXhkRWVNT1VLcmImWnU5dkhpdnNVTHpOS0Y0Ukx5L1NIYkc0SHRUTnJZOUZa>> (consulté le 25 mars 2019)



## LISTE DES TABLEAUX

Tableau 1 : Les principales écoles de périodisation du développement de la langue chinoise.....	24
Tableau 2 : Catégories et distribution géographique des dialectes en Chine.....	32
Tableau 3 : Exemples de classificateurs nominaux en chinois.....	35
Tableau 4 : Exemple montrant l'absence de différenciation de nombre d'un nom chinois.....	35
Tableau 5 : Exemple montrant la polysémie d'un nom sans déterminants numériques et articles .....	35
Tableau 6 : Exemples portant sur le temps et l'aspect du chinois simplifié .....	36
Tableau 7 : Exemple portant sur le changement de structure influencé par l'ordre des mots .....	36
Tableau 8 : Autre exemple portant sur le changement de structure influencé par l'ordre des mots .....	37
Tableau 9 : Les traits fondamentaux de l'écriture idéographique du chinois simplifié.....	39
Tableau 10 : Tableau des unités graphiques de l'alphabet chinois (pinyin) .....	42
Tableau 11 : Les quatre tons chinois avec la syllabe « ma » .....	43
Tableau 12 : évolution des relations sémantiques entre néologie et néologisme selon Sablayrolles (2000).....	89
Tableau 13 : Les procédés de création (Pruvost et Sablayrolles 2012:117) .....	101
Tableau 14 : Les procédés de création (Pruvost et Sablayrolles 2016:115) .....	103
Tableau 15 : Tableau raisonné et hiérarchisé des matrices (Sablayrolles 2019:127).....	105
Tableau 16 : Mots clés associés au terme 网络语言 selon le moteur de CNKI.....	119
Tableau 17 : Liste du palmarès des 25 termes les plus populaires présentée dans le rapport 《2014 中国 网络语象报告》 .....	142
Tableau 18 : Récapitulatif synthétique des nouvelles entrées de deux dictionnaires selon le club d'orthographe de Grenoble .....	153
Tableau 19 : Liste du corpus A .....	165
Tableau 20 : Liste du corpus B .....	171
Tableau 21 : Liste du Corpus C .....	174
Tableau 22 : Récapitulatif des nouvelles entrées cités par les organes de presse (Dictionnaire <i>Le Petit Robert</i> ) - Corpus Af .....	180
Tableau 23 : Récapitulatif de nouvelles entrées citées par les organes de presse (Dictionnaire <i>Le Petit Larousse</i> ) - Corpus Bf.....	180
Tableau 24 : Emprunts identifiés dans le corpus de listes chinois et leurs sources .....	185
Tableau 25 : Liste des termes .....	200
Tableau 26 : Les nouveautés langagières en commun dans les sous-corpus A et B.....	238
Tableau 27 : Nouveautés langagières communes traduites et contextualisées .....	247
Tableau 28 : Echantillon de nouveautés langagières choisi pour la collecte de tweets sur <i>Sina Weibo</i>	



.....	255
Tableau 29 : Résultats de la situation d'emploi par recherche filtrée.....	260
Tableau 30 : Echantillon de nouveautés langagières choisi pour la collecte de tweets sur <i>Twitter</i> .....	267

## LISTE DES FIGURES

Figure 1 : Aperçu du processus de la normalisation de la langue chinoise .....	27
Figure 2 : Le rapport interprété entre la langue, le dialecte et le parler.....	30
Figure 3 : Carte géographique des dialectes en Chine.....	33
Figure 4 : Disposition d'une des écritures japonaises (à gauche) et de l'écriture chinoise (à droite) ...	44
Figure 5 : Interfaces de <i>Twitter</i> montrant les différences entre le comptage des signes des caractères chinois et d'un mot français.....	45
Figure 6 : Disposition des touches du mode de saisie Wubi .....	46
Figure 7 : Mode de saisie pinyin logicielle assistée de caractères chinois simplifiés .....	47
Figure 8 : L'architecture de l'Internet.....	52
Figure 9 : Données de Google Trends .....	66
Figure 10 : Relation entre médias sociaux et RSN selon Girard et Fallery (2009) .....	69
Figure 11 : Cartographie des médias sociaux par Coutant et Stenger (2013).....	70
Figure 12 : Méthodologie de recherche documentaire dans CNKI .....	116
Figure 13 : Nombre de publications annuelles des études portant sur 网络语言 (langage d'Internet) (1998-2008) .....	117
Figure 14 : Une scène du sketch 不差钱 (bú chā qián) présenté au Gala de Nouvel An 2009 .....	137
Figure 15 : La couverture du magazine 《咬文嚼字》 (octobre 2016) .....	139
Figure 16 : Billet du compte <i>Sina Weibo</i> de magazine annonçant le résultat de la sélection 2012 .....	140
Figure 17 : Capture d'écran d'une émission d'information de la Chaîne CCTV (le 19 décembre 2015) qui rapporte l'activité 《汉语盘点》 2014.....	141
Figure 18 : Affiche du Festival XYZ (édition 2011) .....	146
Figure 19 : Affiche du Festival du Mot (édition 2018).....	148
Figure 20 : Couverture de la brochure pédagogique « dis-moi dix mots » (édition 2016-2017) .....	149
Figure 21 : Captures d'écran d'articles de presse qui présentent les nouvelles entrées des dictionnaires .....	152
Figure 22 : Proportions des matrices internes et externe dans le corpus de listes chinois.....	184
Figure 23 : Proportions des différentes matrices internes dans le corpus de listes chinois .....	193
Figure 24 : Proportions des matrices internes et externe dans le corpus de listes français .....	202
Figure 25 : Proportions des matrices internes et externe dans le corpus de listes français à partir d'une définition restrictive de l'emprunt .....	204
Figure 26 : Proportions des différentes matrices internes dans le corpus de listes français .....	205
Figure 27 : Différents affichages de tweets en captures d'écran selon Paveau (2013:7-9) .....	221
Figure 28 : Capture d'écran du message de Sarkozy présenté par Yanoshevsky (2010:4) .....	224

Figure 29 : Capture d'écran de la présentation des tweets de M. Le Pen par Longhi (2013:3) .....	224
Figure 30 : <i>attachiant</i> se trouve dans une réponse .....	228
Figure 31 : <i>attachiant</i> se trouve dans un retweet avec commentaire.....	228
Figure 32 : Réponse contenant le mot « attachiant » présentée dans la page de résultat .....	229
Figure 33 : Conversation liée à une réponse .....	229
Figure 34 : Affichage de la réponse dans l'intégralité de l'interaction.....	230
Figure 35 : Interface d'affichage des résultats de recherche par Hashtag sur <i>Sina Weibo</i> .....	233
Figure 36 : Interface d'affichage du résultat de recherche par Hashtag sur <i>Twitter</i> .....	233
Figure 37 : Publication du terme 元芳，你怎么看 datée du 9 juin 2020, 20:17.....	257
Figure 38 : Publication du terme 洪荒之力 datée du 9 juin 2020.....	258
Figure 39 : Capture d'écran du tweet n°3.5 avec hashtag en position initiale.....	274
Figure 40 : Capture d'écran du tweet n°2.3 avec hashtag en position finale.....	275
Figure 41 : Capture d'écran du tweet n°1.5 avec hashtag en position finale.....	276
Figure 42 : Capture d'écran du tweet n°1.2 avec le texte du tweet constitué par un hashtag.....	276
Figure 43 : Capture d'écran du tweet n°2.4.....	278
Figure 44 : Capture d'écran du tweet n°4.1.....	279
Figure 45 : Capture d'écran du tweet n°3.2.....	280
Figure 46 : Capture d'écran du tweet n°3.4.....	281
Figure 47 : Capture d'écran du tweet n°4.2.....	282
Figure 48 : Capture d'écran du tweet n°1.10.....	282
Figure 49 : Capture d'écran du tweet n°2.10.....	284
Figure 50 : Capture d'écran du tweet n°5.4.....	285
Figure 51 : Capture d'écran des tweets n°4.10 et n°3.10 .....	286
Figure 52 : Capture d'écran d'un tweet de 2014 qui s'interroge sur le sens des nouveautés langagières .....	288
Figure 53 : Capture d'écran d'un tweet employant « tuerie » dans son sens ancien.....	296
Figure 54 : Capture d'écran du tweet n°7.1.....	297
Figure 55 : Capture d'écran du tweet n°7.5.....	297
Figure 56 : Capture d'écran du tweet n°4.4 et page complète de la Une, image rapportée.....	299
Figure 57 : Capture d'écran du tweet n°7.4 et page complète de la Une, image rapportée.....	301
Figure 58 : Capture d'écran du tweet n°6.1.....	303
Figure 59 : Capture d'écran du tweet n°6.2.....	303
Figure 60 : Capture d'écran du tweet n°2.1.....	304
Figure 61 : Capture d'écran du tweet n°4.1.....	305
Figure 62 : éléments d'identité déclarative sur le compte d'Emmanuel Macron.....	306
Figure 63 : Capture d'écran du tweet n°6.3.....	307

Figure 64 : Capture d'écran du tweet n°6.6.....	308
Figure 65 : Capture d'écran du tweet n°6.10.....	308
Figure 66 : Capture d'écran du tweet n°3.5.....	309
Figure 67 : Capture d'écran du tweet n°2.2.....	310
Figure 68 : Capture d'écran du tweet n°2.5.....	311
Figure 69 : Capture d'écran du tweet n°1.5.....	312
Figure 70 : Capture d'écran du tweet n°2.6.....	312
Figure 71 : Capture d'écran du tweet n°2.4.....	313
Figure 72 : Capture d'écran du tweet n°2.4.....	316
Figure 73 : Capture d'écran du tweet n°3.8.....	317
Figure 74 : Capture d'écran du tweet n°6.9.....	318
Figure 75 : Capture d'écran du tweet n°4.10.....	318



# TABLE DES MATIÈRES

REMERCIEMENTS .....	3
SOMMAIRE .....	5
INTRODUCTION GÉNÉRALE.....	7
PARTIE I : CONTEXTUALISATION.....	15
CHAPITRE I : Présentation du chinois : caractéristiques historiques, linguistiques et sociolinguistiques .....	17
Introduction .....	17
1.1.Généralités sur la langue chinoise .....	18
1.2. Aperçu historique .....	21
1.2.1. Plusieurs divisions en périodes .....	21
1.2.2. Le chinois contemporain (现代汉语, xiàn dài hàn yǔ).....	25
1.3. Les dialectes chinois.....	27
1.4. Les caractéristiques du chinois contemporain.....	34
1.4.1. Caractéristiques grammaticales comparées.....	34
1.4.2. Les systèmes d'écriture .....	37
1.5. La saisie informatique du chinois.....	45
Conclusion.....	47
CHAPITRE II : Contexte numérique et nouveaux dispositifs médiatiques.....	49
Introduction .....	49
2.1. L'Internet et le Web .....	50
2.1.1. Aperçu historique .....	50
2.1.2. Qu'est- ce le Web 2.0 ?.....	53
2.2. Réseaux sociaux, RSN, médias sociaux, quelles sont les différences ?.....	55
2.2.1. Réseaux sociaux et réseaux sociaux numériques (RSN).....	56
2.2.2. Les caractéristiques des RSN .....	57
2.2.3. Evolution des RSN .....	60
2.2.4. Des notions en discussion dans la communauté des chercheurs .....	61
2.2.5. Médias sociaux et RSN .....	65
2.3. Le microblog, en tant que RSN .....	72
2.3.1. Présentation générale.....	72
2.3.2. <i>Twitter</i> et <i>Sina Weibo</i> .....	75
2.4. Rôles des usagers et des chercheurs dans le contexte numérique .....	79
2.4.1. La culture participative.....	80

2.4.2. L'observation participante .....	81
Conclusion.....	83
CHAPITRE III : Des néologismes aux nouveautés langagières en contexte numérique : définition de l'objet de recherche.....	85
Introduction .....	85
3.1. Les métadiscours savants en question.....	86
3.1.1. Les notions de <i>néologie</i> et <i>néologisme</i> dans les dictionnaires français.....	86
3.1.2. Ressources chinoises .....	90
3.1.3. Cadre conceptuel pour l'interprétation des néologismes et de leur formation.....	94
3.1.4. Les typologies des néologismes .....	98
3.2. Créations lexicales en contexte numérique .....	109
3.2.1. Caractéristiques communicationnelles générales des créations lexicales .....	109
3.2.2. Nouveaux phénomènes langagiers en contexte numérique : le point de vue des chercheurs chinois .....	112
3.2.3. Nouveaux phénomènes langagiers en contexte numérique : le point de vue des chercheurs francophones et anglophones .....	123
3.3. Nommer notre objet d'étude .....	127
Conclusion.....	130
PARTIE II : MÉTHODOLOGIE ET ANALYSES.....	131
CHAPITRE IV : Présentation des corpus et de la méthodologie de constitution des corpus	133
Introduction .....	133
4.1. Questionnements et principales options relatives au corpus .....	133
4.2. Démarches préparatoires en vue de la collecte des corpus .....	136
4.2.1. Observations et démarche de collecte pour le corpus (en) chinois .....	136
4.2.2. Observations et démarche de collecte pour le corpus (en) français .....	144
4.3. Une démarche parallèle mais différenciée pour un double corpus .....	154
Conclusion.....	156
CHAPITRE V : Constitution et analyse des corpus de liste .....	159
Introduction .....	159
5.1. Constitution du corpus de liste chinois .....	159
5.1.1. Corpus A : 《咬文嚼字》十大流行语 (les 10 mots et expressions les plus populaires de l'année).....	160
5.1.2. Corpus B : 《汉语盘点》十大网络用语 (le top des 10 termes utilisés en ligne)....	167
5.1.3. Corpus C : 2014 网络流行语排行 (les mots populaires de l'Internet de l'année 2014) .....	173

5.2. Procédures de constitution du corpus français .....	175
5.2.1. Les nouvelles entrées des dictionnaires citées dans les articles de presse .....	175
5.3. Analyse linguistique comparée des corpus de listes.....	180
5.3.1. Une grille d'analyse commune pour comparer les procédés de création.....	181
5.3.2. Le corpus de listes chinois .....	183
5.3.3. Le corpus de listes français .....	201
5.3.4. Eléments de comparaison.....	213
Conclusion.....	215
CHAPITRE VI : Constitution d'un corpus de tweets en vue d'une analyse discursive : approche théorique et méthodologique .....	217
Introduction .....	217
6.1. Eléments théoriques et méthodologiques pour le traitement des corpus RSN.....	217
6.1.1. La créativité lexicale et <i>Twitter</i> selon Reboul- Touré.....	217
6.1.2. La contextualisation technorelationnelle de Paveau .....	219
6.1.3. La méthodologie contextuelle et outillée de Longhi appliquée aux tweets politiques. ....	222
6.1.4. Options et questionnements .....	224
6.2. La constitution du corpus de tweets chinois en 4 étapes.....	237
6.2.1. Choix des nouveautés langagières pour collecter des tweets : l'étape des éléments communs .....	237
6.2.2. Une étape intermédiaire de pré-analyse contextualisée .....	238
6.2.3. L'étape de sélection d'un échantillon pour l'analyse discursive des tweets .....	254
6.2.4. Démarche de sélection des tweets .....	256
6.3. La constitution du corpus de tweets français .....	261
6.3.1 Premières observations sur le corpus de listes français et ses sources.....	261
6.3.2 Une démarche adaptée de sélection des entrées.....	263
Conclusion.....	270
CHAPITRE VII : Analyse discursive de deux corpus de tweets .....	273
Introduction .....	273
7.1. Analyse discursive d'un corpus de tweets chinois .....	273
7.1.1. Discussion sur la catégorisation des hashtags selon leur emplacement .....	274
7.1.2 Fonctions des hashtags .....	277
7.1.3. Les relations entre hashtags et nouveautés langagières .....	287
7.1.4. Observations sur la circulation des nouveautés langagières .....	293
7.2. Analyse discursive d'un corpus de tweets français .....	295
7.2.1. Caractéristiques linguistiques sémantiques .....	295



7.2.2. Un discours déterminé par les visées des producteurs .....	302
7.2.3. Caractéristiques technodiscursives sur <i>Twitter</i> .....	313
Conclusion.....	319
CONCLUSION GÉNÉRALE.....	321
BIBLIOGRAPHIE .....	327
LISTE DES TABLEAUX .....	341
LISTE DES FIGURES.....	343
TABLE DES MATIÈRES .....	347



**Titre :** Une étude discursive de la circulation des nouveautés langagières sur les réseaux sociaux numériques chinois et français : le cas de *Sina Weibo* et *Twitter*

**Mots clés :** Néologie, Néologisme, Réseaux sociaux numériques, discours numérique, analyse comparée chinois-français, linguistique d'Internet

**Résumé :**

Cette recherche s'intéresse à des productions langagières néologiques diffusées sur les dispositifs numériques de micro blogging qui connaissent souvent une circulation rapide et élargie. Elle prend pour objet des nouveautés langagières en chinois et en français sur une période de 9 années allant de 2009 à 2017 et reconnues comme telles par des sources institutionnelles et ou médiatiques. Les productions langagières sont appréhendées dans l'environnement technorelationnel des réseaux socio-numériques *Sina Weibo* (chinois) et de *Twitter* (version française). L'étude qui se veut à la fois linguistique et communicationnelle relève d'une analyse du discours numérique dans une

perspective écologique et comparée qui tient compte des spécificités des deux langues considérées. Cette perspective amène à explorer la notion de culture numérique participative et les usages des nouveaux dispositifs médiatiques.

L'analyse linguistique d'un corpus d'entrées néologiques et l'analyse discursive d'un corpus de tweets visent à observer si les nouvelles modalités d'expression et les pratiques des usagers influencent la créativité langagière et à voir comment leur circulation va de pair avec les évolutions de sens.

**Title:** A discursive study of newly invented language productions on the French and Chinese social networks: *Sina Weibo* and *Twitter*

**Keywords:** neology, neologism, social networks, digital discourse, Chinese-French comparative analysis, Internet linguistics

**Abstract:**

This research is interested in newly invented language productions and its spread on Social Networks of microblog. This kind of language creation is often propagated rapidly and extensively. It focuses on new language productions in Chinese and French from 2009 to 2017, collected from institutional and/or media sources. Those productions will be observed in the technorelational environment of the social networks *Sina Weibo* (Chinese) and *Twitter* (French version). The study, which is in both aspects of linguistics and communication, is based on an analysis of digital discourse in an ecological and comparative perspective that takes into

account the specifics of two languages. This perspective leads to explore the notion of participatory culture and the uses of new media devices. Meanwhile, the characteristics that define neology and neologism are going to be applied to discuss the nature of new language productions. Then, to figure out the impact of new communication medium and its user practices on language creativity, and the evolution of its meaning during its circulation on social networks, a comparative linguistics analysis of a corpus of neological entries and a discursive analysis of a corpus of tweets are going to be conducted.



**THÈSE DE DOCTORAT DE L' UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ  
PRÉPARÉE À L'UNIVERSITÉ DE FRANCHE-COMTÉ**

École doctorale n°592

Lettres, Communication, Langues, Arts (LECLA)

**Doctorat de sciences du langage**

Par

GAO Shan

**Une étude discursive de la circulation des nouveautés langagières  
sur les réseaux sociaux numériques chinois et français : le cas de *Sina Weibo* et *Twitter***

Volume II

Thèse présentée et soutenue à Besançon, le 17 mai 2021

Composition du Jury :

M. PRUVOST Jean,	Professeur émérite à Cergy Paris Université,	Président de Jury & Rapporteur
Mme MITROPOULOU Eleni,	Professeure à l'Université de Haute-Alsace,	Rapporteure
Mme CHEN Min,	MCF à l'Université de l'An'hui (Chine),	Examinatrice
Mme CHAUVIN-VILENO Andrée,	Professeure émérite à l'Université de Franche-Comté,	Directrice de thèse
M. MADINI Mongi,	MCF HDR émérite à l'Université de Franche-Comté,	Codirecteur de thèse



# ANNEXES

## Sommaire

ANNEXES.....	1
Annexe 1 : Tableau chronologique des dynasties chinoises .....	3
Annexe 2 : Documents relatifs au corpus de listes chinois .....	4
2a : Tableau du corpus A (《咬文嚼字》十大流行, les 10 mots et expressions les plus populaires de l'année).....	4
2a-T : Tableau du corpus A (traduit en français).....	5
2b : Tableau du corpus B (《汉语盘点》十大网络用语, le top des 10 termes utilisés en ligne).....	10
2b-T : Tableau du corpus B (traduit en français).....	11
2c : Tableau du corpus C (2014 网络流行语排行, les mots populaires de l'Internet de l'année 2014).....	16
2c-T : Tableau du corpus C (traduit en français).....	17
2d : Tableau des entrées du corpus de listes chinois classées par matrice.....	19
2f : Tableau des nouveautés langagières communes traduites et contextualisées .....	20
Annexe 3 : Documents relatifs au corpus de listes français .....	29
3a-1 : Tableau des nouvelles entrées des dictionnaires rapportées par <i>Le Figaro</i> .....	29
3a-2 : Tableau des nouvelles entrées des dictionnaires rapportées par <i>20 Minutes</i> .....	31
3a-3 : Tableau des nouvelles entrées des dictionnaires rapportées par <i>Le Parisien</i> .....	32
3a-4 : Tableau des nouvelles entrées des dictionnaires rapportées par <i>RFI</i> .....	33
3b-1 : Tableau récapitulatif de nouvelles entrées citées par les organes de presse (Dictionnaire <i>Le Petit Robert</i> ) corpus Af.....	34
3b-2 : Tableau récapitulatif de nouvelles entrées citées par les organes de presse (Dictionnaire <i>Le Petit Larousse</i> ) corpus Bf .....	34
3c : Tableau récapitulatif des entrées du corpus de listes français classées par matrice .	35
3d : Tableau préparatoire au classement par matrices des entrées du corpus de listes français .....	38
Annexe 4 : Le corpus de tweets chinois.....	39
4a : Le terme n°1 元芳, 你怎么看 (yuán fāng, nǐ zěn me kàn) .....	39
4b : Le terme n°2 女汉子 (nǚ hàn zi).....	44
4c : Le terme n°3 高大上 (gāo dà shàng).....	49
4d : Le terme n°4 主要看气质 (zhǔ yào kàn qì zhì).....	53
4e : Le terme n°5 洪荒之力 (hóng huāng zhī lì).....	58
4f : Le terme n°6 蓝瘦, 香菇 (lán shòu, xiāng gū).....	63
4g : Le terme n°7 打 call (dǎ call).....	67
Annexe 5 : Le corpus de tweets français.....	72
5a : Le terme n°1 <i>tuerie</i> .....	72
5b : Le terme n°2 <i>marche blanche</i> .....	74
5c : Le terme n°3 <i>nomophobe</i> .....	77

5d : Le terme n°4 <i>zadiste</i> .....	80
5e : Le terme n°5 <i>tendu comme un string</i> .....	85
5f : Le terme n°6 <i>vivre ensemble</i> .....	88
5g : Le terme n°7 <i>loup solitaire</i> .....	91
Annexe 6 : Métadonnées des corpus.....	95
6a : Captures d'écran des billets de blog du magazine « yǎo wén jiào zì» (2009-2017).	95
6b : Etymologie et définition des sélections selon le Rapport 2014 sur le langage d'Internet en Chine.....	112
6c : Captures d'écran de certains articles de presse portant sur les nouvelles entrées des dictionnaires français .....	113



## Annexe 1 : Tableau chronologique des dynasties chinoises<sup>1</sup>

Dynastie	Début et fin de règne
Dynastie des Xia	Environ XXIIe-XXVIe siècle av. j.-C
Dynastie des Shang	Environ XXVIe-XIe siècle av. j.-C
Zhou de l'Ouest	Environ XIe siècle-770 av. j.-C
Zhou de l'Est (Epoque des Printemps et Automne, et des Royaumes combattants)	770-221 av. j.-C
Dynastie des Qin	221-207 av. j.-C
Han de l'Ouest	206 av.j.-C-24 ap. j.-C
Han de l'Est	25-220 ap. j.-C
Les Trois Royaumes(Wei, Shu, Wu)	220-265 ap. j.-C
Jin de l'Ouest	265-316 ap. j.-C
Jin de l'Est	317-420 ap. j.-C
Dynastie du Sud et du Nord	420-589 ap. j.-C
Dynastie des Sui	581-618 ap. j.-C
Dynastie des Tang	618-907 ap. j.-C
Les Cinq Dynasties	907-960 ap. j.-C
Dynastie Song du Nord	960-1127 ap. j.-C
Dynastie Song du Sud	1127-1279 ap. j.-C
Dynastie des Yuan	1271-1368 ap. j.-C
Dynastie des Ming	1368-1644 ap. j.-C
Dynastie des Qing	1644-1911 ap. j.-C
La République de Chine	1912-1949 ap. j.-C
La République populaire de Chine	1949- ap. j.-C

<sup>1</sup> Source: <http://www.mfa.gov.cn/fra/ljzg/658/t25223.htm> (Consulté en 3 mai 2016)

## Annexe 2 : Documents relatifs au corpus de listes chinois

2a : Tableau du corpus A (《咬文嚼字》十大流行, les 10 mots et expressions les plus populaires de l'année)

2009	2010	2011	2012	2013	2014	2015	2016	2017
不差钱	给力	亲	正能量	中国梦	顶层设计	获得感	洪荒之力	不忘初心
躲猫猫	神马都是浮云	伤不起	元芳, 你怎么看	光盘	新常态	互联网+	吃瓜群众	砥砺前行
低碳	围脖	Hold住	舌尖上	倒逼	打虎拍蝇	颜值	工匠精神	共享
被就业	围观	我反正信了	躺着也中枪	逆袭	断崖式	宝宝	小目标	有温度
裸	二代	坑爹	高帅富	微XX	你懂的	创客	一言不合就XX	流量
纠结	拼爹	卖萌	中国式	大V	断舍离	脑洞大开	友谊的小船, 说翻就翻	可能XXX 假XXX
钓鱼	控	吐槽	压力山大	女汉子	失联	任性	供给侧	油腻
秒杀	帝	气场	赞	土豪	神器	剁手党	葛优躺	尬
蜗居	达人	悲催	最美	奇葩	高大上	网红	套路	怼
蚁族	穿越	忐忑	接地气	点赞	萌萌哒	主要看气质	蓝瘦, 香菇	打call

## 2a-T : Tableau du corpus A (traduit en français)

2009	2010	2011	2012	2013	2014	2015	2016	2017
不差钱 bù chà qián  ne manque pas d'argent	给力 gěi lì  c'est génial	亲 qīn  [chéri(e)] <sup>2</sup> cher client (conversation commerciale en ligne) <sup>3</sup>	正能量 zhèng néng liàng  qui donne une énergie positive	中国梦 zhōng guó mèng  rêve chinois (ambition pour le pays issu du discours politique du président Xi <sup>4</sup> )	顶层设计 dǐng céng shè jì  traiter d'un point de vue général en visant l'efficacité (disc.politique)	获得感 huò dé gǎn  satisfaction d'avoir obtenu , gagné qqch. (pour le peuple : disc.politique)	洪荒之力 hóng huāng zhī lì  Voir n° 12 * <sup>5</sup>	不忘初心 bù wàng chū xīn  ne pas oublier les objectifs initiaux dans l'action (disc.politique)
躲猫猫 duǒ māo māo  [jeu de cache cache] (connotation grave : suspicion en lien avec actualité du décès d'un suspect en prison)	神马都是浮云 shén mǎ dōu shì fú yún  rien n'a d'importance	伤不起 shāng bù qǐ  difficile à supporter	元芳，你怎么看 yuán fāng, nǐ zěnmē kàn  Voir n° 1*	光盘 guāng pán  finir toute l'assiette (initiative gouvernementale pour éviter gaspillage)	新常态 xīn cháng tài  nouvelle norme (notion économique)	互联网+ hù lián wǎng +  l'autoroute de l'information (disc. politique 1er ministre chinois Li keqiang)	吃瓜群众 chī guā qún zhòng  Voir n° 13 *	砥砺前行 dǐ lì fèn jìn  avancer malgré les difficultés (disc. politique)

<sup>2</sup> Les éléments entre crochets correspondent à une traduction littérale quand celle-ci est différente du sens global ou du nouveau sens du mot ou de l'expression.

<sup>3</sup> Les éléments entre parenthèses indiquent brièvement le contexte ou le domaine thématique s'il est nécessaire à la compréhension.

<sup>4</sup> La plupart des éléments issus du discours politique viennent des propos du président Xi Jinping ou du gouvernement, tels qu'ils sont repris par les médias.

<sup>5</sup> Les éléments suivis d'astérisque sont étudiés de manière développée dans le Tableau 2f.

<p>低碳 dī tàn</p> <p>baisse du carbone (notion écologique ; disc. politique)</p>	<p>围脖 wéi bó</p> <p>petite écharpe ; cache-cou (nom du microblog dans l'usage courant)</p>	<p>Hold住 Hold zhù</p> <p>tenir qqch. sous contrôle</p>	<p>舌尖上 shé jiān shàng</p> <p>Voir n° 2 *</p>	<p>倒逼 dào bāi</p> <p>Stimuler d'une manière inverse</p>	<p>打虎拍蝇 dǎ hǔ pāi yíng</p> <p>[frapper les tigres comme les mouches] (disc. pol. ; anticorruption concerne tout le monde, les cadres comme les agents)</p>	<p>颜值 yán zhí</p> <p>évaluation d'après une belle apparence</p>	<p>工匠精神 gōng jiàng jīng shén</p> <p>être créatif, professionnel et sérieux dans le travail [L'esprit d'artisanat] (disc. politique maîtrise des techniques)</p>	<p>共享 gòng xiǎng</p> <p>partager (notion économique; système de location ; pratique sociale)</p>
<p>被就业 bèi jiù yè</p> <p>être employé à son insu (commentaire ironique d'actualité)</p>	<p>围观 wéi guān</p> <p>regarder en tant que spectateur ou témoin passif</p>	<p>我反正信了 wǒ fǎn zhèng xìn le</p> <p>[je suis convaincu] je ne crois pas (antiphrase)</p>	<p>躺着也中枪 tǎng zhe yě zhōng qiāng</p> <p>Voir n° 3*</p>	<p>逆袭 nì xí</p> <p>Voir n° 5 *</p>	<p>断崖式 duàn yá shì</p> <p>comme une falaise qui s'écroule (disc. politique à propos de qch qui baisse soudainement, ex. : les prix)</p>	<p>宝宝 bǎo bǎo</p> <p>Voir n° 10*</p>	<p>小目标 xiǎo mù biāo</p> <p>Voir n° 14 *</p>	<p>有温度 yǒu wēn dù</p> <p>qqch. ou qqn qui procure le bien- être (disc. politique à propos de la construction d'une ville agréable, plus humaine)</p>

<p>裸 luǒ</p> <p>[être nu]</p> <p>ne rien préparer ; être délogé de l'aspect matériel</p>	<p>二代 èr dài</p> <p>[2<sup>e</sup> génération] qui ne fait pas d'effort mais bénéficie d'avantages qu'offre la famille</p>	<p>坑爹 kēng diē</p> <p>[tromper son père]</p> <p>Quelle tromperie !</p>	<p>高帅富 gāo shuài fù</p> <p>Voir n° 4 *</p>	<p>微 XX wēi XX</p> <p>micro, notamment pour la taille</p>	<p>你懂的 nǐ dǒng de</p> <p>c'est sûr que tu as compris (pour indiquer qu'il ne convient pas d'en parler en public)</p>	<p>创客 chuàng kè</p> <p>qqn de très doué pour la technique qui aime partager (trad.anglais <i>culture maker</i>)</p>	<p>一言不合就 XX yī yán bù hé jiù XX</p> <p>si on n'est pas d'accord, tout de suite on va faire XX (qqch.)</p>	<p>流量 liú liàng</p> <p>acteurs ou célébrités qui ont beaucoup de fans et d'influence (valeur attraction éventuellement commerciale)</p>
<p>纠结 jiū jié</p> <p>[plantes ou fils qui s'entrelacent] hésiter</p>	<p>拼爹 pīn diē</p> <p>[comparer les pères] profiter des avantages des pères</p>	<p>卖萌 mài méng</p> <p>montrer qu'on est mignon (cf. n° 9*)</p>	<p>中国式 zhōng guó shì</p> <p>à la chinoise (titre d'un feuilleton ex. : « Divorce à la chinoise »)</p>	<p>大 V dà V</p> <p>comptes vérifiés sur microblog qui ont une influence importante</p>	<p>断舍离 duàn shě lí</p> <p>il faut abandonner qqch., mener un style de vie simple (emprunt au jap.)</p>	<p>脑洞大开 nǎo dòng dà kāi</p> <p>[il y a un grand trou dans le cerveau] libérer l'imagination</p>	<p>友谊的小船，说 翻就翻 yǒu yì de xiǎo chuán, shuō fān jiù fān</p> <p>Voir n°15*</p>	<p>可能 XXX 假 XXX kě néng XXX jiǎ XXX</p> <p>possible que qqch. XXX soit une contrefaçon (plaisanterie sur la mauvaise qualité ou qqch. qui n'est pas conforme à l'attente)</p>

<p>钓鱼 diào yú</p> <p>[pêcher] Tendre un piège à qqn</p>	<p>控 kòng</p> <p>aimer beaucoup qqch. (emprunt phonétique au jap.)</p>	<p>吐槽 tù cáo</p> <p>[vomir ; ustensile pour gaver] se plaindre ; commenter (emprunt au jap.)</p>	<p>压力山大 yā lì shān dà</p> <p>[Alexandra : paronyme]  stresser beaucoup</p>	<p>女汉子 nǚ hàn zi</p> <p>Voir n°6*</p>	<p>失联 shī lián</p> <p>perdre contact (dialecte)</p>	<p>任性 rèn xìng</p> <p>[capricieux] faire comme on veut, être désinvolte</p>	<p>供给侧 gōng jǐ cè</p> <p>[fournir un côté] entité qui offre (notion économique)</p>	<p>油腻 yóu nì</p> <p>[gras] Voir n°19 *</p>
<p>秒杀 miǎo shā</p> <p>[seconde tuer] offre de dernière minute</p>	<p>帝 dì</p> <p>[L'empereur] qqn talentueux</p>	<p>气场 qì chǎng</p> <p>[champ d'énergie autour des personnes] charme ou caractère de qqn (notion de physique)</p>	<p>赞 zàn</p> <p>[génial]  j'aime (<i>like</i> sur microblog)</p>	<p>土豪 tǔ háo</p> <p>Voir n°7 *</p>	<p>神器 shén qì</p> <p>[objet sacré] outils performants efficaces qui facilitent une tâche</p>	<p>剁手党 duò shǒu dǎng</p> <p>[gang qui coupe les mains] qqn qui achète trop de choses, qui aime acheter</p>	<p>葛优躺 gě yōu tāng</p> <p>Voir n° 16 *</p>	<p>尬 gà</p> <p>gênant</p>

蜗居 wō jū  [maison d'escargot] Se loger dans un espace exigu pour des raisons financières, souvent dans des grandes villes (conditions de vie des jeunes)	达人 dá rén  [qqn de rationnel] qqn qui a du talent	悲催 bēi cuī  triste	最美 zuì měi  [le plus beau]  lié à actualité pour décrire le caractère positif héroïque de qqn	奇葩 qí pā  [jolie fleur] qqn de bizarre (ironie)	高大上 gāo dà shàng  Voir n°8 *	网红 wǎng hóng  [Internet rouge]  influenceur (rouge associé à célébration)	套路 tào lù  Voir n°17 *	怼 duǐ  Voir n°20 *
蚁族 yǐ zú  [fourmis] jeunes diplômés en situation précaire vivant regroupés comme dans une fourmilière en banlieue	穿越 chuān yuè  [traverser un espace physique] voyager dans le temps (science-fiction) être perdu (sentiment)	忐忑 tǎn tè  s'inquiéter de qqch.	接地气 jiē dì qì  [toucher la terre]  connaître qqch. c'est le connaître dans le réel au quotidien	点赞 diǎn zàn  [j'aime, je like]  dire des choses positives sur qch	萌萌哒 méng méng da  Voir n°9 *	主要看气质 zhǔ yào kàn qì zhi  Voir n°11 *	蓝瘦，香菇 lán shòu, xiāng gū  Voir n°18 *	打 call dǎ call  Voir n°21*

2b : Tableau du corpus B (《汉语盘点》十大网络用语, le top des 10 termes utilisés en ligne)

2012	2013	2014	2015	2016	2017
中国好声音	中国大妈	我也是醉了	重要的事情说三遍	洪荒之力	打 call
元芳, 你怎么看	高端大气上档次	有钱就是任性	世界那么大, 我想去看看	友谊的小船	尬聊
高富帅, 白富美	爸爸去哪儿	蛮拼的	你们城里人真会玩	定个小目标	你的良心不会痛吗
你幸福吗	小伙伴们都惊呆了	挖掘机技术哪家强	为国护盘	吃瓜群众	惊不惊喜, 意不意外
江南 style	待我长发及腰	保证不打死你	明明可以靠脸吃饭, 却偏偏要靠才华	葛优躺	皮皮虾, 我们走
躺着也中枪	喜大普奔	萌萌哒	我想静静	辣眼睛	扎心了, 老铁
屌丝, 逆袭	女汉子	时间都去哪了	吓死宝宝了	全是套路	还有这种操作
舌尖上的中国	土豪金	我读书少你别骗我	内心几乎是崩溃的	蓝瘦香菇	怼
最炫民族风	摊上大事了	画面太美我不敢看	我妈是我妈	老司机	你有 freestyle 吗
给跪了	涨姿势	且行且珍惜	主要看气质	厉害了我的哥	油腻



**2b-T : Tableau du corpus B (traduit en français)**

2012	2013	2014	2015	2016	2017
<p>中国好声音 zhōng guó hǎo shēng yīn</p> <p>[la belle voix chinoise] nouveau mode de sélection avec jury (programme <i>The Voice</i> en Chine)</p>	<p>中国大妈 zhōng guó dà mā</p> <p>[les dames chinoises seniors] groupe de dames chinoises au pouvoir d'achat élevé qui pratiquent par ex. l'achat de l'or (sens social et économique)</p>	<p>我也是醉了 wǒ yě shì zuì le</p> <p>[je suis ivre] sentiment de ne rien y pouvoir</p>	<p>重要的事情说三遍 zhòng yào de shì qíng shuō sān biàn</p> <p>il faut dire trois fois les choses importantes</p>	<p>洪荒之力 hóng huāng zhī lì</p> <p>Voir n°12 *</p>	<p>打 call dǎ call</p> <p>Voir n°21*</p>
<p>元芳，你怎么看 yuán fāng, nǐ zěn me kàn</p> <p>Voir n°1*</p>	<p>高端大气上档次 gāo duān dà qì shàng dàng cì</p> <p>Voir n°8 *</p>	<p>有钱就是任性 yǒu qián jiù shì rèn xìng</p> <p>avec de l'argent on peut faire n'importe quoi</p>	<p>世界那么大，我想去看看 shì jiè nà me dà, wǒ xiǎng qù kàn kàn</p> <p>[le monde est grand et je voudrais le visiter] (à partir du cas d'une jf qui démissionne de son travail pour voyager, renvoie à attitude de liberté et de bien-être)</p>	<p>友谊的小船 yǒu yì de xiǎo chuán</p> <p>Voir n°15 *</p>	<p>尬聊 gà liáo</p> <p>bavarder de façon futile</p>

<p>高富帅，白富美 gāo fù shuài, bái fù měi</p> <p>Voir n°4 *</p>	<p>爸爸去哪儿 bàba qù nǎ'er</p> <p>[tu vas où Papa ?] on va où ? (titre émission télé-réalité)</p>	<p>蛮拼的 mán pīn de</p> <p>travailler dur</p>	<p>你们城里人真会玩 nǐ men chéng lǐ rén zhēn huì wán</p> <p>[Vous les citadins savez bien vous amuser] renvoie à des pratiques surprenantes insolites (ironie)</p>	<p>定个小目标 dìng gè xiǎo mù biāo</p> <p>Voir n°14 * [se fixer un petit objectif] gagner une grosse somme d'argent (ironie)</p>	<p>你的良心不会痛吗 nǐ de liáng xīn bù huì tòng ma</p> <p>[tu te sens mal à ton bon cœur ?] tu as mauvaise conscience ? (ironie)</p>
<p>你幸福吗 nǐ xìng fú ma</p> <p>[est-ce que tu es heureux ?] (ironique)</p>	<p>小伙伴们都惊呆了 xiǎo huǒ bàn men dōu jīng dāi le</p> <p>[mes copains et moi nous sommes tous étonnés] c'est étonnant</p>	<p>挖掘机技术哪家强 wā jué jī jì shù nǎ jiā qiáng</p> <p>[quelle école donne la meilleure formation pour une belle mécanique hydraulique ?] (publicité pour recruter ou indiquer qualité)</p>	<p>为国护盘 wèi guó hù pán</p> <p>[protéger la Bourse pour le pays] aider le pays à garder sa stabilité économique</p>	<p>吃瓜群众 chī guā qún zhòng</p> <p>Voir n°13 *</p>	<p>惊不惊喜，意不意外 jīng bù jīng xǐ, yì bù yì wài</p> <p>c'est une surprise (fin de message)</p>

<p>江南 style jiāng nán style</p> <p>Gangnam style (titre chanson coréenne)</p>	<p>待我长发及腰 dài wǒ zhǎng fā jí yāo</p> <p>[quand j'aurai les cheveux longs jusqu'à la ceinture...] j'attends un moment pour faire qqch.</p>	<p>保证不打死你 bǎo zhèng bù dǎ sǐ nǐ</p> <p>[je t'assure je ne vais pas te battre à mort] viens pour que je te tape (plaisanterie)</p>	<p>明明可以靠脸吃饭,却偏 偏要靠才华 míng míng kě yǐ kào liǎn chī fàn, què piān piān yào kào cái huá</p> <p>[qqn qui peut gagner de l'argent simplement avec sa beauté mais a choisi de travailler avec ses talents] qqn qui est beau mais a aussi des talents</p>	<p>葛优躺 gě yōu tāng</p> <p>Voir n°16 *</p>	<p>皮皮虾, 我们走 pí pí xiā, wǒ men zǒu</p> <p>[les squillidae crevettes-mantes, on y va] au revoir (image crevette signale l'humour)</p>
<p>躺着也中枪 tǎng zhe yě zhōng qiāng</p> <p>Voir n°3 *</p>	<p>喜大普奔 xǐ dà pǔ bēn</p> <p>c'est une très bonne nouvelle et je vais courir pour que tout le monde le sache (structure qui imite une locution en 4 caractères ; sens très compact)</p>	<p>萌萌哒 méng méng da</p> <p>Voir n°9*</p>	<p>我想静静 wǒ xiǎng jìng jìng</p> <p>Laisse-moi tranquille</p>	<p>辣眼睛 là yǎn jīng</p> <p>[Ça pique les yeux] se dit de choses peu intéressantes ou horribles</p>	<p>扎心了, 老铁 zhā xīn le, lǎo tiě</p> <p>[mon frère tu as blessé mon cœur] je suis triste, tu me fais mal (humour pour refuser une conversation)</p>

屌丝，逆袭 <i>diǎo sī, nì xí</i>  Voir n°5 *	女汉子 <i>nǚ hàn zi</i>  Voir n°6 *	时间都去哪了 <i>shí jiān dōu qù nǎ le</i>  [Où va le temps] (titre chanson) gaspiller le temps sans en avoir conscience	吓死宝宝了 <i>xià sǐ bǎo bǎo le</i>  Voir n°10 *	全是套路 <i>quán shì tàolù</i>  Voir n°17 *	还有这种操作 <i>hái yǒu zhè zhǒng cāo zuò</i>  [il y a une opération pareille ?] c'est surprenant
舌尖上的中国 <i>shé jiān shàng de zhōng guó</i>  Voir n°2*	土豪金 <i>tǔ háo jīn</i>  Voir n°7 *	我读书少你别骗我 <i>wǒ dú shū shǎo nǐ bié piàn wǒ</i>  [je ne lis pas beaucoup et tu ne me trompes pas] j'ai bcp de connaissances et tu n'arrives pas à me tromper (antiphrase)	内心几乎是崩溃的 <i>nèi xīn jī hū shì bēng kuì de</i>  [au fond du cœur c'est presque écrasant] je suis dépassé	蓝瘦，香菇 <i>lán shòu xiāng gū</i>  Voir n°18 *	怼 <i>duǐ</i>  Voir n°20 *
最炫民族风 <i>zuì xuàn mín zú fēng</i>  [le meilleur style folklorique]  (titre chanson)	摊上大事了 <i>tān shàng dà shì le</i>  rencontrer un terrible événement, être frappé par le malheur	画面太美我不敢看 <i>huà miàn tài měi wǒ bù gǎn kàn</i>  [l'image est très belle et je n'ose pas regarder]	我妈是我妈 <i>wǒ mā shì wǒ mā</i>  [ma mère c'est ma mère] (ironique surtout à l'égard des administrations qui	老司机 <i>lǎo sī jī</i>  [vieux chauffeur]	你有 freestyle 吗 <i>nǐ yǒu freestyle ma</i>  est-ce que tu as le free style ?

		je ne peux pas supporter de regarder ça (antiphrase pour qqch. de choquant)	demandent de prouver cela)	qqn qui a de l'expérience dans un domaine	
给跪了 gěi guì le  [s'agenouiller] admirer qqn	涨姿势 zhǎng zīshì  j'ai appris qqch. (paronyme)	且行且珍惜 qiě xíng qiě zhēn xī  [on marche tout en chérissant qqch.] il faut être compréhensif	主要看气质 zhǔ yào kàn qì zhì  Voir n°11 *	厉害了我的哥 lì hài le wǒ de gē  bravo mon ami,e	油腻 yóu nì  Voir n°19*

2c : Tableau du corpus C (2014 网络流行语排行, les mots populaires de l'Internet de l'année 2014)

Sélection

Classement

Résultats de recherche dans système de journalisation (utilisation dans les titres des articles)

表 2014 网络流行语排行

排名	网络热词	中文报刊检索结果 (标题中使用)	微博提
1	萌萌哒	1410	214,720,184
2	逗比	276	150,043,492
3	我也是醉了	298	79,679,804
4	且行且珍惜	878	73,477,308
5	涨姿势	222	70,124,496
6	心塞	132	50,601,480
7	暖男	1455	50,564,364
8	买买买	247	41,685,392
9	也是蛮拼的	643	36,984,032
10	逼格	65	30,109,324
11	画面太美我不敢看	133	29,820,644
12	小鲜肉	808	20,879,812
13	那么问题来了	309	16,912,524
14	no zuo no die	99	16,792,928
15	脑洞大开	162	15,572,224
16	有钱就是任性	855	14,904,136
17	时间都去哪了	365	10,701,780
18	保证不打死你	7	8,124,280
19	我读书少你别骗我	25	7,282,984
20	习大大	325	6,161,256
21	上天台	58	4,379,688
22	什么仇什么怨	78	4,272,464
23	你家里人知道吗?	39	3,175,480
24	挖掘机技术哪家强	38	2,033,132
25	APEC 蓝	730	1,831,056

Nombre de citation sur le SINA WEIBO

(数据来源: 新浪微博、人民网舆情监测室中文报刊系统)

数据截取日期: 2014 年 1 月 1 日-12 月 31 日)

## 2c-T : Tableau du corpus C (traduit en français)

1	萌萌哒 <b>méng méng da</b>	voir n°9 *
2	逗比 <b>dòu bǐ</b>	quelqu'un de drôle
3	我也是醉了 <b>wǒ yěs hì zuì le</b>	[je suis ivre] sentiment de ne rien y pouvoir
4	且行且珍惜 <b>qiě xíng qiě zhēn xī</b>	[on marche tout en chérissant quelque chose] il faut être compréhensif
5	涨姿势 <b>zhǎng zī shì</b>	j'ai appris qqch. (paronyme) j'ai pris connaissance
6	心塞 <b>xīn sāi</b>	[le cœur est bloqué] je ne me sens pas bien
7	暖男 <b>nuǎn nán</b>	[chaud homme] homme très attentionné
8	买买买 <b>mǎi mǎi mǎi</b>	acheter
9	也是蛮拼的 <b>yě shì mán pīn de</b>	travailler dur
10	逼格 <b>bī gé</b>	avoir de la classe
11	画面太美我不敢看 <b>huà miàn tài měi wǒ bù gǎn kàn</b>	[l'image est très belle et je n'ose pas regarder] je ne peux pas supporter de regarder ça (antiphrase pour qqch. de choquant)
12	小鲜肉 <b>xiǎo xiān ròu</b>	[petite viande fraîche] jeune homme
13	那么问题来了 <b>nà me wèn tí lái le</b>	voici que viennent les questions ?
14	No zuo no die	il ne faut pas jouer avec le feu
15	脑洞大开 <b>nǎo dòng dà kāi</b>	[il y a un grand trou dans le cerveau] libérer l'imagination
16	有钱就是任性 <b>yǒu qián jiù shì rèn xìng</b>	avec de l'argent on peut faire n'importe quoi
17	时间去哪儿了 <b>shí jiān qù nǎ'èr le</b>	[Où va le temps ?] gaspiller le temps sans en avoir conscience (titre chanson)
18	保证不打死你 <b>bǎo zhèng bù dǎ sǐ nǐ</b>	[je t'assure je ne vais pas te battre à mort] viens pour que je te tape (plaisanterie)
19	我读书少你别骗我 <b>wǒ dú shū shǎo nǐ bié piàn wǒ</b>	[je ne lis pas beaucoup et tu ne me trompes pas] j'ai bcp de connaissances et tu n'arrives pas à me tromper (antiphrase)
20	习大大 <b>xí dà dà</b>	[Xi grand grand] Oncle Xi (surnom du président Xi, sens positif)
21	上天台 <b>shàng tiān tái</b>	[aller sur le toit] se suicider (plaisanterie pour se moquer quand on subit un échec)
22	什么仇什么怨 <b>shén me chóu shén me yuàn</b>	[quelle haine, quelle hostilité] se demander pourquoi
23	你家里人知道吗?	[les membres de ta famille sont au courant?]

	nǐ jiā lǐ rén zhīdào ma?	ce n'est pas bien (ironie) ou c'est super (sens positif)
24	挖掘技术哪里强 wā jué jì shù nǎ lǐ qiáng	[quelle école donne la meilleure formation pour une belle mécanique hydraulique?] (publicité pour recruter ou indiquer qualité)
25	APEC 蓝 APEC lán	[le bleu APEC] la bonne qualité de l'air (référence au ciel bleu rare à Pékin lors de l'APEC Chine 2014 en raison de la campagne de réduction des émissions dirigée par le gouvernement chinois)



## 2d : Tableau des entrées du corpus de listes chinois classées par matrice

Tableau récapitulatif

Les dialectes	MS	SS		PH	
不差钱	Hold住	躲猫猫	亲	蜗居	剁手党
坑爹	大V	裸	最炫民族风	蚁族	网红
失联	互联网+	纠结	爸爸去哪儿	给力	主要看气质
怼	你有freestyle吗	钓鱼	挖掘机技术哪家强	神马都是浮云	吃瓜群众
<b>ME</b>	APEC蓝	围观	时间都去哪了	拼爹	不忘初心
控	No zuo no die	二代	那么问题来了	伤不起	砥砺前行
达人	涨姿势	帝	躺着也中枪	我反正信了	共享
卖萌	围脖	穿越	我读书少你别骗我	高帅富	供给侧
吐槽	蓝瘦，香菇	气场	任性	中国式	葛优躺
逆袭	一言不合就XX	悲催	洪荒之力	中国梦	压力山大
打call	可能XXX假XXX	忐忑	工匠精神	女汉子	萌萌哒
断舍离	微XX	赞	小目标	点赞	习大大
颜值	友谊的小船	最美	套路	打虎拍蝇	秒杀
低碳	舌尖上	接地气	有温度	断崖式	正能量
新常态	待我长发及腰	光盘	流量	你懂的	创客
江南style	元芳，你怎么看	倒逼	油腻	高富帅，白富美	重要的事情说三遍
暖男	中国好声音	土豪	尬	给跪了	脑洞大开
	被就业	奇葩	中国大妈	小伙伴们都惊呆了	厉害了我的哥
		顶层设计	辣眼睛	摊上大事了	买买买
		神器	老司机	我也是醉了	逼格
		获得感	心塞	蛮拼的	小鲜肉
		宝宝		保证不打死你	我想静静
				世界那么大，我想去看看	内心几乎是崩溃的
				高大上	我妈是我妈
				你们城里人真会玩	你的良心不会痛吗
				为国护盘	惊不惊喜，意不意外
				明明可以靠脸吃饭，却偏偏要靠才华	皮皮虾，我们走
				还有这种操作	扎心了，老铁
				上天台	什么仇什么怨
				你家里人知道吗？	喜大普奔
				且行且珍惜	你幸福吗
				逗比	

## 2f : Tableau des nouveautés langagières communes traduites et contextualisées

N°	Nouveautés langagières (caractères et pinyin) + [procédé création]	Traduction littérale française selon unité de sens + traduction littérale globale <sup>6</sup>	Sens contextualisé
1	元芳, 你怎么看 yuán fāng, nǐ zěn me kàn [MS]	元芳/你/怎么/看 yuán fāng / nǐ / zěn me / kàn Yuan Fang [nom d'un personnage] / tu / comment / voir <i>Yuan Fang, qu'en penses-tu?</i>	Cette phrase est issue d'un feuilleton télévisé policier de 2004. Elle est souvent utilisée pendant la conversation entre le procureur et son assistant Yuan Fang pour discuter les détails d'un procès. La réponse correspondante est : « Il y a quelque chose de louche ». Elle est largement employée par les internautes en 2012 au sujet du décès mystérieux d'une fille qui est considéré finalement par la police comme un suicide. Ensuite, cette phrase devient populaire pour appeler à commenter les actualités. Ainsi, des imitations de cette structure surgissent en changeant des éléments, comme le prénom. Tous ces usages ont favorisé la formation de «la structure Yuan Fang » et son nouveau sens ironique.
2	舌尖上 (corpus A) shé jiān shàng [MS]  舌尖上的中 国 (corpus B) shé jiān shàng de zhōng guó [MS]	舌尖/上 (corpus A) shé jiān / shàng le bout de la langue / sur <i>sur le bout de la langue</i>  舌尖/上/的/中国 (corpus B) shé jiān / shàng / de / zhōng guó le bout de la langue / sur / préposition / Chine La Chine sur le bout de la langue	《舌尖上的中国》 est le titre d'un documentaire chinois destiné à présenter les spécialités culinaires chinoises. Il signifie « Goûter la cuisine chinoise ». La version française du titre qui circule en ligne est « La Chine au bout de la langue ». Ce documentaire a connu un très grand succès. Ensuite, cette expression est devenue populaire pour exprimer des choses liées à la culture culinaire. On constate aussi des variations comme « 舌尖上的快乐 » (shé jiān shàng de kuài lè), le plaisir de goûter les plats ; « 舌尖上的爱情 » (shé jiān shàng de ài qíng), relation amoureuse nouée par la gastronomie. « 舌尖上 » est sa forme raccourcie.

<sup>6</sup> Les barres obliques sont utilisées comme marqueur de séparation.

3	躺着也中枪 tǎng zhe yě zhòng qiāng [SS]	躺着/也/中枪 tǎng zhe / yě / zhòng qiāng s'allonger / aussi / se prendre une balle <i>se prendre une balle malgré la position allongée.</i>	Généralement, la position allongée est une position de protection face aux menaces de coups de feu, comme cela, la probabilité d'être abattu est faible. Initialement employée dans le dialogue d'un film comique, cette expression signifie: « Être mêlé à une affaire sans aucune raison ». En utilisant cette expression, on manifeste son étonnement en créant un effet humoristique. Les formes raccourcies comme 躺中、躺枪 expriment le même sens. Une nouvelle expression figée et son nouveau sens se sont formés .
4	高帅富 (corpus A) gāo shuài fù [PH] 高富帅, 白富美 (corpus B) gāo fù shuài, bái fù měi [PH]	高/帅/富 (corpus A) gāo / shuài / fù grand / beau / riche <i>garçon ou homme en possession d'avantages d'apparence physique et financière.</i> 高/富/帅, /白/富/美 (corpus B) gāo / fù / shuài, / bái / fù / měi grand / beau / riche / blanc / riche / belle <i>garçon ou homme ou fille ou femme en possession d'avantages d'apparence physique et financière</i>	« 高帅富 » est une nouvelle expression figée, combinant trois adjectifs positifs qui sont simples et directs, destinés à décrire les conditions majeures à remplir par un époux parfait selon des valeurs traditionnelles. La version féminine de cette expression est « 白富美 ». Elle est originalement employée par les usagers des forums pour décrire les caractères des héros dans les séries pour les ados ou dans les mangas, par exemple, un prince ou une princesse. L'utilisation de 3 adjectifs successifs créent un effet de redoublement. Selon le contexte, ces termes peuvent être péjoratifs ou positifs.

5	逆袭 (corpus A) <i>nì xí</i> [Emprunt]  屌丝, 逆袭 (corpus B) <i>diǎo sī, nì xí</i> [Emprunt]	逆袭 (corpus A) <i>nì / xí</i> inverse / attaquer <i>contre-attaquer</i> 屌丝/, 逆袭 (corpus B) <i>diǎo sī / nì / xí</i> perdant / inverse / attaquer <i>la réussite d'un perdant</i>	<p>Cette expression d'origine japonaise signifie contre-attaquer. Après avoir fréquemment été utilisée dans des films et productions ACGN<sup>7</sup> japonais, elle est ensuite introduite en Chine dans le lexique du MMOG<sup>8</sup> pour désigner une stratégie d'attaque.</p> <p>Aujourd'hui, on constate un changement de sens et un élargissement d'application dans la société chinoise. Elle signifie actuellement qu'on réussit à faire quelque chose malgré des conditions difficiles ou défavorables, y compris la situation familiale, financière, le statut social.</p> <p>Lors de son utilisation, elle est souvent combinée avec une autre nouveauté langagière «屌丝» qui est l'antonyme de «高帅富», et fait référence à un jeune homme d'apparence médiocre et de statut social faible. Né dans une famille modeste, il n'a ni voiture, ni maison, ni relations. C'est plutôt une autodérision des jeunes chinois face à la pression sociale. Cette expression a aussi été appliquée dans les productions culturelles chinoises : les romans, les films, etc.</p>
6	女汉子 <i>nǚ hàn zi</i> [PH]	女汉子 <i>nǚ / hàn zi</i> féminin / mec <i>femme masculine</i>	<p>La première utilisation de ce terme est celle d'une mannequin et animatrice LI Ai sur <i>Sina Weibo</i> pour décrire les filles ou femmes indépendantes, capables de faire quelque chose sans l'aide des hommes, particulièrement des tâches traditionnellement attribuées aux hommes, comme le bricolage.</p> <p>Ce caractère féminin est différent des stéréotypes sur les femmes, tant sur l'apparence que sur la personnalité. Il met en valeur l'image des femmes contemporaines.</p>
7	土豪 (corpus A) <i>tǔ háo</i> [SS]  土豪金 (corpus B) <i>tǔ háo jīn</i> [SS]	土豪 (corpus A) <i>tǔ / háo</i> rustique / généreux <i>les riches</i> 土豪/金 (corpus B) <i>tǔ / háo / jīn</i> rustique / généreux / couleur d'or <i>la couleur d'or qui représente les riches</i>	<p>Ce terme existe depuis longtemps dans la langue chinoise ; originellement, il désigne les riches des campagnes et particulièrement les personnes qui exercent une autorité tyrannique et excessive sur leur entourage dans les années 50 du siècle dernier.</p> <p>Son nouveau sens est né durant la conversation entre les joueurs des jeux d'Internet. On critique les autres joueurs qui gagnent facilement par l'achat d'« équipements » très chers mais non par des stratégies. Ce genre d'action détruit les loisirs et l'égalité du jeu. Aujourd'hui, il signifie : les nouveaux riches qui dépensent de l'argent sans réfléchir et juste exhiber leur fortune. Les productions achetées sont peu importantes, par ex. : acheter un café pour 1000 euros. Ces personnes sont riches mais pas nobles.</p> <p>En adoptant ce terme, on peut exprimer la jalousie, le mépris ou l'admiration.</p>

<sup>7</sup> Voir note n°175.

<sup>8</sup> Sigle de l'anglais *Massively Multiplayer Online Game* qui signifie jeu en ligne massivement multijoueur.

			<p>« 土豪金 » est une dérivation du terme « 土豪 ». « 金 » est la couleur d'or, un symbole de fortune en Chine. Cette expression désigne généralement tous les objets de couleur d'or, avec un teint exagéré. Les personnes qui utilisent ce genre de produit sont considérés d'avoir aucun goût d'esthétique.</p> <p>On observe aussi d'autre dérivation comme : « 土豪, 我们做朋友吧 » : nouveaux riches, je veux être ton ami/e.</p>
8	<p>高大上 (corpus A)</p> <p><i>gāo dà shàng</i></p> <p>[PH]</p> <p>高端大气上档次 (corpus B)</p> <p><i>gāo duān dà qì shàng dàng cì</i></p> <p><i>dàng cì</i></p> <p>[PH]</p>	<p>高/大/上 (corpus A)</p> <p><i>gāo / dà / shàng</i></p> <p>supérieur / bien / meilleur</p> <p>haut de gamme, grandiose, classe</p> <p>高端/大气/上档次(corpus B)</p> <p><i>gāo duān / dà qì / shàng dàng cì</i></p> <p>haut de gamme / grandiose / classe</p> <p><i>haut de gamme, grandiose, classe</i></p>	<p>« 高大上 » est la forme réduite de « 高端大气上档次 ». Il est destiné à décrire quelque chose de luxueux, de haut de gamme. Selon le contexte d'utilisation, on l'emploie aussi ironiquement. Il a été utilisé pour la première fois dans une série comique chinoise 《武林外传》(wǔ lín wài zhuàn). Depuis sa propagation en 2013, on constate son utilisation dans les forums, les programmes de divertissement, les médias, la publicité, etc.</p>
9	<p>萌萌哒</p> <p><i>méng méng da</i></p> <p>[PH]</p>	<p>萌萌 / 哒</p> <p><i>méng méng / da</i></p> <p>mignon / une particule modale</p> <p><i>très aimable</i></p>	<p>Cette expression est influencée par la culture japonaise « Moe » (萌え), « <i>faisant référence à des sentiments de forte affection orientée principalement vers des personnages fictifs, issus des domaines de l'anime, du manga et du jeu vidéo</i> »<sup>9</sup>. Ces personnages ont des points communs comme : grands yeux, oreilles d'animal comme celle du chat ou du renard. Ils s'expriment d'une façon mignonne. 萌 signifie à l'origine : germer. « 萌萌哒 » vise à décrire quelque chose ou quelqu'un qui a l'air mignon. Il est difficile de retrouver sa première apparition en Chine mais il est largement utilisé en ligne sur divers réseaux sociaux ou autres plateformes de communication. On suppose que sa structure est le résultat de l'imitation d'une autre nouveauté langagière « 么么哒 » (<i>mo mo da</i>, qui signifie bisous). L'emploi de deux mots identiques renforce le sentiment à exprimer. En effet « 哒 » est une onomatopée, fréquente dans la culture Moe. Ce genre d'expression qui essaie de se rapprocher de la langue des enfants, a pour but de créer une atmosphère de communication plus détendue et agréable.</p>

<sup>9</sup> Wikipédia : [https://fr.wikipedia.org/wiki/Moe\\_\(Japon\)](https://fr.wikipedia.org/wiki/Moe_(Japon)), consulté le 9 avril 2018.

10	宝宝 (corpus A) bǎo bǎo [SS]  吓死宝宝了 (corpus B) xià sǐ bǎo bǎo le [SS]	宝宝 (corpus A) bǎo bǎo bébé <i>un bébé innocent</i>  吓死/宝宝/了 (corpus B) xià sǐ / bǎo bǎo / le mourir de peur / bébé / placé comme particule à la fin pour indiquer l'achèvement de l'action ou changement <i>cela a effrayé le bébé</i>	Le terme «宝宝» est dérivé de l'expression «吓死宝宝了». Selon la source encyclopédique <i>Baidu Baike</i> , pour sa création, il existe 8 possibilités qui sont liées par exemple aux productions audiovisuelles, aux programmes télévisés, à des personnes telles que les célébrités, les influenceurs, etc. Bébé est le sens original de «宝宝». Actuellement, il est souvent employé par l'énonciateur pour s'appeler lui-même, c'est-à-dire « je ». C'est une manière mignonne de s'exprimer. Cela est étroitement lié à la culture Moe. Au lieu de dire, « ça me fait peur », dire « cela a effrayé le bébé » peut rapprocher les interlocuteurs.
11	主要看气质 zhǔ yào kàn qì zhì [PH]	主要/看/气质 zhǔ yào / kàn / qì zhì principalement / voir / tempérament <i>c'est le tempérament qui compte.</i>	Cette expression devient populaire grâce à une chanteuse qui a publié une photo sur Sina Weibo dans le but de promouvoir la vente de son album. Cette photo la montre mangeant un hamburger or certains usagers pensent qu'il n'est pas convenable de présenter ce genre de nourriture fast food dans une production culturelle et que le design de son image de communication manque de goût. Sa réponse (« On n'évalue pas quelqu'un par son apparence ou ses habits mais par son caractère et ses connaissances») a été beaucoup likée et partagée par les internautes. Sa circulation inter-plateforme est intéressante : un jeu social sur Wechat <sup>10</sup> a été lancé alors que cette expression se propageait. Les usagers du jeu sont invités à partager dans la rubrique « moments » la photo dont ils pensent qu'elle s'accorde avec cette expression.
12	洪荒之力 hóng huāng zhī lì [SS]	洪荒/之/力 hóng huāng / zhī / lì Temps immémorial / reliant le qualificatif et le mot qualifié / pouvoir <i>le pouvoir du temps immémorial</i>	Selon la mythologie chinoise, au début des temps, « 洪荒之力 » est le pouvoir de déluge qui peut détruire le monde. L'expression a été connue récemment par les lecteurs d'un roman fantastique chinois et est devenue populaire suite à la mise en scène de ce roman en 2015, dans le feuilleton télévisé « huā qiān gǔ ». Lors les jeux Olympiques de Rio en 2016, son utilisation par la championne de natation Fu yuanhui lors d'une interview avec un journaliste après avoir battu un record asiatique, a changé le sens et créé un phénomène de partage sur les réseaux sociaux, surtout sur <i>Sina Weibo</i> , sous des formes différentes (images, émoticônes, etc.). Elle signifie actuellement : faire tout ce qu'on peut pour réaliser quelque chose.

<sup>10</sup> Une application de messagerie textuelle et vocale instantanée en Chine.

13	<p>吃瓜群众 chī guā qún zhòng [PH]</p>	<p>吃瓜/群众 chī guā qún zhòng manger / pastèque / public <i>le public ou les spectateurs qui mangent de la pastèque</i></p>	<p>Dans l'expression « 吃瓜群众 », « 瓜 » il s'agit à l'origine des graines de tournesol « 瓜子 » (guā zi) et non de la pastèque. Dans la vie quotidienne, quand les gens se réunissent pour bavarder, les graines de tournesol constituent l'une des nourritures préférées. Cette expression est créée pour décrire certains internautes qui réagissent à un billet en ligne mais n'ont pas pour but d'exprimer leurs opinions ou d'aider à résoudre les questions posées. Ils parlent d'autres sujets qui n'ont rien à voir avec la discussion ou ils mentionnent dès le début qu'ils sont des spectateurs et non des acteurs.</p> <p>On suppose que si « 瓜子 » devient « 瓜 », c'est pour économiser du temps pour taper les caractères et réaliser une communication rapide et de ce fait les graines de tournesol deviennent pastèque. Suite au succès de cette expression, son utilisation est élargie, elle désigne aujourd'hui, les spectateurs curieux (et non seulement les internautes) du déroulement de quelque chose, soit parce qu'ils ne savent pas ce qui s'est passé, soit parce qu'ils accordent leur attention à quelque chose sans intervenir.</p>
14	<p>小目标 (corpus A) xiǎo mù biāo [SS]</p> <p>定个小目标 (corpus B) dìng gè xiǎo mù biāo [SS]</p>	<p>小/目标 (corpus A) xiǎo / mù biāo petit / objectif <i>petit objectif</i></p> <p>定个/小/目标(corpus B) dìng / gè / xiǎo / mù biāo fixer / un mot spécifique / petit / objectif <i>fixer un petit objectif</i></p>	<p>Ce terme a été employé par un homme d'affaire très connu en Chine durant un talk-show au cours duquel il était invité à partager des conseils aux jeunes Chinois qui veulent devenir riches. Il a répondu que c'était une bonne direction mais qu'il valait mieux commencer par se fixer un petit objectif, par exemple gagner 100 millions. Ses paroles ont provoqué une vive réaction du public pour qui cet objectif est impossible à réaliser.</p> <p>Tout de suite, ce terme est devenu populaire en ligne et employé par ironie comme une antiphrase. C'est-à-dire que le petit objectif est en effet ambitieux et difficile à atteindre.</p>

15	友谊的小船，说翻就翻 (corpus A) yǒu yì de xiǎo chuán, shuō fān jiù fān [PH]  友谊的小船 (corpus B) yǒu yì de xiǎo chuán [PH]	友谊/的/小船/说翻就翻 (corpus A) yǒu yì / de / xiǎo chuán / shuō fān jiù fān l'amitié / précédé d'un nom exprime la relation d'appartenance / petit bateau / se renverser à tout moment <i>le bateau d'amitié peut se renverser du jour au lendemain, ou à tout moment.</i>  友谊/的/小船(corpus B) yǒu yì / de / xiǎo chuán l'amitié / précédé d'un nom exprime la relation d'appartenance / petit bateau <i>le bateau d'amitié</i>	Cette expression signifie en réalité la fragilité de la relation interpersonnelle. On peut ne plus être amis pour des raisons très simples : par exemple on ne mange pas épicé mais on est invité pour un repas de fondue chinoise très relevé. « 友谊的小船 » au sens littéral signifie : le bateau d'amitié. Il se peut qu'il soit issu du mot anglais « friendship » (amitié) et que la traduction chinoise ait divisé le mot en deux : « friend » traduit comme l'amitié et « ship » comme bateau. On peut aussi constater ce jeu de mot dans les dialogues de la fameuse sitcom américaine <i>Friends</i> . Cette expression devient populaire en ligne grâce aux bandes dessinées de Nan Dongni qui montrent deux pingouins sur un petit bateau qui chavire à cause de différents changements de ces deux animaux (ex. : l'un a maigri).  Ensuite, les internautes ont créé diverses dérivations en imitant la structure. Par exemple : le bateau du mariage, le bateau du travail. Cette expression a tellement de succès qu'on peut être compris par les autres sans écrire la phrase complète.
16	葛优躺 gě yōu tāng [PH]	葛优/躺 gě yōu / tāng nom d'un acteur chinois / s'allonger Geyou s'allonge	Depuis 2016, une série d'images macro <sup>11</sup> est apparue en ligne. Le héros de ces images, est un personnage du feuilleton télévisé « 我爱我家, wǒ ài wǒ jiā », qui s'allonge sur le canapé après avoir bien mangé. Comme l'acteur s'appelle 葛优 (gě yōu), les internautes ont nommé cette position paresseuse par le nom d'acteur. On suppose que la réussite de cette position et ce terme répond aux besoins psychologiques de ces usagers qui subissent la pression sociale et un rythme de vie accéléré.
17	套路 tàò lù [SS]	套/路 tàò / lù série / moyen <i>un ensemble de méthodes</i>	L'expression « 套路 » existe déjà dans la langue chinoise. Elle signifie une série de mouvement d'arts martiaux ou un ensemble de méthodes. En ligne, surtout dans les jeux, cette expression a été dotée d'un nouveau sens. « 套 » est interprété comme le piège, donc « 套路 » devient « des stratégies de tromperie ». Ces tromperies sont souvent préplanifiées en suivant certaines modalités. Après avoir connu une utilisation élargie, elle est souvent employée péjorativement.

<sup>11</sup> Une image macro (terme d'origine anglaise) est un élément de support numérique composé par une image et des éléments textuels superposés pour créer souvent des effets humoristique ou ironique. Il s'appelle en chinois 表情包 (biǎo qíng bāo).



18	<p>蓝瘦，香菇</p> <p>lán shòu, xiāng gū</p> <p>[MS]</p>	<p>蓝/瘦/香菇</p> <p>lán / shòu / xiāng gū</p> <p>bleu / mince / champignon shiitake</p> <p><i>bleu et mince, champignon shiitake</i></p>	<p>« 蓝瘦，香菇 » est paronymie de « 难受，想哭 » (nán shòu, xiǎng kū) qui signifie trop triste et avoir envie de pleurer. Donc, il est difficile de connaître le vrai sens sans avoir cette connaissance préalable.</p> <p>L'origine de cette expression est liée à un internaute venant de la province de Guangxi. Il a posté en ligne une vidéo exprimant sa tristesse après s'être séparé de sa copine. En raison de son fort accent, les internautes n'entendent pas clairement « 难受，想哭 » et ensuite le transcrivent et créent cette expression drôle « 蓝瘦，香菇 ».</p> <p>Nous supposons que le choix du mot « 蓝 » (bleu) n'est pas dû au hasard. On joue sur deux aspects : 1) la couleur bleue a une connotation spéciale, la tristesse. 2) « lan » à la place de « nan » renforce la présence de l'accent (les gens du Sud de la Chine ont de la difficulté à distinguer la prononciation L et N). En plus, « 瘦 » signifie maigrir. Il est raisonnable de penser que quand on est trop triste, on perd du poids. Le remplacement de « 想哭 » par « 香菇 », est aussi destiné à souligner l'influence de l'accent en ajoutant des effets humoristiques.</p> <p>Cette expression est devenue populaire en très peu de temps et a circulé dans toute la Chine quelques heures après sa création. Par la suite, elle est présente sur des supports différents, en ligne et hors ligne. Dans certaines cantines, on a aussi créé et nommé des plats avec cette expression, avec bien sûr des champignons comme ingrédient. On observe même la création d'une société utilisant cette expression comme nom de marque.</p>
19	<p>油腻</p> <p>yóu nì</p> <p>[SS]</p>	<p>油/腻</p> <p>yóu / nì</p> <p>gras / grassex;</p> <p>gras et lourd</p>	<p>« 油腻 » signifie initialement quelque chose de trop gras. En 2017, un usager de <i>Sina Weibo</i> a publié un tweet dans lequel ce terme est employé pour décrire certains caractères péjoratifs des personnes d'un âge moyen. Depuis, ce terme est porteur d'un nouveau sens. Il est destiné à montrer les apparences et les comportements qui provoquent un sentiment désagréable voire écœurant, pas seulement concernant des personnes d'âge mûr.</p> <p>En même temps, on peut constater la reconnaissance et l'utilisation de ce terme dans les tweets des célébrités, des influenceurs, etc.</p>
20	<p>怼</p> <p>duǐ</p> <p>[MI]<sup>12</sup></p>	<p>怼</p> <p>duǐ</p> <p>haïr du fond du cœur</p>	<p>La prononciation initiale de « 怼 » est <i>duì</i> qui signifie haïr (emploi comme verbe) ou cruel (adjectif). La racine de ce mot est « 心 » (xīn, le cœur). Cela montre un sentiment haineux profond, venant du fond du cœur. Dans le dialecte de la province Henan, « 怼 » peut être utilisé comme un verbe universel : offenser, critiquer selon le contexte.</p>

<sup>12</sup> Dans le chapitre V, nous avons ajouté une catégorie indépendante aux matrices MS, SS, PM et PH dans les matrices internes pour souligner les influences des dialectes en chinois et en français, les influences des variants de français régional et non hexagonal. Ici nous utilisons MI (matrice interne) pour dénommer cette catégorie et pour distinguer

		<i>haineux</i>	Son nouveau sens populaire en ligne, est issu d'une émission de télé-réalité, où ce terme est employé avec le sens de « réfuter ». C'est un emprunt à un dialecte mais certains estiment que, c'est peut-être une fausse utilisation du dialecte puisque qu'il existe un autre caractère « 搥 » avec la même prononciation et un sens rapproché. Suite à sa propagation, de nouvelles interprétations comme : disputer exprès, rivaliser, sont aussi ajoutées pour ce terme. Il est aussi beaucoup présent dans des articles de presse.
21	打 call dǎ call  [ME]	打/call dǎ /call taper / téléphone <i>téléphoner</i>	Cette expression constitue un terme de la culture des fans, particulièrement les fans de vedettes (wota), dans le domaine des concerts Live. Ce genre de culture s'est d'abord développé en Corée du Sud et au Japon. Il englobe différents moyens de soutien des fans pour leurs idoles. « 打 call » est créé à partir du mot japonais « コール ». Par « 打 call », les fans crient sur place ou agitent par exemple des tiges fluorescentes, éventails, bandeaux pour montrer leur enthousiasme en créant une atmosphère chaleureuse. C'est un moyen d'interaction entre les fans et leurs idoles. Il faut souligner que ces interactions sont strictement préplanifiées selon le rythme de chanson. Nous pouvons supposer que l'emploi du caractère chinois « 打 » est pour décrire l'action ; « call » est une transcription phonétique de l'expression japonaise. Une fois introduit en Chine, son utilisation n'est pas restreinte au domaine mentionné ci-dessus. Il signifie généralement, encourager quelqu'un à faire quelque chose, comme « allez » dans la langue française. Parfois, les usagers négligeant cet aspect culturel, considère « call » comme téléphone. Cela faisant, on exprime l'encouragement aussi par « 打电话 » (电话 signifie téléphone). Or, l'expression « 打电话 » signifie originellement « téléphoner » en chinois.

---

des autres procédés de la typologie de référence.

### Annexe 3 : Documents relatifs au corpus de listes français

3a-1 : Tableau des nouvelles entrées des dictionnaires rapportées par *Le Figaro*

<a href="#">Petit Larousse édition 2014</a>	cougar	<a href="#">Petit Larousse édition 2012</a>	belgitude	<a href="#">Petit Robert édition 2011</a>	mamavirus	<a href="#">Petit Larousse édition 2011</a>	fashionista
	botoxé		virelangue		nanosciences		bisphénol
	speed dating				biogaz		footeux
	Nomophobe				smoothie		Métrosexuel
	googliser				cranberry		Locavore
	prioriser				l'auto-scooter		Google
	textoter				frisko		Wikipédia
	post				panty		pop-up
	hashtag				Agender		cheese-cake
	Flashcode				chocolatine		scrapbooking
	cyberdéfense				aidant		coloc
	se désinscrire				se reconstruire		nerd
	démondialiser				diversité		
	biométhane				NAC		
	écoblanchiment				prime à la casse		
	lombricompostage				droit à l'image		
	climatoseptique				réalité augmentée		
	chialeux, euse				enfants soldats		
					homéotique		
					nocicepteur		
					biologie cellulaire		
					biologie moléculaire		
					virologie		

<a href="#">Petit Larousse édition 2017</a>	Burrata	<a href="#">Petit Larousse édition 2016</a>	Crudivore	<a href="#">Petit Robert édition 2016</a>	baltringue	<a href="#">Petit Larousse édition 2015</a>	cigarette électronique
	ciabatta		Bistronomie		Big data		vapoter
	argouse		Cuisine moléculaire		réenchanter		vapoteur, euse
	phô		vegan		hystériser		hystérisation
	wrap		cœur-de-bœuf		écoconduite		psychoter
	yuzu		électrosensibilité		écoblanchiment		surinterpréter
	arty		algologie		agroécologie		zénitude
	cinégénique		arthroscanner		butternut		écoresponsable
	fanfiction		génériquable		cari		marée verte
	Mook		histocompatibilité		wakamé		biopétrole
	Rétrofuturisme		palper-rouler		entomophagie		empreinte carbone
	spin-off		capitaliser		webdesign		noniste
	Fablab		MOOC		cybercrimes		candidater
	flasher		clivant		cybersexe		annus horribilis
	open source		surréagir		adulcents		post-ethnique
	QR code		zénifiant		zadiste		Accumulateur de froid
	troll		glamouriser		déconnade		Animatronique
	émoticône		rétropédaler				Consolidation budgétaire
	déréférencement		Avoir le melon				Crise de liquidité
	défaçage (ou défacement)		community manager				État-parti
	téléverser		selfie				Iconique
	Autocentré		chnequer				Imprimante 3D
	covoiturer		siester				Le made in
	socialiser		réciproquer				Micropaiement
	zone de confort		dosha				Motion capture
	économie		shakti				Multivers
	collaborative		fictionniser				Notation souveraine
	zadiste		microfiction				Obsolescence programmée
	loup solitaire		dézonage				Périurbanisation
	antisystème		Atomiser				Psychologisation
	complotisme		mémériser				Psychoter
	déclinisme		Dédiabolisation				Scud
	europhobe						Surinterpréter
	exfiltrer						Tag
	lobotomiser						Tourisme spatial
	REP						Viralité
	Zika						Vol suborbital
	S'enkyster						
	patate						
	ambiancer						
	sans-allure						

### 3a-2 : Tableau des nouvelles entrées des dictionnaires rapportées par 20 Minutes

<a href="#">Petit Larousse édition 2017</a>	yuzu	<a href="#">Petit Robert édition 2016</a>	covoiturer	<a href="#">Petit Larousse édition 2016</a>	bolos
	phô		écoconduite		mémériser
	wrap		recyclerie		tutos
	mook		zadistes		selfies
	seul-en-scène (one-man-show)		faucheurs volontaires		
	autocentré		décroissants		
	émoticône		bitcoin		
	loup solitaire		captcha		
	zadiste		big data		
	antisystème		baltringues		
	europhobe		prendre cher		
	complotiste		tendu comme un string		
	déclinisme		envoyer des scuds		
	balado		beuh		
	s'ambiancer		partir en cacahuète		
	sans-allure				
	yoyette				
	débit de kalach				

<a href="#">Petit Larousse édition 2015</a>	Scud	<a href="#">Petit Larousse édition 2014</a>	speed dating
	noniste		monter au filet
	vapoter		googliser
	botoxé		hashtag
	wushu		post
			extoter
			cyberdéfense
			même
			climatosceptique
			cougar
			flashmob
			marche blanche
			poilade
			orwellien
			méthanisation
			lombricompostage
			microfinance
			réindustrialiser
			voxographie
			préquelle
			slopestyle
			telemark

### 3a-3 : Tableau des nouvelles entrées des dictionnaires rapportées par *Le Parisien*

<a href="#">Petit Larousse édition 2017</a>	<a href="#">Petit Robert édition 2017</a>	<a href="#">Petit Robert édition 2013</a>
yuzu	youtubeur (euse)	corium
wrap	twittosphère	liquidateur
Troll	geeker	mox
émoticône	émoji	parc éolien
zadiste	nomophobie	marée verte
complotiste	ubériser	gaz de schiste
QR code	pad thaï	agence de notation
déréférencement	piquillo	dette souveraine
Zad	alfalfa	Indignés
Antisystème	viandard	oscariser
europhobe	s'enjailler	comater
mook	chouille	pipeauter
néokeynésianisme	vivre-ensemble	psychoter
	matinalier	marrade
	fixie	subclaquant
	aquabike	gloups
		acérola
		bento
		bobun
		makrout
		notebook
		netbook
		biopic
		lol
		passer dans le beurre
		les bobettes
		panade
		prépension

### 3a-4 : Tableau des nouvelles entrées des dictionnaires rapportées par RFI

<a href="#">Petit Larousse édition 2017</a>		<a href="#">Petit Larousse édition 2016</a>		<a href="#">Petit Robert édition 2016</a>	
	yuzu		sape		bitcoin
	émoticône		selfie		ZAD
	zadistes		big data		zadiste
	avoir la patate		open data		big data
	covoiturer		community manager		captcha
	argouse		particule fine		covoiturage
	burrata		circulation alternée		gyropode
	wrap		bolos		beuh
	troll		bistronomie		tendu comme un string
	loup solitaire		anthropocène		maquillée comme un camion volé
	ZAD		dédiabolisation		
	stand up paddle		bolos		
	pole dance				
	spin-off				
	mook				

<a href="#">Petit Larousse édition 2015</a>		<a href="#">Petit Robert édition 2015</a>		<a href="#">Petit Robert édition 2013</a>	
	tuerie		abdo-fessiers		Groupes
	vapoter		pilate		Dette souveraine
	se loguer		vétetiste		indignés
	cachaça		zumba		belgitude
			selfie		cyberdépendance
			emmieuter		billet
					nuage
					informatique
					mémoire flash
					permalien
					réseautique
					rétrolien
					anosognosie
					Psychoter
					comater
					subclaquant
					à l'arrache
					biopic
					dystopie
					œuvre orpheline
					oscariser
					passer dans le beurre
					taxage
					prépension
					panade

### 3b-1 : Tableau récapitulatif de nouvelles entrées citées par les organes de presse

(Dictionnaire *Le Petit Robert*) corpus Af

les organes de presse		Figaro	20 Minutes	Le parisien	RFI
Année d'édition Petit Robert	2017			16	
	2016	17	15		10
	2015				6
	2014				
	2013			29	24
	2012				
	2011	23			

### 3b-2 : Tableau récapitulatif de nouvelles entrées citées par les organes de presse

(Dictionnaire *Le Petit Larousse*) corpus Bf

les organes de presse		Figaro	20 Minutes	Le parisien	RFI
Année d'édition Petit Larousse	2017	40	18	13	15
	2016	32	4		11
	2015	37	5		4
	2014	18	22		
	2013				
	2012	2			
	2011	12			



### 3c : Tableau récapitulatif des entrées du corpus de listes français

classées par matrice

Les influences de l'anglais
les influences des autres langues étrangères

MS-CA				
agender	déclinisme	histocompatibilité	prioriser	vétetiste
agroécologie	déconnade	homéotique	psychologisation	virologie
algologie	dédiabolisation	hystérisation	psychoter	voxographie
anthropocène	défaçage (ou défacement)	hystériser	recyclerie	youtubeur (euse)
antisystème	démondialiser	locavore	réenchanter	zadiste
argouse	déréférencement	marrade	réindustrialiser	zénifiant
arthroscanner	dézonage	matinalier	réseautique	zénitude
autocentré	dystopie	mémériser	Rétrofuturisme	
belgitude	écoblanchiment	méthanisation	rétrolien	
biogaz	écoconduite	microfiction	sape	
biométhane	écoresponsable	microfinance	se désinscrire	
biopétrole	électrosensibilité	micropaiement	se loguer	
bisphénol	entomophagie	multivers	siester	
botoxé	europhe	nanosciences	socialiser	
capitaliser	exfiltrer	néokeynésianisme	subclaquant	
chocolatine	fictionniser	noniste	surinterpréter	
cinégénique	footeux	orwellien	surréagir	
comater	geeker	oscariser	téléverser	
complotisme	génériquable	périurbanisation	textoter	
covoiturer	glamouriser	pipeauter	ubériser	
crudivore	googliser	poilade	vapoter	
	gyropode	post-ethnique	vapoteur, euse	

MS-CC		SS
abdo-fessiers	lombricompostage	aidant
accumulateur de froid	loup solitaire	atomiser
adolescents	marche blanche	billet
agence de notation	marée verte	bolos
animatronique	mémoire flash	clivant
anosognosie	métrosexuel	décroissants
biologie cellulaire	mook	diversité
biologie moléculaire	nocicepteur	Flashcode
bistronomie	notation souveraine	flasher
cigarette électronique	nuage informatique	iconique
circulation alternée	obsolescence programmée	Indignés
climatoseptique	œuvre orpheline	liquidateur
cœur-de-bœuf	parc éolien	lobotomiser
consolidation budgétaire	particule fine	monter au file
crise de liquidité	permalien	palper-rouler
cuisine moléculaire	prime à la casse	pilate
cybercrimes	réalité augmentée	rétropédaler
cyberdépendance	scrapbooking	scud
cybersexe	seul-en-scène	se reconstruire
dette souveraine	téléverser	s'enkyster
droit à l'image	tourisme spatial	telemark
économie collaborative	twittosphère	troll
émoticône	virelangue	tuerie
empreinte carbone	vol suborbital	viandard
enfants soldats		viralité
État-parti		vivre-ensemble
faucheurs volontaires		zika
gaz de schiste		
imprimante 3D		

MS-ID	PM	ME		
baltringue	coloc	acérola	fixie	spin-off
beuh	lol	alfalfa	flashmob	tag
gloups	mamavirus	aquabike	hashtag	wakamé
s'enjailler	MOOC	arty	Le made in	webdesign
	mox	bento	makrout	wrap
<b>les variants du Français</b>	NAC	big data	même	wushu
ambiancer	nomophobe	biopic	motion capture	yuzu
auto-scooter	QR code	bitcoin	nerd	zumba
balado	REP	bobun	netbook	
cari	tutos	burrata	notebook	
chialeux, euse	vegan	butternut	open data	
chneuquer		cachaça	open source	
chouille	<b>PH</b>	captcha	pad thaï	
emmieuter	à l'arrache	cheese-cake	panty	
frisko	annus horribilis	ciabatta	phô	
les bobettes	avoir le melon	community manager	piquillo	
panade	débit de kalach	corium	pop-up	
passer dans le beurre	envoyer des scuds	cougar	post	
prépension	maquillée comme un camion volé	cranberry	préquelle	
réciproquer	partir en cacahuète	dosha	selfie	
sans-allure	patate/avoir la patate	émoji	shakti	
taxage	prendre cher	Fablab	slopestyle	
yoyette	tendu comme un string	fanfiction	smoothie	
	zone de confort	fashionista	speed dating	

### **3d : Tableau préparatoire au classement par matrices des entrées du corpus de listes français**

#### **NOTA BENE**

L'annexe 3 d ne figure que dans la version électronique de la thèse et hors pagination pour des raisons de lisibilité des cellules du classeur Excel, le fichier Excel est donc autonome.

Cette annexe reproduit le travail préalable au classement par matrices des nouveautés langagières du corpus de listes français. Chaque feuille du classeur reprend par média les entrées lexicales qui sont ensuite classées par année d'intégration aux dictionnaires avec pour les colonnes renvoyant au dictionnaire un lien hypertexte pour accéder aux articles consultés.

Chaque entrée figure dans une cellule qui comporte les informations suivantes :

entrée [**MATRICE PRINCIPALE** explications et éléments de définition obtenus par consultation de sources lexicographiques et encyclopédiques diverses] *définition brève*

Les colonnes grisées (par exemple résultats du journal *Le Monde* finalement non retenus ou les entrées de l'année 2018) ne sont pas prises en compte pour le récapitulatif du classement par matrices présenté dans l'annexe 3c et le calcul proportionnel des matrices présenté dans le chapitre V.

## Annexe 4 : Le corpus de tweets chinois

### 4a : Le terme n°1 元芳，你怎么看 (yuán fāng, nǐ zěn me kàn)

1.

🌟 热门



厦门天气在线 ✓

【投票：#鼠年太长该高兴还是难过# 🐭】即将到来的农历庚子鼠年共384天，双闰年、双闰月、双立春...不过这也意味着#鼠年要多上一个月的班#。面对超长待机的鼠年，该高兴还是难过？#元芳,你怎么看# 小投票↓↓ 鼠年要多上一个月的班，对于此事

01月05日 11:02 来自 微博 weibo.com

收藏

转发 5

评论 2

👍 45

2.



OdTu ✓

#元芳，你怎么看#



2016年07月24日 17:33 来自 三星android智能手机

收藏

转发

评论

👍

3.



主编赵旭东

#长租公寓转手致租客被撵#老赵说事#长租公寓转手致租客被撵#你在经营不善，把租客的钱收了就要赶人，有1600套房源的悦如公寓因经营不善被转手，结果悦如却没把租金都付给房东，房东将房屋换锁了。这种公司是不是应该负法律责任？ #元芳,你怎么看#

2019年10月11日 20:56 来自 三星GALAXY A

收藏

转发 3

评论

👍 2

4.



Grams21

最流行奶奶灰加半扎丸子头#元芳, 你怎么看#



2016年04月06日 11:00 来自 iPhone 6 Plus

收藏

转发

评论

👍

5.



交通言究社

今天看到新闻说长春一辆坦克追尾了一辆厢式货车😱！这可是真事！一般来说，道路上发生追尾事故，基本都是后车全责，但那是针对机动车而言。现在好家伙，追尾的是一辆坦克诶！交警蜀黍们打算怎么判责？保险能赔吗😱？#元芳，你怎么看# → 网页链接



2015年04月17日 11:02 来自 微博 weibo.com

收藏

转发 4

评论 3



6.



江淮晨报

【孙楠谈退赛 原谅我的老不更事】中途退出#我是歌手#总决赛, @孙楠 受访称:“我跟汪涵很多年的朋友, 他不会生气...我听说这一段让收视提高不少, 芒果台应该感谢我。”这一言论引发不小争议。今天下午, 孙楠微博致歉:“请原谅我的鲁莽、任性、老不更事。”有人说是故意炒作 #元芳, 你怎么看#



2015年03月29日 16:37 来自 360安全浏览器

收藏

转发 4

评论

14

7.

 **idog** ★

追完第一季。如果不考虑福叔的因素，我最多把它当#名侦探柯南#看。因为#福山雅治#的原因，我权且继续跟进第二季和剧场版吧。之前对该剧期望值颇高，现在.....比韩剧的逻辑还是多一点。[#元芳，你怎么看#](#) 😊

**@idog** ★

抛弃日剧好久好久。这些年看韩剧流失的智商，就从#神探伽利略#开始往回补吧。[#福山雅治#](#)真是不怕老啊！第一季第一集就有[#唐泽寿明#](#) 😲 surprise!!! 🤔 [玫瑰]@糗糊的人生



2014年07月26日 00:58 来自 HTC Phone 转发 2 | 评论 6 | 1

2014年07月27日 21:41 来自 iPad客户端

收藏 | 转发 | 评论 2 | 1

8.

 **大连市普兰店区法院** ✓

[#元芳，你怎么看#](#)【店主抓小偷泄愤 挂牌跪地引争议】小编语：小偷固然可恶，可是国有国法，滥用私刑体罚侮辱小偷践踏了小偷的人权，也触犯了法律。还是该送至派出所，这样处罚小偷才名正言顺~~

**@中国警方在线** ✓

【店主抓小偷泄愤 挂牌跪地引争议】近日，温州某服装市场，一女子行窃时被店主抓获。女子被挂一块写“我是小偷(应为偷)，该打”的牌子，跪在地上，表情痛苦。几经“折腾”，女子被保安带走。[温州店主抓女贼泄愤 挂牌跪地引发争议](#) 提示：滥用私刑擅自体罚侮辱小偷，属以暴制暴，既践踏了人权，也触犯了法律。[你怎么看？](#)



2014年07月13日 17:23 来自 百度浏览器 转发 163 | 评论 301 | 52

2014年07月14日 08:44 来自 微博桌面

收藏 | 转发 | 评论 | 1



9.

 **松阪牛** V  
#元芳，你怎么看#



2013年05月10日 13:57 来自 Weico微可拍

收藏 | 转发 2 | 评论 8 | 

10.

 **广州公安** V  
#2012十大流行语公布#昨日，语文期刊《咬文嚼字》正式发布了“2012年十大流行语”，分别是：[#正能量#](#)，[#元芳，你怎么看#](#)，[#舌尖上#](#)，[#躺枪#](#)，[#高富帅#](#)，[#中国式#](#)，[#亚历山大#](#)，[#最美#](#)，[#赞#](#)，[#接地气#](#)。童鞋们，您们认为2012年还有什么更流行的流行语呢？[走你][江南style][圆蛋快乐] via@平安北京

2013年01月01日 13:22 来自 iPhone客户端

收藏 | 转发 33 | 评论 10 | 

## 4b : Le terme n°2 女汉子 (nǚ hàn zi)

1.



狐狸养的笨小喵 🐱

打造出租屋是很累的😓

女汉子上线，以后一定买个电动螺丝刀  
手动的太累了😓 #出租屋改造##女汉子#



05月09日 00:09 来自 喵喵拳Android

收藏

转发

评论 8

👍 3

2.



hmmhmmmm ★

为什么别人一抬手都是Tiffany Catier Pandora...

我们一抬手就是.....发圈😂😂 #女汉子# ?



2017年10月27日 12:19 来自 iPhone

收藏

转发 1

评论 1

👍 2

3.

 **Me\_石榴** ▼

[#女生节的正确打开方式#](#)各种买买买。。。电动车，电饭煲，布衣柜，锅碗瓢盆……一个人搞定，请叫我[#女汉子#](#) 🤖 📍 [新乡·长垣县](#)



2017年03月07日 23:08 来自 OPPO R9

---

收藏 | 转发 | 评论 5 | 5

4.

 **看新西兰** ▼

新西兰最名副其实的[#女汉子#](#) --瓦莱丽·阿达姆斯，目前保持着女子铅球的世界纪录（与另外一位运动员一起），同时她也独霸着新西兰、大洋洲、英联邦等等赛事的[女子铅球纪录](#)。她的个人比赛最好成绩曾经扔出过21.24米的距离，是名副其实的“超级[女汉子](#)”。



2016年08月01日 17:31 来自 iPhone 6s Plus

---

收藏 | 转发 | 评论 | 5

5.

 **992一路畅行** ▼

**#992晚高峰#** 【武汉：为不让男友湿鞋，“中国好女友”背男友过马路】昨天，湖北武汉市，光谷大道与金融港三路交叉口附近积水未全退，市民仍需脱鞋涉水过马路，一女子为不让男友湿鞋，执意要背着他过马路，尝试一下背男友的感觉。今日话题**#女汉子#**你周围有**女汉子**么？说说她们的故事。



2016年07月05日 17:20 来自 360安全浏览器

---

收藏 | 转发 5 | 评论 21 |  2

6.

 **大圣大战来福** ▼

**@鹧鸪动物园** 园长，大圣这算不算**#奇葩睡姿#** 或者**#女汉子#**



2016年06月22日 23:09 来自 iPhone 6

---

收藏 | 转发 | 评论 | 

7.



诚信广东 ✓

**#女汉子#**【一声怒吼震住歹徒 美女教官刀下救人】13日下午，广州公安边防部队高等专科学校女教官章荣下班跑步锻炼时，见一男子正用刀捅另一男子。她迅速堵在歹徒身前：“你在干吗？给我住手！”后歹徒落荒而逃，伤者随后获救。据悉，章教官“平时就仗义”，更是学生心中的女神级人物。 [☞一声怒吼震住歹徒 女神教官刀下救人](#) [收起全文](#) ^



2016年05月21日 12:34 来自 三星android智能手机

收藏

转发

评论



8.



依含羞草 ✓

**#女汉子#** 女汉子的定义应该有纯爷们的性格，纯爷们儿能做的事，女人也都行啊！那还要男人干啥呢？女人可以做到不依附于男人，坚持自我，对命运不论顺境逆境都能安排好，不需要小鸟依人，柔情似水，唉，做自己的女神吧！



2016年04月11日 00:28 来自 华为P8

收藏

转发

评论



9.

 **Sun\_ny小希望** 

**#今日贴纸打卡##女汉子#** 发现一个真理，男生都喜欢小小的一只的女生，而大骨架的大只的我，注定得不到疼爱，我就是传说中的女汉子喽    [@大连·海军广场街区](#)



2016年04月01日 19:15 来自 iPhone客户端

收藏 | 转发 | 评论 4 |  1

10.

 **女汉子真爱公式** 

**#女汉子#**听说很多地方的高考分数线已经出炉，是时候再次把学霸锦鲤祭出来了！小五兄提前预祝各位参加高考的童鞋都能选中理想大学，走上人生巅峰哦~**#2015高考分数线#**



2015年06月23日 17:12 来自 微博 weibo.com

收藏 | 转发 20 | 评论 15 |  177

#### 4c : Le terme n°3 高大上 (gāo dà shàng)

1.



杭城陈杭 🏆

#高大上# 我们所谓的高大上，无非就是洋人们的生活日常，遵循自己才是真正的高大上。



2019年06月28日 07:47 来自 iPhone客户端

收藏

转发

评论 2

👍 5

2.



加泰罗尼亚旅游网 ✓

📍西班牙·桃乐丝酒庄 #葡萄酒# #高大上# 闻名世界的桃乐丝酒庄(Bodegas Torres), 创立于1870年。😍到巴塞罗那旅游的小伙伴们, 都是必去桃乐丝酒庄的。酒庄位于距离巴塞罗那市约45公里处的佩内德斯地区, 中文导游将贴心的给你带来葡萄酒和美食的难忘体验之旅, 逼格满满啊!



2017年10月26日 12:34 来自 微博 weibo.com

收藏

转发 4

评论 5

👍 11

3.



老杜找乐儿

CNN用无人机拍摄宾州投票现场#高大上##美国大选#



2016年11月08日 23:50 来自 iPhone 6s Plus

收藏

转发 5

评论 2

6

4.



共青团牡丹区委

#高大上#青年看过来，联合国教科文组织青年论坛请你参会 让更多人听到当代中国青年的理性声音 @菏泽学院校园微博 @青春菏泽 @山东共青团

@新华国际

#UNESCO联合国教科文组织青年论坛招募# 有为青年们，联合国叫你参会啦！只要你青春活力、乐于沟通，只要你热爱地球、关注气候，就有可能成为31位优秀青年代表之一，赴巴黎代表中国与世界互动。快戳链接报名吧 网页链接别犹豫，行动吧，无参与，不青春！ @完美世界教育 @全球青年领袖实验室



2015年05月29日 08:59 来自 专业版微博

转发 9606 | 评论 4537 | 53

2015年06月03日 09:00 来自 微博 weibo.com

收藏

转发 1

评论

5



5.



带着面具的人生

#高大上#的生活!!!



2016年06月07日 13:50 来自 荣耀7 世界有点不同

收藏

转发

评论



6.



中央民族大学

#民大资讯##高大上# 3月17日, 国际知名学术期刊NATURE上发表了我校生命与环境科学学院学科带头人桑卫国教授在研究论文China Draws Lines to Green Future。网页链接文章系统地阐述了我国土地保护红线的概念,该研究成果不仅为国家生态保护政策提供依据,还对我校生态学科建设与发展具有重要 展开全文



【超话】民大资讯

218 帖子170粉丝

+ 关注

2016年03月25日 11:17 来自 民大资讯超话

收藏

转发 3

评论 1

13

7.



海峡都市报闽南版

#高大上#【泉州公交卡月底刷遍全国50城】本月底, 泉州市民只需手持一张泉州公交IC卡, 便可在全国50个城市刷卡坐公交、地铁了。"刷卡金额即按照当地普通卡待遇, 如当地的一卡通可乘坐地铁、轮渡, 泉州的公交卡在当地也可实现。"网页链接



2014年10月11日 09:36 来自 微博 weibo.com



收藏

转发 34

评论 40

14


8.

 **扒拉扒拉的吧的吧**  ▼



小时候就应该学画画而不是乐器 学画画的能时不时在朋友圈里发下自己的作品 画得不好还可以说是抽象 我能怎么办 难道要录一段放上去😭 [#高大上#](#)


2015年05月01日 11:06 来自 iPhone客户端

---

收藏 | 转发 | 评论 | 


9.

 **蕾蕾\_酱**  ▼

我是个半高不矮，没发型，近视眼，吃货胖子，爱哭鬼，做事粗心三分钟热度，没心机的女汉子，希望在新的一年里2014里我能变的[#高大上#](#)往全能女神的方向靠近，每一天都能开心幸福 

2013年12月31日 20:25 来自 iPhone客户端

---

收藏 | 转发 | 评论 |  1

10.

 **介似天津卫**  ▼

[@然并卵的萌男](#):论天津市随处可见的[#英语神翻译#](#) 。真心敬佩咱天津人的英文水准。华都食品机械商店——Chinese have food machinery store，公交专用——BUS ZHUANYONG。好[#高大上#](#)的翻译啊  ！



2016年06月09日 20:22 来自 360安全浏览器

---

收藏 | 转发 2 | 评论 2 |  7

#### 4d : Le terme n°4 主要看气质 (zhǔ yào kàn qì zhi)

1.



**袁岳** ✓

【美女，真的有标准吗? | Mr.Data [美女，真的有标准吗? | Mr.Data](#)】  
什么是女性之美? 是唐寅的“赏花赏月赏秋香”，是王小波的“视力没有大毛病的都能欣赏”，还是一句时下流行的“我可以”? [#主要看气质#](#)




**美女，真的有标准吗? | Mr.Data**  
什么是女性之美? 是唐寅的“赏花赏月赏秋香”，是王小波的“视力没有大毛病的都能欣赏”，还是一句时下流行的“我可以”?

04月17日 13:46 来自 微博 weibo.com

收藏 | 转发 1 | 评论 6 | 9

2.



**\_MaryHuang** 👑

[#365天审美积累##主要看气质#](#) 对于渴望追求美的女孩们，自己的建议：“仅仅看服装并不够，一定要与艺术亲近，多逛博物馆看展览，多看电影，听听音乐会，还要多了解关于艺术的背景知识。这对于提升审美才是最重要的。📍长沙”

03月11日 21:28 来自 iPhone 8 Plus

收藏 | 转发 | 评论 2 | 2

3.




**阿坤的家kun** 👑

岁月会把一个女孩变成妇女  
而读书会把一个女人变成富女  
[#读书##主要看气质##婚姻故事#](#)

03月08日 21:32 来自 我怕穷Android

收藏 | 转发 | 评论 | 1


4.

 **Emotion-delayDE凉子** ▼

**#主要看气质#**昨天下午帮忙接侄儿，因为他还在打扫课室所以我就上楼等。结果其他小朋友有给我敬礼的，有对我喊完老师好才敢从我面前经过的...所以上楼的时候其他给我打招呼的大人也是把我当成老师了么 😊🙇

2019年07月06日 06:39 来自 HUAWEI P30

---

收藏 | 转发 | 评论 | 

5.

 **rob00oort** ▼

**#主要看气质#** 在2016年4月18日刚刚结束的由魏丽莉老师主讲的“现代礼仪与个人修养”讲座中，经济学一班的全体同学都受到了意义非凡的职场礼仪教育。发比衣重要，鞋比衣重要；搭比穿重要，内比外重要。讲真，**气质最重要**。 [📍兰州·夏官营镇](#)



2016年04月18日 22:34 来自 魅族 MX5

---

收藏 | 转发 | 评论 | 

6.



华农校园植物

#主要看气质# 兰科#绶草(Spiranthes sinensis)因其花序似绶带而得名，盛花期很容易辨认，平常只能看气质啦。所有野生兰科植物都是国家二级保护以上，乱采乱挖是犯法的哦。



2016年04月13日 23:42 来自 荣耀畅玩4C

收藏

转发 3

评论 4

👍 11

7.



北京无极堂书画院

#国画 主要看气质# 人生的意义在于创造，艺术可以提升人的生命境界，每做一件事，记得多加些巧思，在每件所做的事情上，发挥才能，加入你自己。@画家刘文顺



2016年06月07日 18:29 来自 超话

收藏

转发

评论

👍

8.



lucy1818\_nvq 👑

Morning!#主要看气质# 大家都说我气质好，当然咯，我从小就是在麻麻的书房里长大的，要上音，体，美，诗歌，外语，游戏等课，品种繁多。虽然辛苦，可是谁让古人的话让人中毒太深呢：“书中自有黄金屋，书中自有颜如玉”，“万般皆下品，唯有读书高”，“以貌侍君岂能长久”呢？



2016年07月24日 09:39 来自 华为P9手机摄影再突破

收藏

转发

评论

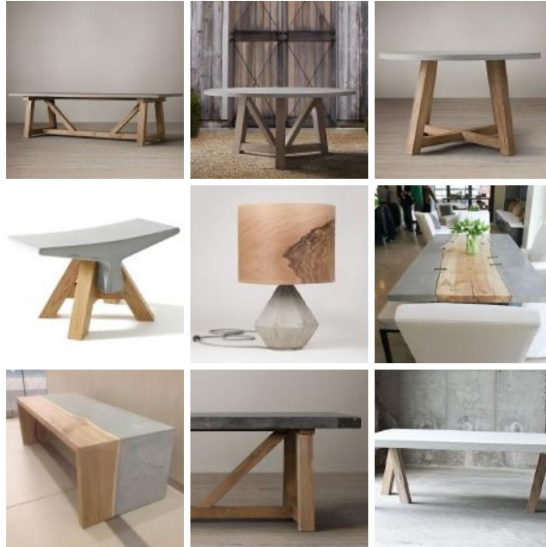


9.



木匠龙岛主 ✓

#主要看气质# 每一种材质都有它独特的无可替代的气质。



2016年08月21日 14:12 来自 搜狗高速浏览器

收藏

转发 1

评论 2



10.



鼠贷金融 **V**



**#主要看气质#** 如何选择P2P理财平台, **主要看“气质”!** 鼠贷金融, 一个有情怀有担当有良心的P2P理财平台, 你值得选择! 最重要的是选择鼠贷, 加息不断, 年化收益11.0+1%, 投资期限3个月, 你还在等什么? [网页链接](#)



2016年08月22日 10:01 来自 360安全浏览器

收藏

转发

评论



#### 4e : Le terme n°5 洪荒之力 (hóng huāng zhī lì)

1.



**天空是蓝蓝滴** ▼

**#洪荒之力#**

苍天，赐予我**洪荒之力**吧！  
让我顺利通过考试。😭😭😭

2017年08月17日 19:45 来自 OPPO智能手机

---

收藏 | 转发 | 评论 | 1

2.



**西安趣事儿** ▼

**#我为生活带颜#** 为了儿子，Micheal从一个180斤的胖子逆袭成健身达人，也从一个浪荡男孩蜕变成一个成熟的好爸爸。这是一个父亲的**#洪荒之力#**，为了爱，成为更好的自己❤️




2017年06月19日 20:29 来自 iPhone

---

收藏 | 转发 | 评论 | 1

3.



**好剧分享资源社** ▼

美食诚可贵，健康价更高。只是吃了一顿素的我都压抑不住我的**#洪荒之力#**了，心好累😭😭

😭想吃肉。没有肉的世界好悲惨。

2017年03月18日 12:45 来自 vivo智能手机

---

收藏 | 转发 | 评论 | 1



4.



小C管理5

#洪荒之力#是要靠平时的累积的，无论是游泳还是别的方面，所以我一有空就用 @听果 收听文学名著，补充能量啦~ @修德本人 @谷zzling @蓝白黑7优房



2017年05月11日 23:34 来自 听果

收藏

转发

评论



(publication virale) et ses variétés



big---bird

我的#洪荒之力#要用完了，需要在 @听果 听听有声小说补充一下啦，马上又和#傅园慧#一样很满意啦，我正在用 @听果 聆听 #涅槃重生——犹太人 第1集# 民族起源早期历史 主讲:西北工业大学 蒲瑶 @金牛蔚然Vivi



2017年05月04日 16:46 来自 听果

收藏

转发

评论



5.

 **湖北高速交警**  

蜀黎用尽**#洪荒之力#**为大家出行保平安 

**@湖北高速交警十堰大队** 

**#高速公路奇葩事#**一辆载货汽车所载的卡车轮毂散落一地，为了尽快打通道路，先期到达现场的 **@湖北高速交警十堰大队** 蜀黎用双手一个个将货物搬到路旁。总共散落了多少个呢？160个，每个70斤。蜀黎表示用尽了“**洪荒之力**”了！！ **@湖北交警@湖北高速交警 @湖北随州交警 @一叶知秋微直播 @公安部交通安全微发布 @湖北卫视 @湖北交通频道** [收起全文](#) ^



2017年06月10日 10:48 来自 360安全浏览器 转发 5 | 评论 2 |  3

2017年06月10日 10:54 来自 HUAWEI Mate 9

---

收藏 | 转发 3 | 评论 1 |  3

6.

 **举个栗子Lucky**  

**#洪荒之力#** 有了它妈妈再也不用担心我蹬不动 动感单车啦！so easy!  榆树·榆树大街



2017年07月30日 19:47 来自 全村的希望iPhone

---

收藏 | 转发 | 评论 |  3

7.



Ex-l



**#洪荒之力#**今天值班发现，单位**#海味儿单词串讲群#**大院里的石榴姑娘们已经胖到不行了，一开始我以为它们都是长不大那种的，就是土话说的看石榴。谁知道它们越长越大，大到手机屏幕都装不下它们了，这简直是用了**洪荒之力**在可劲长肉啊，想想也是，一下雨我们大院就发洪水，**洪荒之力**，果真霸气嚣张！石榴姐



2017年08月04日 09:28 来自 微博 weibo.com

收藏

转发

评论



8.



谢姣姣V



**#洪荒之力#** 这周写了无数无数的稿子和文件啊！红红火火恍恍惚惚！需要补脑！补脑！

“我爱工作，工作使我快乐 😊”



2017年08月25日 17:31 来自 iPhone 7

收藏

转发

评论



9.



讨厌鬼都看不到我微博



是我力气太大了吗 还是我的#洪荒之力# 显现了!



2018年05月06日 13:06 来自 OPPO R9

收藏

转发

评论 12

👍 6

10.



亲爱滴Mina 👑



以后尽量不当#老师#，我怕控制不住体内的#洪荒之力#，管不住想损人的那张嘴。被举报就糟糕了🐶🐶

2019年07月25日 09:43 来自 HUAWEI nova 2 Plus

收藏

转发

评论 1

👍

4f : Le terme n°6 蓝瘦，香菇 (lán shòu, xiāng gū)

1.



---南岸青栀

#蓝瘦，香菇# 吃了你，但愿明天鼻子不要堵了🙏



2018年08月24日 20:52 来自 iPhone客户端

收藏

转发

评论

👍 1

2.



我哋一片云

小狗狗失去了自由。#蓝瘦，香菇# 深圳·爱上咖啡 LoveCafe



2018年03月22日 11:38 来自 360手机Q5Plus

收藏

转发

评论 3

👍 1

3.



刮风要下雨嘞哦 🙏

#蓝瘦, 香菇#

从年初的一无所有

到年终的身无分文

一年白忙 哈哈.....😂😂😂📍奎屯市

2018年01月08日 14:12 来自 iPhone客户端

收藏

转发

评论 1



4.



新浪湖南教育

#蓝瘦, 香菇#近日, 网络热词“蓝瘦香菇”让众网友脑洞大开, 日前, 有食堂更将“蓝瘦香菇”搬上餐桌。19日, 湖南大学第二食堂正式推出创意料理“蓝瘦香菇”, 这也引来许多学生前来尝鲜。

😂😂😂📍湖南高校食堂推“蓝瘦香菇”



2016年10月20日 15:51 来自 微博 weibo.com

收藏

转发

评论



5.



醉美生活官网 ✓

都在问蓝瘦香菇是什么意思? 我来告诉你, 你没有办法免费吃喝玩乐游, 拥有醉美生活VIP的伙伴们可以免费吃喝玩乐游, 你就蓝瘦, 你没有醉美生活VIP卡, 各种不如意, 不滋润, 你就会蓝瘦到香菇! 🙏嗯, 大概就是这个词~ ~#蓝瘦, 香菇#

2016年10月14日 14:47 来自 微博 weibo.com

收藏

转发

评论



6.

 **有未旅行** ✓


[#蓝瘦, 香菇#](#) 蓝天碧海~~在落潮时分显露出来一条白色的陆上走廊的长沙滩, 会抚慰你的“[蓝瘦](#)” 🐼 潜水圣地~~在拥抱海洋且被海洋拥抱的那一刻, 会让你忘掉“[香菇](#)” [#爱在帕劳#](#)



2016年10月14日 13:58 来自 微博 weibo.com

收藏 | 转发 | 评论 | 1

7.

 **聪而不惠** ✓

孩子二年级了, 学习状态还不如一年级, 作业少, 错误多, 也不愿意多做题目弥补自己的不足。在今时今日这种竞争激烈的教育环境中, 完全淡定不了, 每天的咆哮和歇斯底里, 换来的结果却依然很惨[#蓝瘦, 香菇#](#)

2016年10月13日 21:39 来自 iPhone 7

收藏 | 转发 | 评论 | 1

8.

 **Lengow法国乐售** 🇫🇷 ✓

尴尬! 库克深圳买咖啡Apple Pay支付失败 当时库克的心情可以用四个字形容“[#蓝瘦, 香菇#](#)”


🔗尴尬! 库克深圳买咖啡Apple Pay支付失败



2016年10月13日 20:15 来自 亿邦动力网

收藏 | 转发 | 评论 | 1

9.

 **LovelyDona** ▼

[#蓝瘦, 香菇#](#)真不知道这种话题是怎样火起来的, 让人觉得特别无聊、低级趣味, 请问这里面有一丝一毫的意义吗? 而大众为何又要跟随这种无聊的“潮流”, 从滑板鞋到叶良辰, 人云亦云之后几乎是永远的寂静……背后的推手又是何意呢。这种现象, 真是占用了我们宝贵的认知资源, 且勿听勿言勿传吧。😓

2016年10月13日 11:13 来自 iPhone 6s

---

收藏 | 转发 | 评论 1 | 5

10.

 **仪征气象** ✓ ▼

今天多云到阴有时有小雨, 偏东风3到4级, 今天最高气温19到20°C, 明晨最低气温16到17°C。早晨天阴沉沉的, 雷达上周边零散的回波不少, 出门外套雨披是必备的啦! 这种天气, 感觉今年的秋天biu一下就完了, 完全没有秋高气爽的情怀[#蓝瘦, 香菇#](#)



**厉Hello world哥!**

2016年10月13日 06:33 来自 微博 weibo.com

---

收藏 | 转发 3 | 评论 2 | 5



## 4g : Le terme n°7 打 call (dǎ call)

微博搜索 #打call# 搜索 高级

 #打call# 分享 申请主持人  
阅读926.8万 讨论1.7万

综合 实时 热门 视频 图片

1.

 银力体育官微  

#健身生活##打call#

每个人的梦想都不是渺小的  
希望你永远保持热爱  
在自己的世界里  
闪闪发亮  
#周末##早安#



06月28日 08:44 来自 iPhone客户端

收藏 | 转发 | 评论 | 

2.

 myxis 

今晚鹤唳华亭#打call#

04月21日 15:22 来自 OPPO随光而变R17

收藏 | 转发 | 评论 | 

3.



苏州高新区文体旅局

#打Call#厉害了！蓝海创意云平台入选年度“一带一路”国际合作重点项目！厉害了！蓝海创意云平台入选年度“一带一路”国际合作重点项目！



2019年09月26日 15:51 来自 皮皮时光机

收藏

转发

评论



4.



一只九头鸟

#为朱Queen 打Call #，继续加油。

@新浪排球

朱婷 土超决赛第二回合，瓦基弗银行3:1伊萨奇巴希，朱婷怒砍22分冠绝全场：朱婷22分土超决赛瓦基弗3-1



2019年04月27日 09:32 来自 专业版微博

转发 10 | 评论 14 | 157

2019年04月27日 09:41 来自 iPhone客户端

收藏

转发

评论



5.

 **湖南司法警官学院警院之声广播站**  ▼

#涨姿势# #打call# 科普那些你不了解的区域

**@中科院之声** 

【天文学家捕获的首张黑洞照片公布】该黑洞图像揭示了室女座星系团中超大质量星系 Messier 87中心的黑洞。该黑洞距离地球5500万光年，质量为太阳的65亿倍。图中心的暗弱区域即为“黑洞阴影”。



@中科院之声

2019年04月10日 21:07 来自 微博 weibo.com 转发 88896 | 评论 36236 | 点赞 207541

2019年04月10日 22:10 来自 Android

---

收藏 | 转发 | 评论 | 点赞 1

6.

 **法国蔚蓝海岸戛纳**  ▼

#法国#最美集市评选活动正在法国#TF1#电视台火热竞争中戛纳#福尔韦尔市场#喊你来为它#打call#哟 🤔🤔🤔🤔



2019年03月26日 14:27 来自 专业版微博

---

收藏 | 转发 2 | 评论 | 点赞 3

7.



空调忽冷忽热 🍷

#流浪地球# 观后感

我觉得作为一个爆米花电影，真的很棒了，背景故事交代清晰，画面真的很大片。不懂挑刺的观众的心态，比如后天，2012，环太平洋，星际穿越，普罗米修斯等等好莱坞大片不都是靠一个小队甚至一个人翻盘全局么。不管怎样，为中国科幻电影#打call#

2019年02月11日 19:40 来自 OPPO Find X

收藏

转发

评论



8.



任洁洁upup抬起头

为什么表示支持叫做#打CALL#

快速信息化的时代，见面>打电话>>在心里默默想念却不表达  
要要常约饭，常吐槽，增进友谊，身心愉悦#九月再见#



2018年09月29日 18:01 来自 iPhone客户端

收藏

转发

评论 6



9.

 米驴旅行网 ✓

震后初夏的#九寨沟# 经历了“8.8”地震的九寨沟景区迎来了震后初夏，经过近一年的自然恢复，长海、犀牛海、树正群海、芦苇海等景点光彩依旧。为九寨沟#打call# 🥰🥰



2018年05月21日 12:07 来自 360安全浏览器

收藏 | 转发 | 评论 | 1

10.

 西北大学社团管理中心 ✓

#幸好有你，为你打call# 疫情当头，是坚守工作岗位的医生，军警，科研小组，教师，在守护我们，坚守在一线。而我们能做的，就是为他们打call。快来挖掘疫情当头不广为人知的英雄事迹，通过话题#西大社管——为xxx 打call # 参与，不要忘记@西北大学社团管理中心。快来为你心目中的英雄加油助威吧! 🙌🙌🙌 收起全文^




03月13日 11:35 来自 OPPO K3 硬核少年

收藏 | 转发 | 评论 | 3





## Annexe 5 : Le corpus de tweets français

### 5a : Le terme n°1 *tuerie*

1.

 **Thibaut** @TPiccioni · 23 sept. 2010

C'est dingue. Le dernier vrai bon PES, c'était le 6. Et depuis, pfiouuu..  
FIFA a gagné du chemin chaque année, et maintenant  
[#CestUneTuerie](#)

2.

 **Liv** @acklesholic · 14 août 2018


J'ai fait une glace à la crème de coco, aux noix de pécan et au caramel  
de sucre de coco et comment vous dire 🤔🤔🤔 [#cestunetuerie](#)

 1   2 

3.

 **peke** @Martinoooo\_ · 22 févr. 2013


S'enjailler avec ses potes devant l'album de [@lafouine78](#)  
[#CestUneTuerie](#)

4.

 **Carole** @carole\_dh · 19 juin 2013

Je viens de regarder la vidéo que j'ai prise quand ils dansaient [#Abusé](#)  
[#Cestunetuerie](#)

5.

 **Juliette Corroyez** @JCorroyez · 20 avr. 2013 ▼  
Il y a un niveau de ouf cette année a The Voice !! [#cestunetuerie!](#)

6.

 **Quentin**. @QuentinLPe · 27 sept. 2012 ▼  
Je viens de découvrir la photo panoramique sur Iphone5  
[#cestunetuerie](#)


   





7.

 **monsieur crapule** @davidhertz0 · 19 sept. 2012 ▼  
Je suis content que le dessin que j'ai fais pour le [#tattoo](#) de mon gars @geezy\_officiel redonne aussi bien [#CestUneTuerie](#)

8.

 **Made in Vietnam** @AlineDesoubrie · 14 oct. 2014 ▼  
J'ai dû être dans un autre monde, je viens de découvrir @QueensChristine [#CestUneTuerie](#) !!!


   

9.

 **Sadoun** @MichaelSadoun · 7 août 2012 ▼  
les olives en Grèce [#CestUneTuerie](#)

10.

 **Dami** 🤖 @DamienOvril · 10 mai 2013 ▼  
Claudia quand elle est énervé [#CestUneTuerie](#)

## 5b : Le terme n°2 *marche blanche*

1.

 **Christophe** @ChristophePretr · 22 mai

Quelqu'un peut m'expliquer ce nouveau concept de **#MarcheBlanche** en soutien à la délinquance ? **#Argenteuil** ...

19 62 214

2.

 **Charlotte Baechler** @Chach\_Bchlr · 17 nov. 2019

Près de 300 personnes sont venues dire aurevoir à Sylvia, 40ans, morte sous les coups de son mari.  
Rdv à 12h30 sur @Europe1 **#femicide** **#Oberhoffensurmoder**  
**#Alsace** **#marcheblanche**



6 7

3.

 **Philippe** @psikonawak · 16 déc. 2011

Liège: marche blanche controversée samedi et hommage officiel mardi  
[bit.ly/tHvyfJ](http://bit.ly/tHvyfJ) @ladh **#liège** **#marcheblanche**

6 7




4.

 **L'Est Républicain**  @lestrepublicain · 13 janv. 2012   
#MeurtredeMaxime / #marcheblanche demain à #Etouvans ( 25 ) -  
L'Est Républicain - [bit.ly/A1H6j2](http://bit.ly/A1H6j2)

5.

 **Toulouseinfos** @Toulouseinfos · 17 mars 2013   
#Toulouse : 1000 personnes rendent hommage aux victimes de  
#Merah [ow.ly/j55bb](http://ow.ly/j55bb) #marcheblanche



   





6.

 **Anas Oulmidi** @Oulmidianas · 4 mai 2013   
N'oubliez pas la #marcheblanche demain a 10h devant le megarama,  
je fé le déplacement de marrakech donc casawi, soyez la. #pedophilie  
#maroc



  1  





7.

 **Aïda Amor** @Aida\_Amor · 1 mai 2013   
Tous les crimes sexuels devraient être criminalisés, c'est une honte que  
ce ne soit pas déjà le cas! #LT #justicemaroc #marcheblanche

8.

 **Grégoire de FOURNAS** @gdefournas · 5 avr. 2014   
Soutient à la famille d'Alexis Moulinier, étudiant de 19 ans lâchement  
assassiné d'un coup de couteau à Talence. #marcheblanche  
#insecurité

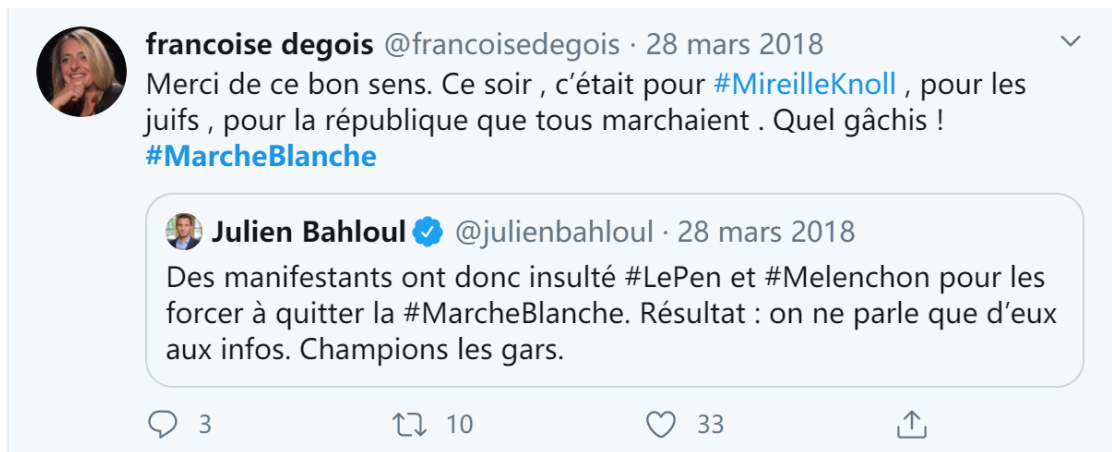
  1  1 

9.

 **Manuel Valls**  @manuelvalls · 16 juin 2016   
Pensée émue pour nos deux policiers lâchement assassinés par la  
barbarie. #MarcheBlanche

 108  102  180 

10.



**francoise degois** @francoisedegois · 28 mars 2018

Merci de ce bon sens. Ce soir , c'était pour [#MireilleKnoll](#) , pour les juifs , pour la république que tous marchaient . Quel gâchis !  
[#MarcheBlanche](#)

**Julien Bahloul** @julienbahloul · 28 mars 2018

Des manifestants ont donc insulté [#LePen](#) et [#Melenchon](#) pour les forcer à quitter la [#MarcheBlanche](#). Résultat : on ne parle que d'eux aux infos. Champions les gars.

3 10 33

### 5c : Le terme n°3 *nomophobe*

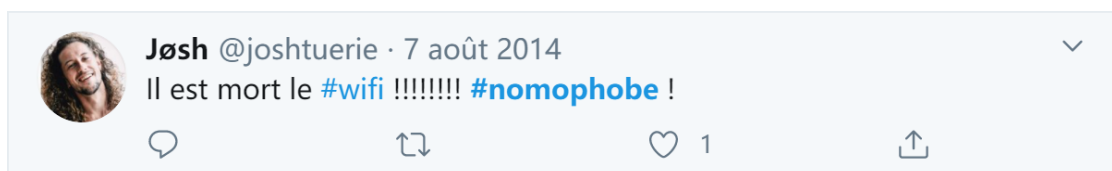
1.



2.



3.



4.

 **abbé Dominique Fabien Rimaz** @DonDomFabien · 1 févr. 2018

Le terme **#nomophobe** est un mot-valise construit par contraction de l'expression anglaise « no mobile-phone phobia » et désigne alors la peur excessive d'être séparé de son téléphone mobile



**NOMOPHOBE** (n. m.) :  
PERSONNE AYANT  
PEUR D'ÊTRE SÉPARÉE  
DE SON SMARTPHONE

1 1 1

5.

 **Del\_H** @MatthieuDelach · 20 févr. 2012

Etes-vous **#nomophobe** (souffrir de ne pas avoir son téléphone portable à portée de main) ? (Slate) [slate.fr/lien/50349/SAN...](http://slate.fr/lien/50349/SAN...)

1 1 1

6.

 **Mickael Pallu** @mikapallu · 23 déc. 2015

Je suis **#nomophobe** ! Pourquoi est-on accro aux notifications sur les réseaux sociaux ? [madame.lefigaro.fr/bien-etre/doi...](http://madame.lefigaro.fr/bien-etre/doi...) via @MadameFigaro

1 1 1

7.



**T'as tout compris** @tastoutcompris · 24 mars 2016

Accros à votre [#smartphone](#) ? Vous êtes peut-être [#nomophobe](#) alors ! Dans [#TasToutCompris](#) on fait le point ;-)  
[education.francetv.fr/emissions/t-as...](http://education.francetv.fr/emissions/t-as...)

🗨️      ↻ 9      ❤️ 1      ↗

8.



**Flo** @cueillesfleurs · 3 août 2016

Bon, 36 heures sans Wi-Fi je tiens plus... Tant pis, je mets vite fait la 4G 😊 [#nomophobe](#) [#accroàTwitter](#)

🗨️      ↻ 2      ❤️ 4      ↗

9.




**Philippe Etienne** @dudijonnais · 6 févr. 2015

C'est aujourd'hui la journée sans portable...OK je dégainé ma tablette 8" [#nomophobe](#)

🗨️      ↻      ❤️ 2      ↗

10.



**Elodie de Sélys** ✓ @edselys · 18 oct. 2016

Oublier mon téléphone à la maison. Impression d'être à l'école en pantoufles. [#cauchemar](#) [#nomophobe](#) #çavaaller#çavaaller#çavaaller [#çavapas](#)

🗨️ 11      ↻ 1      ❤️ 47      ↗

## 5d : Le terme n°4 *zadiste*

1.

 **Chantal morel #PortezLeMasque** @dc\_morel · 16 avr. 2018

Démocratie "alternative" version **#zadiste** pour les bloqueurs des **#universités** 😞  
**#LoiOre #LoiVidal**

 **FAGE** @La\_FAGE · 16 avr. 2018

"En Lorraine, on a eu 50% de participation et 75% de personnes pour la reprise des cours, les bloqueurs n'ont pas reconnu le vote" @Theo\_FAGE sur @franceculture

🗨️ 11 🍷 3 📤

2.

 **émilie Gasc** @emiliegasc · 21 sept.

Je découvre que **#Zadiste** est dans Le Petit Robert : « Militant qui occupe une ZAD pour s'opposer à un projet d'aménagement qui porterait préjudice à l'environnement ». Y'a aussi **#ZAD** pour « Zone à défendre », cool pour le Scrabble !

🗨️ 1 🔄 1 🍷 4 📤

3.

 **Cath** @Cath\_\_ · 2 août 2013

le mot **#zadiste** va-t-il rentrer dans le dico ? Des mouvements partout à l'image de **#nddl #cestbonpourlemoral** [la-croix.com/Actualite/Fran...](http://la-croix.com/Actualite/Fran...)

🗨️ 🔄 🍷 📤

4.

 Lecteur CharlieHebdo @CHVPSL · 10 avr. 2018  
Ah ah ! #CharlieHebdo a débusqué "Le zadiste de l'Elysée" pour la Une du 11 avril !  
#Macron #Zadiste #JeanPierrePernaut

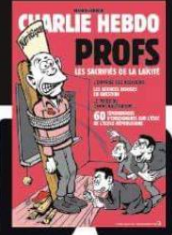


LAÏCITÉ À L'ÉCOLE  
LES PROFS TÉMOIGNENT  
NOUVEAU HORS-SÉRIE EN KIOSQUE  
CHARLIE HEBDO  
LES SECRETS DE LA LAÏCITÉ  
LE ZADISTE DE L'ÉLYSÉE

1 48 73

LAÏCITÉ À L'ÉCOLE  
LES PROFS TÉMOIGNENT

NOUVEAU  
HORS-SÉRIE  
EN KIOSQUE



# CHARLIE HEBDO



5.

 **RAVEN φ** @ltvan · 28 nov. 2014

J'étais pas contre, mon fils voulait être policier, maintenant il veut devenir **#Zadiste** et j'en suis fier ! **#RemiFraise** **#CazeneuveDémission**

  1  1 



6.

 **Patache** @pjover44 · 22 déc. 2017

On change le terme, **#zadiste** devient néo-ruraux, pour ne pas pas parler des vrais problèmes . DE LA COM bien orchestrée . Moi, je renomme les "médiateurs" en "bidouilleurs" ! ! !

 **GF\_Courage** @GF\_Courage · 21 déc. 2017

Quel terme édulcoré, néo-ruraux, pour désigner les zadistes qui occupent illégalement des terrains! [twitter.com/Chris\\_\\_444/st...](https://twitter.com/Chris__444/st...)

  8  10 

7.

 **Laurent Potelle** 🐻 @Laurent\_Potelle · 24 avr. 2018

♪ J'aurais voulu être un **#Zadiste**  
Être l'ami des oiseaux  
Et pas un crétin sur la piste  
du macadam, des lacrymos ♪

 1  16  47 

8.

 **La Dépêche 31** ✓ @ladepeche31 · 23 janv. 2015

**#zadiste** de **#Sivens** blessée à la main : ouverture d'une information judiciaire [depe.ch/6wAQ](https://depe.ch/6wAQ) **#Tarn**



9.



**Camille** @000Vince · 21 oct. 2016

Pour votre soif d'[#Hemoglobine](#), mangez un [#Tartare](#) ([#Vegan](#) biensûr) pas un [#Zadiste](#)...!  
[#NDDL](#)



1



2



10.



**Bruno Dalpra** @brunodalpra67 · 7 janv. 2016

[#JeSuisCharlie](#) : être solidaire n'a pas la même signification selon si tu t'appelles [#Valls](#) ou si tu es [#zadiste](#)... [#Convergence](#) [#NDDL](#) [#GCO](#)




1



## 5e : Le terme n°5 *tendu comme un string*


1.

 **Toma de saulieu** @Toma\_de\_Saulieu · 1 avr. 2015

Rdv dans 15mn avec les mécènes de ma future exposition pour leur présenter ma sélection .....  
[#Tenducommeunstring](#)

1      ↻      3      ↗

2.

 **Eva Samaï** @samaisamva · 25 avr. 2013

Si vous saviez comment l'épisode de Samaï Chasseuse de Crafts d'aujourd'hui est trop EPIC ! [#TenduCommeUnString](#) sortie 17h30

1      ↻ 1           ↗

3.

 **Samy Rabih** @SamyRabih · 23 juil. 2014

De ces jours où je ne supporte plus du tout les pertes de temps et les discussions qui ne vont pas à l'essentiel [#tenducommeunstring](#)

     ↻ 1           ↗


4.

 **Nico Puchat** @NicoPuchat · 27 mars 2014


Pas possible de jouer ces matches de playoffs plus tôt dans la journée ? C'est interminable... [#HeyOnYVa](#) [#TenduCommeUnString](#)

1      ↻ 2      4      ↗


5.

 **arnaudt** @arnaudt · 29 sept. 2010 ⌵





Attends désormais deux choses d'Apple : iOS4.2 pour iPad, et le MBA 11'6.  
Je m'impatiente !  
[#tenducommeunstring](#)

6.

 **Nico** @nleal12 · 7 nov. 2012 ⌵

Je m'intéresse plus aux élections américaines que celles de France,  
[#passionnant](#) [#TenduCommeUnString](#) [#obama2012](#)


7.

 **arnaud hamzaoui** @a02halo · 1 avr. 2014 ⌵


4 semaines pour une CNI et 3 semaines pour un passeport..je dois être  
a Birmingham le 27 avril.. [#tenducommeunstring](#)

 1   

8.

 **MysternoW** @Mystern0W · 27 nov. 2014 ⌵

Contacté par MAGIX qui a créé le logiciel de musique que j'utilise: ils  
veulent ma participation à un tournoi musical :D  
[#TENDUCOMMEUNSTRING](#)

 1   

9.

 **...CarpeDiem...** @kevrama · 13 nov. 2012 ⌵

Salut les [#Twittos](#) dites moi quelqu'un sait il comment deconnecter  
un compte twitter sur l'application iphone ? [#tenducommeunstring](#)

 1   

10.

 **cpolitic** @blogcpolitic · 17 nov. 2016  
Impressionnant la joie sur le visage des astronautes...  
[#tenducommeunstring](#)

 **Del\_H**  @MatthieuDelach · 17 nov. 2016  
EN DIRECT - Suivez avec nous le décollage de Thomas Pesquet à bord de la fusée Soyuz qui rejoindra l'ISS  
[Lci.fr/high-tech/en-d...](http://Lci.fr/high-tech/en-d...)



   1 

5f : Le terme n°6 *vivre ensemble*

1.



**CZeLibérArsèn** #FJW #ObjectifFrance #FBPE @... · 13 déc. 2019

Les miracles du **#VivreEnsemble** : un Israélien violemment agressé dans le métro à **#Paris** parce qu'il parlait hébreu



Paris : un Israélien violemment agressé parce qu'il parlait hébreu  
Un étudiant israélien de 31 ans a été victime d'une agression dans le métro parisien. Le parquet de Paris a retenu le caractère ...  
[rtl.fr](https://www.rtl.fr)



4



16



19



2.



**France Jamet** @FranceJamet · 23 oct. 2017

Nouvelle provocation dans le monde du **#VivreEnsemble** : ils appellent leur petit garçon **#Jihad**.



Ils appellent leur petit garçon... Jihad  
Le choix du prénom de leur petit garçon ne fait pas l'unanimité. En août, deux parents vivant à ...  
[ledauphine.com](https://www.ledauphine.com)



14



6



3.

 **Robert Ménard**  @RobertMenardFR · 9 déc. 2016 ▼  
Le [#VivreEnsemble](#) est la plus grande escroquerie du monde ...  
[#islam](#) [#Immigration](#)

 52       205       217      

4.

 **CFDT**  @CFDT · 8 déc. 2015 ▼  
Comment traiter la question du fait religieux en entreprise? Des éléments de réponse dans notre guide [#VivreEnsemble](#)





Le fait religieux en entreprise - édition 2018  
La collection « Vivre ensemble, travailler ensemble » que la CFDT propose est au service des ...  
[cfdt.fr](#)

       25       5      

5.

 **François Mandil** @francoismandil · 29 déc. 2015 ▼  
Plus je lis les infos, plus l'éduc pop m'apparaît comme une des principales réponses de long terme à la crise de société  
[#VivreEnsemble](#)

 2       10       14      

6.

 **Christophe Clergeau**  @clergeau · 9 déc. 2015 ▼  
110 ans de la [#loi1905](#) sur la [#laïcité](#), une loi de tolérance, de liberté et de [#VivreEnsemble](#)

       14       3      

7.

 **Allocations Familiales**  @cnaf\_actus · 23 nov. 2015 ▼  
Demain, @centressociaux et la Cnaf présenteront [#Senacs](#), Observatoire des centres sociaux [#vivreensemble](#) [#cohésionsociale](#)

       6       2      

8.

 **Fabrice Robert**  @\_fabricerobert · 21 déc. 2014 

Un djihadiste "français" qui appelle les musulmans à tuer des Français.  
Le miracle du [#VivreEnsemble](#) ? [goo.gl/ToQJyk](http://goo.gl/ToQJyk)

 3  36  13 

9.

 **ladilafé** @LadiLafeFR · 26 août 2014 

L'[#HistoireDeLaRéunion](#) ne se résume pas à des cicatrices  
Elle ns a aussi appris à [#VivreEnsemble](#) en Frères [#WeAre974](#)



 1  17  10 

10.

 **SniperDeDroite**  @SniperDeDroite · 20 août 2013 

Les [#socialistes](#) qui parlent de [#sécurité](#) c un peu comme si Marine  
[#LePen](#) parlait du [#VivreEnsemble](#) [#Crédibilité0](#)

  4  



## 5g : Le terme n°7 *loup solitaire*

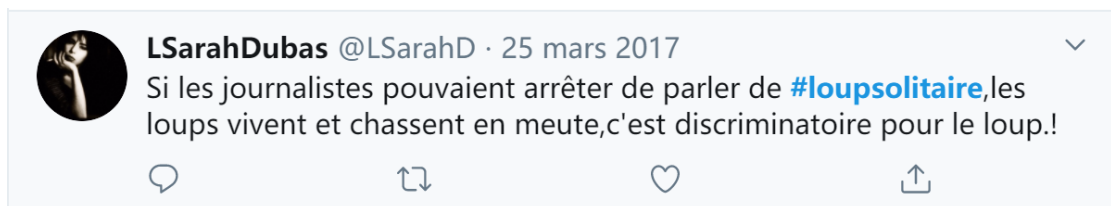
1.



2.



3.



4.

 **Del Ponte João** @PonteJoa · 12 déc. 2017

Y a pas de [#LoupSolitaire](#) Abattez le troupeau ! Merci [#CharlieHebdo](#) pour cette couverture!

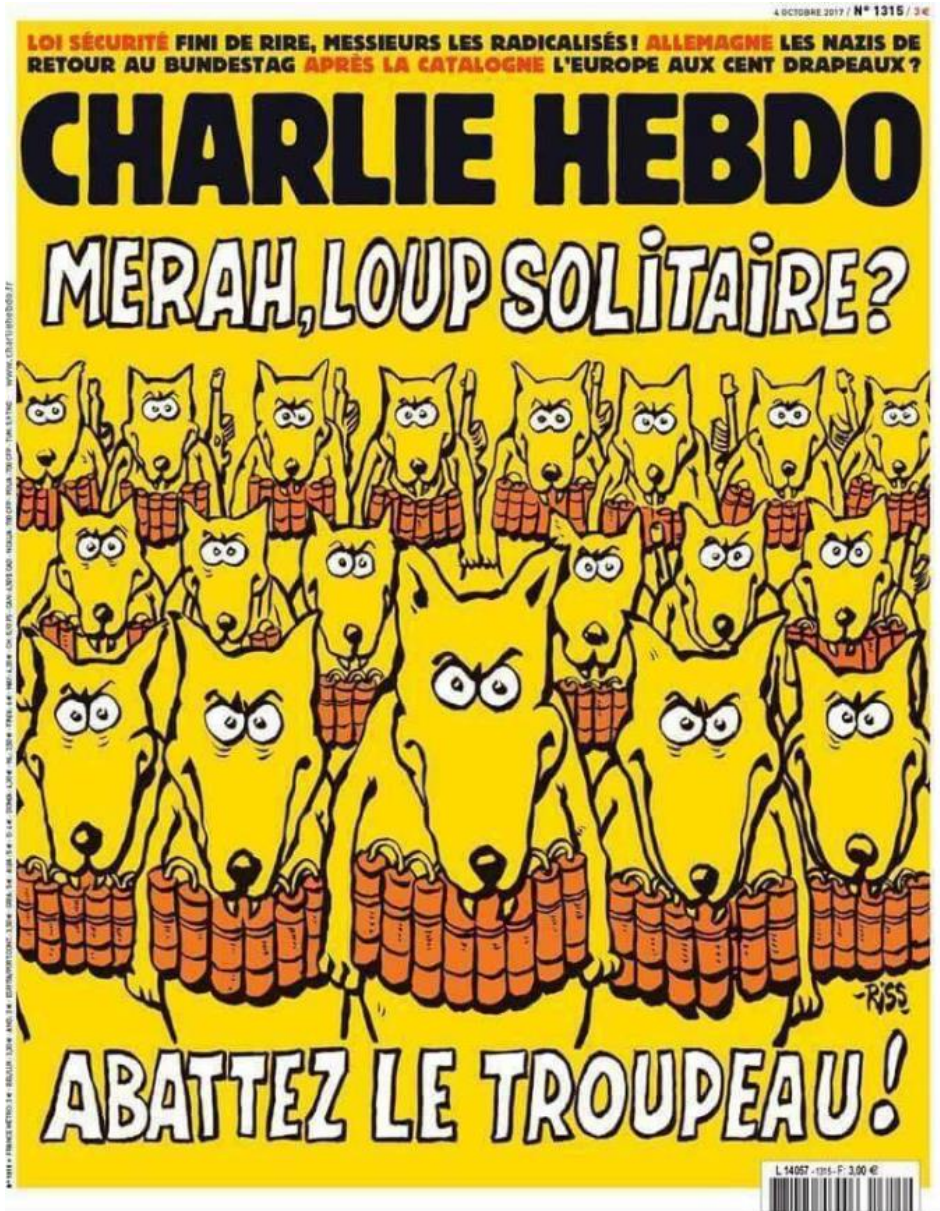


LES BARRIÈRES FINI DE CIRE, MESURENT LES BARRIÈRES? ALLONGER LES MARS DE  
BETON AU BARRIÈRE JOUR LA L'ÉTRANGER L'ÉTRANGER AUX TOUT DÉPENSER ?

**CHARLIE HEBDO**

**MERAH, LOUP SOLITAIRE?**



5.



**Hiitzy Sanchez** @HiitzyPharaon · 30 sept. 2017

Je ne me soucis plus des autres. J'suis devenu trop égoïste je pense.

[#LoupSolitaire](#)



6.



**La Médiatrice** @mediatriceRF · 25 janv. 2018

Qu'est ce qu'un [#LoupSolitaire](#) ? Pierre Manen, linguiste, explique le sens du mot. [mediateur.radiofrance.fr/videos/loup-so...](http://mediateur.radiofrance.fr/videos/loup-so...)




7.

 **Jocelyne Robert** @JocelyneRobert · 30 juil. 2016  
Nommer [#chiensolitaire](#) et non [#loupsolitaire](#) le terroriste esseulé.  
Loup revêt noblesse dans imagerie populaire. Pas le chien  
[@dessinemoi](#) [#El](#)

1 7 8

8.

 **Alfred Maxar** @alfredmaxar · 21 juil. 2016  
Le loup chasse toujours en meute.  
Le [#LoupSolitaire](#) n'existe pas, c'est un mythe.


7

9.

 **Sylvain Couture** @blackshield77 · 30 nov. 2016  
Livre dont vous êtes le Héros : R.I.P Joe Dever [#lonewolf](#)  
[#loupsolitaire](#) [#jdr](#) [#rpg](#)

7

10.

 **robert grossmann** @robgrossmann · 3 juin 2014  
[#attentats](#) [#Loupsolitaire](#) une invention des journalistes. Comment  
baptiser ainsi de lâches assassins ? Pourquoi insulter le loup?

1

## Annexe 6 : Métadonnées des corpus

### 6a : Captures d'écran des billets de blog du magazine «yǎo wén jiào zì» (2009-2017)

2009

荐 2009年十大流行语 (2010-01-21 09:40:00) + 转载 ▾

标签: 杂谈

## 2009年十大流行语

《咬文嚼字》编辑部

说明

2009年十大流行语是由本刊广大读者推荐，经语言文字专家多次讨论评议后共同选定的。以下三种情况未列入流行语范围：第一，一般的新闻词语，像指人的“周立波”，指物的“曹操墓”，指事的“甲流”等。第二，只出现在网络中，尚未被纸质媒体认可并广泛使用的词语，像以谐音方式替代原词的网络用词“杯具”（悲剧）、“洗具”（喜剧）、“餐具”（惨剧）、“茶具”（差距）等。第三，仿造过去的流行格式而产生的次生词语，像“楼脆脆”“桥塞塞”等。

一、不差钱 本是东北地区的方言说法。2009年“春晚”赵本山等演出的小品，用它作了剧名，产生了轰动效应，后在全国流行开来。在实际使用中，多带有调侃的色彩，并不纠缠于钱多钱少。

二、躲猫猫 本是一种儿童游戏，即“捉迷藏”。2009年2月，云南省晋宁县看守所里发生了所谓“躲猫猫”事件，一名24岁的男子因盗伐林木被关入看守所后死去，警方称死者是在玩“躲猫猫”时眼部被蒙，不慎撞墙受伤而死亡的。最终侦查结果表明该男子系被“狱霸”殴打致死。此后，“躲猫猫”就有了多种新义：或隐瞒事实，或逃避监督，或暗箱操作。总之，不让人了解真情。

**三、低碳** 所谓“低碳”，是指降低二氧化碳的排放。二氧化碳是全球气候变暖的罪魁祸首，影响到了人类社会的可持续发展。2009年12月哥本哈根联合国气候变化大会，在世界范围内又一次掀起了“低碳”浪潮。“低碳”经济、“低碳”科技、“低碳”城市、“低碳”生活，“低碳”已成为“绿色”的重要标志之一。

**四、被就业** 2009年7月，一位应届大学毕业生在网上发帖爆料：他在不知情的情况下，学校已经替他签好了“就业协议书”。于是网友发明了“被就业”的说法，以讽刺这种虚报高校就业率的行为。随后，但凡意志遭遇强迫时，均可套上“被××”的帽子，如职工“被全勤”，举报人“被自杀”，交择校费的家长“被自愿”。

**五、裸** 本义为不穿衣服，流行语用的是引申义，即只有事物本身，而没有任何附加物或附加条件。比如“裸婚”是不买房，不买车，不戴婚戒，不办婚礼，不度蜜月，只领取结婚证书；“裸官”是家属孩子存款都在国外，一个人留在国内做官；“裸退”指干部退休后不再担任官方、半官方以及群众组织中的任何职务。

**六、纠结** 本为动词，表示互相缠绕。自从四五年前的一部动画片中有个角色大呼“纠结啊”之后，“纠结”便在网络上走红。2009年起，更广泛见于纸质媒体，并且用法多样。可以作名词，表示解不开的心结；可以作动词，表示陷入复杂而尴尬的境地；可以作形容词，表示思绪的极度困惑和混乱。

**七、钓鱼** 完整的说法叫“钓鱼执法”。这是一种比喻的说法，指执法人通过“钓钩”，诱惑一个原本没有违法意图的人去从事违法活动，然后实施执法。这种说法早已存在，因2009年10月上海发生的一起查“黑车”事件而重新引起关注并流行。凡是违反法律精神，别有用心地诱人上钩，都可以称之为“钓鱼”。

**八、秒杀** 译自英语的seckill，起先是网络游戏的专用词，指玩家在游戏中瞬间被PK出局或者瞬间将对手击倒。2009年9月，某购物网站周年庆推出了“秒杀”活动：在网络拍卖开始后第一个确认的网络买家，可以按照远远低于成本价的秒杀价买到指定商品。于是“秒杀客”“秒杀族”应运而生，并且越来越多。此后，“秒杀”又用于其他领域，比如股市中某股价格在短时间内大幅下跌，也叫“秒杀”。

**九、蜗居** 本来用于谦称自己的住所，指像蜗牛壳一样狭小的房子。这一词语的流行，和2009年热播电视剧《蜗居》有关。在房价节节攀升的大背景下，电视剧中年轻人为了买房而沦为房奴的故事引起了广泛的共鸣，它让人们对于“蜗居”一词有了深刻的印象。“蜗居”不仅可指空间上的狭窄，也可指精神上的狭窄。

**十、蚁族** 青年学者廉思主编的《蚁族——大学毕业生聚居村实录》一书出版后，“蚁族”一词就频繁地在各种媒体上亮相。它指的是高学历低收入的群体。这些人一般聚居大城市的边缘地区和近郊农村，因和蚂蚁有若干相似之处而得名。他们虽然弱小，但胸怀理想，充满活力，具有挑战的意识和顽强的意志。

标签： 杂谈

## 2010年十大流行语

《咬文嚼字》编辑部

### 说明

2010年十大流行语是由本刊广大读者推荐，经语言文字专家多次讨论评议后共同选定的。以下三种情况未列入评选范围：第一，一般的新闻词语，比如“世博会”“房产税”“富士康”等。第二，指某种文体的，比如因其作者获鲁迅文学奖而引起热议的“羊羔体”，“爱××，不爱××”的“凡客体”等。第三，按一种模式形成的临时组合，如涨价系列中的“蒜你狠”“豆你玩”“姜你军”“糖高宗”“油你涨”“苹什么”等。

**一、给力** 字面意思是“给予力量”，引申指“酷”“棒”“爽”。它出自网友配音的日本动画片《西游记·旅程的终点》——唐僧师徒历经磨难到达天竺，发现那里只有一面写着“天竺”二字的小旗，孙悟空抱怨道：“这就是天竺吗？不给力啊！”该动画受到网友追捧，“给力”也由此流行。2010年11月10日，“给力”登上《人民日报》头版头条，这被视为权威媒体认可网络词语的标志性事件，美国《纽约时报》也因此发文介绍“给力”。

**二、神马都是浮云** “神马”非马，而是“什么”的谐音；“浮云”则指转瞬即逝的事物。两者结合，意思就是“什么都不值得一提”。它的流行，源于2010年国庆期间的网络“小月月”事件。横空出世的“小月月”，以匪夷所思的言行雷倒众生，引得网友感叹：“神马网络豪放女，浮云！都是浮云！”流行语“神马都是浮云”由此产生，它可用于抱怨，可用于感叹，还可用于表达超然的心态，成为许多人的口头禅。

**三、围脖** “织围脖”是当今时尚。其中的“围脖”，指的是“微博”，是“微型博客”的简称。“织围脖”即“写微博”。与博客相比，“微博”使用方法更简便，信息传播更迅捷，因此深受用户青睐。如今互联网已迅速进入“围脖时代”，不仅名人热衷于“织围脖”，普通人也乐于成为“织男”“织女”。

**四、围观** 无事围观，曾被鲁迅斥为国人陋习。但如今流行的“围观”义为关注，不再含有贬义。《南方周末》曾发表文章《关注就是力量，围观改变中国》，认为公民的围观，民意的关注，可以“让良知默默地、和平地、渐进地起作用”，最终会促进公权力的合理使用。广州亚运会期间，从抵制番禺垃圾焚烧，到呼吁改变地铁惠民方式，自发的公民“围观”引人注目。

**五、二代** “代”有世系相传的辈分的意思，“二代”即第二代。如今社会，“二代”丛生：官二代、富二代、穷二代、文二代（作家子女）、星二代（明星子女）、农二代（农民工子女）、独二代（第二代独生子女）……，甚至姚明的孩子也被戏称为“姚二代”。“二代”如此流行，是与社会结构的板结化不无关系的。

**六、拼爹** “拼爹”的“拼”是比拼。短兵相接要“拼刺刀”，扳手腕要“拼力气”，马拉松要“拼耐力”。“拼爹”则拼的是“爹”的地位和实力。“我爸是李刚”“我叔是金国友”，都是典型的“拼爹”语言。这类词语的流行，反映的是大众对社会不良现象的不满情绪。

**七、控** 如今，超喜欢长头发的人叫“长发控”，半夜醒来还忍不住要去发一条微博的叫“微博控”，拿着世博护照遍场跑馆盖章的人叫“敲章控”，疯狂参加圣诞活动的叫“圣诞控”。这些“控”，是指极度喜欢某种事物的人。“某某控”的结构是从日语借来的，而日语中的这个“控”则源于英文单词complex（情结）。日语中“控”的读音与complex开头的音相近，日本人就用“控”来表示具有某种情结的人。用“控”对译complex（情结），既有读音的对应，又有意义的结合——深爱某个对象，不就在相当程度上被控制了吗？

**八、帝** 帝，本义是封建社会的最高统治者，如中国的“康熙皇帝”，外国的“彼得大帝”。流行语中指某些领域中成就大、造诣高的人，如“影帝”。如今，只要拥有某一特点即可称“帝”。2010年，各种人物频频称“帝”，如“数钱帝”“表情帝”“体操帝”“贺岁帝”等，而最有名的“帝”却不是人而是鱼——预测南非世界杯足球赛的德国章鱼保罗，它预测八场，场场皆中，被称为“章鱼帝”“保罗帝”“预测帝”。

**九、达人** 中国古代就有“达人”一词，指通达事理的人。“达人”被借入日语后含义发生了变化，如今又从日语传了回来，指见多识广、对某方面懂得透彻的人，如“网络达人”“理财达人”“社交达人”等。2010年东方卫视举办的选秀节目“中国达人秀”，引得万众瞩目，“达人”一词更是遍地开花。

**十、穿越** 传统的“穿越”，对象是空间，如“穿越边境”“穿越沙漠”；现在流行的“穿越”，则是“穿越时空”的简称，是文艺作品中一种展开情节的流行手段。穿越电影、穿越电视剧、穿越小说等共同的特点，是人物会不断地往来于不同的时空。这种时空的错乱，容易造成受众的费解或误解，于是，“穿越”又有了“玄”“乱”“令人莫名其妙”等意思。



标签：文化

## 2011年十大流行语 《咬文嚼字》编辑部 (2011年12月20日)

### 说 明

在广大读者热情推荐的基础上，经过汉语专家反复评议、甄选，本刊拟定了2011年十大流行语。评选仍然坚持几年来一以贯之的原则：第一，不问“出生”在何年何月，只看它在本年度的流行频度。第二，不以网络流行为唯一依据，更要看传统媒体上的影响。第三，注重词形结构的稳定性，即兴式的谐音词一般不选，如“河蟹”（和谐）、“童鞋”（同学）。

**一、亲** “亲”是“亲爱的”的简称。“亲爱的”是一百多年前为了翻译西方语言而“新”造的，用作名词和形容词，大体上跟英语的dear和darling相对应。几年前，“亲”曾在某些群体的小范围中露面；随后进入淘宝网的交易平台：“亲，快来抢购哦！”“亲，包邮哦！”于是“亲，×××”风行起来，人们称之为“淘宝体”。跟“亲爱的”相比较，“亲”显得简洁，也屏蔽了“爱”字的暧昧色彩，亲切感却有增无减。

**二、伤不起** 最初在校内网上以文章标题的形式出现，如“这样的×××你伤不起”，被称为“校内体”，影响并不大。2011年“伤不起”火爆起来，和网络上一篇题为“学法语的人你伤不起啊！”的帖子有关。这个帖子历数学习法语所遇到的种种困难，种种无奈，并且几乎每句话都以“啊”结尾，后面还有一个接一个的感叹号，人称“咆哮体”。随着该帖大热，“伤不起”成了热词，“×××，你伤不起啊”也成了热门句式。

**三、Hold住** 2011年8月9日台湾中天电视综合台的一档综艺节目中，女大学生谢依霖以夸张另类的造型、英汉混杂的台词、扭捏怪诞的举止亮相，陈说在一个时尚场合如何处变不惊，提醒自己不能慌乱，要“hold住”整个场面。7分钟的节目引起巨大的反响，于是“hold住”迅即成为海峡两岸的流行语。Hold指掌控、把握、维持，“hold住”便是掌控住、保持住、管住、抓住之类的意思。

**四、我反正信了** 2011年7月23日，在温州附近发生一起两列动车追尾的重特大交通事故，40人遇难，191人受伤。抢救中有关人员曾把一辆车的车头掩埋于地下。有媒体记者询问原因，铁道部新闻发言人回答：现场有一个泥塘，影响施救工作，把车头埋于其中是为了便于抢救，“至于你信不信，（停顿）我反正信了”。他的话立即引来一片质疑。此后，“我反正信了”被广泛使用，其含义有三：第一，真信，如铁道部那位新闻发言人所说的；第二，说反话，真不信，如“房价是百姓可以接受的，我反正信了”；第三，无厘头的搞笑，指自己说的压根就是一个笑话，如“奥巴马是中国人，我反正信了”。

**五、坑爹** “坑爹”原是詈语，即骂人的话。在北方某些方言中，“坑”有欺骗、欺诈的意思，“爹”指老子，即父亲。“坑爹”的字面义便是“欺骗老子”。连老子都敢欺骗，胆子也忒大了。“坑爹”表达的是一种强烈的愤慨。在流传中，“坑爹”还常用于责备、批评或讽刺、挖苦。如果说话人是女性，有时会把“坑爹”调整为“坑娘”。

**六、卖萌** “卖萌”的“萌”，是从日语中借过来的。日本的动漫爱好者用“萌”形容非常喜好的事物，特别是动漫中的美少女。“萌”进入汉语以后，有了可爱、性感、讨人喜欢的新义项。可以说“萌少女”“萌女郎”等。“卖萌”的“卖”，不是出卖，而是显摆、展露。“卖萌”的意思是装可爱、扮嫩、撒娇，一般具有调侃色彩。

**七、吐槽** 来源于日语，指漫才（类似于中国的相声）里的“突っ込み”（类似相声的捧哏）。台湾把“突っ込み”译成了“吐槽”，后引申指给人难堪、抬杠、掀老底、拆台，多用于嘲笑、讥讽、抱怨，甚至谩骂。当前，主要有两种用法：一是揭人家老底——批评别人；二是揭自己老底——表述心声。

**八、气场** “气场”本指环绕在人体周围的能量场，能显示出一个人的整体状态，包括健康、心理及修为等等。美国心灵励志大师皮克·菲尔著有一本谈比尔·盖茨、奥巴马等世界名人超凡魅力的书，此书引进中国后书名被译成了“气场”。“气场”于是流行开来，指由气质、学识、修养等等的综合表现而形成的超凡魅力。具有强大“气场”的人必定富有吸引力和影响力。本年度的新闻人物报道，流行拿“气场”说事。

**九、悲催** “悲催”是个形容词，形容失意、伤心、难过、哀痛、丧失信心等，由短语“悲惨得催人泪下”缩略而成。从短语的字面上看，似乎“悲”的程度很高；其实不完全如此，常常带有或多或少的调侃或宣泄的意味。可以说“很悲催”“悲催的是”“悲催啊悲催”“悲催的一天”等。活跃在网络上，有三种写法，即悲催、悲摧、悲衰（“衰”读作cuī）。当下在报刊上流行，写法渐趋一致，都写作“悲催”。

**十、忐忑** 龚琳娜演唱的《忐忑》风靡一时。这首歌没有明确的意思，全是让人听不明白的“咳咿呀咿哟……”，但在演唱者眉飞色舞、幽默搞怪的夸张表演下，广大听众为之着迷。《忐忑》被人们戏称为“神曲”，从年过花甲的老太太到年仅几岁的小弟弟都竞相模仿着歌唱，连大牌歌星王菲也不例外。“忐忑”这个词于是火了起来，用法也得到了进一步丰富。本来是形容词，当下又演变出动词（如“忐忑了整整一天”）、名词（如“结束忐忑”）的用法来。有人说，这是“神曲”提振了“神词”。

**荐** 《咬文嚼字》发布“2012年十大流行语”，“正能量”位居榜首 (2012-12-30 18:09:50)

[+ 转载](#)

标签：文化

### 《咬文嚼字》编辑部

1) **正能量** 指的是一种健康乐观、积极向上的动力和情感。“正能量”本是物理学名词，霍金《时间简史》中就用过：“宇宙中的物质是由正能量组成的。”“正能量”的流行源于英国心理学家理查德·怀斯曼的专著《Rip it up》(中文译名《正能量》)，其中将人体比作一个能量场，通过激发内在潜能，可以使人表现出一个新的自我，从而更加自信、更加充满活力。当下，人们为所有积极的、健康的、催人奋进的、给人力量的、充满希望的人和事，贴上“正能量”标签。

2) **元芳，你怎么看？** 出自电视剧《神探狄仁杰》。狄仁杰遇到案情难解之处，时常会问副手李元芳：“元芳，你怎么看？”而李元芳的回答通常都是：“大人，此事蹊跷，背后隐藏着一个惊天阴谋。”今年的流行，是源于一起网络事件。泉州有一女孩疑似被肢解后坠落高楼，警方判断为自杀。一名网友以“元芳，你怎么看”进行嘲讽，暗指案情背后或有蹊跷。该句式于是迅速流行，人们多将它缀于某个句子或语段的末尾，表达某种质疑、嘲讽或公开征询看法。

3) **舌尖上** 2012年中央电视台播出纪录片《舌尖上的中国》，展示了中国各地的美食生态和丰富多彩的饮食文化，引起广泛关注。此纪录片的走红，使得“舌尖上”也受到了人们的青睐，俨然成了饮食文化的“代言人”，灵活地表达与之相关的丰富意蕴。如“舌尖上的中国”，即中国的饮食文化；“舌尖上的快乐”，意思是品味美食的乐趣；而“舌尖上的爱情”则可以理解为与饮食息息相关的爱情生活。

4) **躺着也中枪** 一般而言，枪战中如果躺在地上，中枪的概率是最低的。“躺着也中枪”于是用来形容无缘无故地受到牵连，或被卷进是非。出自周星驰电影《逃学威龙》中的一句台词。剧中双方激烈打斗，某人装死，另一人向地上发了一枪，正中装死的人，装死的人叫道“我靠！躺着都能中枪！”现在也有人使用缩略形式“躺中”“躺枪”，还有人干脆说“中枪”，表达的都是同样的意思。

5) **高富帅** 指高大、富有、帅气的男人，是由三个单音节形容词并列构成的准固定词组。“高富帅”和“白富美”(肤白、钱多、貌美)，是择偶方面最有优势的男人和女人。最初这两个词是网民对青春偶像剧和日本动漫中男女主人公特征的概括，表达了对虚拟世界里“高富帅”与“白富美”之间童话婚恋的艳羡，同时也寄托着对理想生活的向往和对平凡现实的自我解嘲。“高富帅”和“白富美”现已成为日常生活里的热门词语。

6) **中国式** “中国式”的说法最初源自王海鸰的小说《中国式离婚》，改编成电视剧后曾火热一时，衍伸出“中国式XX”等说法。而今年，网友提出的“中国式过马路”，即“凑够一撮人就可以走了，和红绿灯无关”，成为“中国式”爆发的导火线。媒体上出现大批以戏谑口吻议论国人或中国社会问题的“中国式XX”，如“中国式接送孩子”“中国式相亲”“中国式插队”等等，于戏谑中表达自嘲与质疑。

7) **压力山大** 意思是压力像山一样大。这个词是由我国男女老少最为熟悉的外国人名“亚历山大”，通过谐音、暗喻演变过来的。前半部分谐音，后半部分暗喻。后来又出现变体“鸭梨山大”，更加诙谐有趣。“压力山大”一问世，就被人们广泛接受，一来是交际上的需要，二来是念着顺口、听着顺耳，并且幽默感十足。

8) **赞** 先在网络上流行，然后打入传统媒体，有动词、形容词两种用法。作动词的“赞”可以替代许多双音词：赞美、赞赏、赞叹、赞同、赞许，以及称赞、夸赞等。作为形容词的“赞”表示好，一说来源于台湾，理由是台湾有一家报纸评选的2011年台湾年度汉字就是“赞”字，一说吴方言区(包括江苏南部、浙江大部和上海)一向用“赞”表示好，如上海话里“蛮赞”便是很好的意思。

9) **最美** “最美丽”的简称，新兴的用法“最美+身份”的组合作为一种对人物的美称在2012年已成为感动中国的新力量。这种组合的流行始于“最美妈妈”。2011年，一个2岁女童从10楼坠落，正在楼下的吴菊萍冲过去用左臂接住了孩子。小女孩得救了，而吴菊萍左臂粉碎性骨折，人们称赞吴菊萍为“最美妈妈”。此后各种“最美”称号不断涌现，如最美教师张丽莉、最美司机吴斌、最美战士高铁成、最美女法官后莉……

10) **接地气** 本义是“接土地之气”，民间常说的“水土不服”就是不接（当地的）土地之气（泛指自然环境和气候）。流行语“接地气”中的“地”用的是比喻义，指老百姓的生活。“接地气”就是“贴近老百姓真实生活的实际”“反映百姓真实生活情感”。莫言的小说就是“接地气”的一个典型。而党员干部“接地气”，就是要深入基层，广泛接触民众，了解民众的苦与乐、民众的意愿和需求。

## 荐 《咬文嚼字》发布“2013年十大流行语”，“中国梦”位居榜首

2013-12-18 16:44:44

标签：文化

1.中国梦 党的十八大以后，习近平总书记在参观国家博物馆主办的“复兴之路”展览时，提出要实现中华民族伟大复兴的中国梦。梦是理想，是目标。中国梦的基本内涵是，国家富强，民族振兴，人民幸福。中国人民要共筑中国梦，共圆中国梦。必须坚持中国道路，弘扬中国精神，凝聚中国力量，才能实现中国梦。中国梦以其清新的理念和亲和的风格，为大众所认同，并日渐成为主流话语之一，成为2013年度的全民流行语。

2.光盘 光盘就是吃光盘中饭菜的意思。2013年1月，北京一家民间公益组织发起“光盘行动”，倡议市民就餐后打包，“光盘”离开。随后，中央电视台新闻联播，号召大家“节约粮食，从我做起”，“争做节约达人，向舌尖上的浪费说再见”。“光盘行动”迅速席卷全国，“光盘”被捧为时尚新词，“今天你光盘了吗”成为流行语。

3.倒逼 倒逼，即逆向促使。按照常规，甲乙两事物，甲动则乙动；如果一反常规，乙先动逆向促使甲动，这种逆向促动，叫作倒逼。倒逼来源于经济领域中货币供给的倒逼机制。如今扩大了使用范围，“由下而上”“由流溯源”“由果问因”等行为，都可以称“倒逼”。习近平、李克强等国家领导人，在讲话中多次用到“倒逼”，提升了这个词的流行度。

4.逆袭 从日语引进的新词，意思是在逆境中反击成功。比如在球赛中，一支球队先输后赢，可以称逆袭；一支弱队战胜了强队，也可称作逆袭。如今意义和用法拓展，新事物冲击旧事物、后浪推前浪等等，都可称逆袭。逆袭表达了百折不挠、奋发图强、充满正能量的精神。

5.微×× “微”本指小、细、轻、少、弱等，也可表示某些计量单位的百万分之一，如微米、微安（安，安培的简称）。如今，“微”成了一个时尚语素，生活中出现了一批以“微”命名的新事物，如微博、微信、微新闻、微电影等等。社会正在迈进“微时代”，“微××”正在改变人们的生活方式以及思维模式。“微”是现代科技的产物，也是一个时代的标签。一个不起眼的“微”在现代社会中释放出了大能量。

6.大V 指的是在微博上十分活跃并拥有众多粉丝的公众人物，通常把粉丝超过50万的微博用户称为网络大V。大V必须在一个特定的微博平台上获得个人认证，经过认证，其微博昵称后便会附一个“V”形图标。之所以称“V”，有两种说法：一说是VIP（贵宾账户）的首字母，一说是verify（认证）的首字母。大V几乎都是网络上的意见领袖，有着不容小觑的号召力和影响力。因此，大V务必维护好自己的形象，不传播未经核实的信息，不能成为谣言的传播者，否则，大V就成了“大谣”。


7.女汉子 “女汉子”指带有“纯爷们性格”的女性。她们爱美妆喜时尚，还能“干粗活打蟑螂”，集温柔和彪悍于一身，拥有女人的外表和汉子的内质。现代大都市中，竞争激烈，压力山大，许多女性为了生存和发展，不得不把自己变成独立、自主、强悍的女汉子。2013年荧屏流行“女汉子”，像孙俪在《辣妈正传》中扮演的角色，就是有个性、有主见、敢作敢为的现代女性。

8.土豪 本是汉语中固有词语。在20世纪上半叶，特指有财有势横行乡里的地主恶霸。今日之“土豪”源自网络，最初用于吐槽在网络游戏中“无脑”消费人民币的玩家，后来又泛指现实社会中富而不贵的群体。从字面上解释，就是“很土的富豪”，他们大多文化较低，品位不高，极其富有，有的还酷爱炫富。人们对“土豪”的态度，也不是一味地鄙视，而是既排斥又羡慕。“土豪，我们做朋友吧！”这句广为流传的调侃语，最能反映人们对“土豪”的心态。

9.奇葩 汉语中本有其词，指珍奇而美丽的花朵，常用来比喻不同寻常的优秀文艺作品，是地道的褒义词。如今广为流行的“奇葩”，则来源于网络，常用来比喻某人某事或某物十分离奇古怪，世上少有，常人不可理解，含有讽刺或调侃的意味。如有媒体曾经盘点全国众多景区“免收门票”的“奇葩”招数，包括“叫妈免票、属猴免票、背古文免票、穿短裙免票、姓朱免票”等等，招招都让人瞠目结舌。

10.点赞 起源于各大社交网站的“赞”功能。网民对网络上一个帖子、一篇文章或者一条微博的内容表示赞同、欣赏、支持或声援，可以点击一下帖子下面的“赞”图标。“点赞”就是点一个赞，“点”与“赞”是动宾关系。当下频频出现在报刊上的“点赞”，词义有了演变，它成了点评的一种。与点评不同的是，点赞只说好话，即只表示肯定、喜爱、赞美这一类意思。

2015

 2015年十大流行语 (2015-12-15 16:30:29)

 转载 

标签：人文/历史 文化 杂谈

## 2015年十大流行语

《咬文嚼字》编辑部

(2015年12月)

### 一、获得感

“获得感”本表示获取某种利益后所产生的满足感。2015年2月27日，习近平总书记在中央全面深化改革领导小组第十次会议上指出，要科学统筹各项改革任务，推出一批能叫得响、立得住、群众认可的硬招实招，把改革方案的含金量充分展示出来，让人民群众有更多“获得感”。“获得感”一词由此迅速流行，且使用范围出现固化趋势，多用以指人民群众共享改革成果的幸福感。

### 二、互联网+

“互联网+”就是互联网与社会经济领域各生产要素“相加”，即让互联网与传统行业深度融合，重构、再造新的发展业态。符号“+”既表示加入、融入，更表示升级换代、创新发展。“互联网+”既是一种新的思维方式，也是一种新的发展模式，更是一种新的经济形态。这一说法最初是由李克强总理提出的。2015年3月5日，李克强总理在《政府工作报告》中说，要全面制定“互联网+”行动计划，“互联网+”便迅速流行开来。2015年7月，国务院发布《关于积极推进“互联网+”行动的指导意见》，“互联网+”进一步引起社会广泛关注。2015年，随着经济发展新格局的形成，“互联网+”成为媒体的高频用词。

### 三、颜值

“颜”义为面容、容貌，“值”义为数值。“颜值”本表示男女容颜英俊或靓丽的程度，是用数字评价人物的容貌。如同其他数值一样，“颜值”也有衡量标准，可以测量和比较，所以有“颜值高”“颜值爆表”“颜值暴跌”的说法。后来“值”的数值意义淡化，在词义上“颜值”就相当于“颜”，只表示面容和姿色。2015年，其指称范围进一步扩大，由人及物，物品的外表或外观也可用“颜值”表示。

#### 四、宝宝

“宝宝”语出网络潮语“吓死宝宝了”。“宝宝”指“我”，“吓死宝宝了”就是“吓死我了”。女生受到惊吓时常用此语来卖萌。后来“宝宝”独立单用，只要说话氛围协调，几乎在任何语境中，女生都可用“宝宝”或“本宝宝”来称呼自己，如“乐死宝宝了”“笑死宝宝了”“本宝宝这厢有礼了”“本宝宝拜托了”等等。现在，“宝宝”还可用以指称对方，以表达对对方（不论男女）的亲昵态度。购物网上过去常用“亲”来缩短与客户的心里距离，现在“宝宝”也争得一席之地，且大有抗衡之势。

#### 五、创客

“创”指创造，“客”指从事某种活动的人。“创客”本指勇于创新，努力将自己的创意变为现实的人。这个词译自英文单词“Maker”，源于美国麻省理工学院微观装配实验室的实验课题。此课题以创新为理念，以客户为中心，以个人设计、个人制造为核心内容，参与实验课题的学生即“创客”。2015年3月5日，李克强总理在《政府工作报告》中指出，把“大众创业，万众创新”打造成推动中国经济前行的“双引擎”之一。随后国家又陆续出台了若干政策措施，给“创业”“创新”以具体支持。“创客”于是与“大众创业，万众创新”联系在了一起，特指具有创新理念、自主创业的人。

#### 六、脑洞大开

“脑洞大开”由“脑补”衍生而来。“脑补”源自日本动漫，本指观剧者在脑子里通过自己的想象来补充或添加原剧中没有的情节或内容。后来词义扩大，对小说及其他艺术作品甚至现实生活中的情节、情景进行幻想、想象，也称“脑补”。“脑洞”则是大脑中需要用想象力去填充的“洞穴”，即进行“脑补”的场所。“洞穴”越大，想象越丰富。“脑洞大开”义为想象天马行空，联想极其丰富、奇特，甚至到了匪夷所思的地步。2015年2月，网络剧《脑洞大开》播出，“脑洞大开”于是频频亮相于媒体。

#### 七、任性

“任性”本是一个传统词语，意思是“由着自己的性子，不加约束”，常用作贬义，如“任性妄为”。江西曾有一位老伯在网上买保健品被骗，接二连三地给骗子寄钱，发觉受骗后，不但不报警，反而继续寄，总共被骗数十万元。老伯最后还语出惊人：“就是想看看他们究竟能骗我多少钱！”有网友调侃：“有钱就是这么任性。”此语迅速传播开来，以致引爆网络，还衍生出各种变体，如“有权就是这么任性”“有房就是这么任性”“有车就是这么任性”。随着高频率运用，“任性”词汇色彩逐渐变化，也可用来表示“天真率性”“敢作敢为”“勇于担当”等义。



## 八、剁手党

“剁手党”，指沉溺于网络购物的人群，以女性居多。也称“剁手族”“剁手帮”等。她们穿梭于各大购物网站间，搜寻、比价、秒拍、下单，乐此不疲。往往看似精打细算，实则买回了大量没用的东西，造成时间、精力、金钱的浪费。冷静之后，常常追悔莫及，甚至下了再不理性购物就剁手的狠心，但过不了多久又故态重萌。金庸《射雕英雄传》中的九指神丐洪七公，有一次因贪吃误了大事，一发狠剁掉了自己的手指头，但一遇美食仍然把持不住。“剁手”或源于此。

## 九、网红

“网红”即“网络红人”，指被网民追捧而走红的人。“网红”走红的原因很复杂，或因出众的才貌，或因搞怪的行为，或因意外事件，或因网络推手的运作，等等。“网红”是网络时代的产物，其身上投射出的特质，迎合了网络世界的“审美”或“审丑”心理，从而赢得了广大网民的理性或非理性关注。“网红”在2015年热度飙升，和媒体纷纷爆出明星大腕与“网红”女友的恋情有关。郭富城的新女友便是最近的一例。

## 十、主要看气质

“主要看气质”，意思是不要太看重外在形式，内在气质才是决定因素。此语在2015年末突然蹿红，成为一个强势流行语。起因是台湾女歌手王心凌为了配合新专辑的发行，在微博上发布了一张吃汉堡的宣传照。网友纷纷留言评论，其中有人说品位不高，甚至说没文化。一位网友却认为，虽然是一张奇葩照片，但颇有内涵，于是留言：“主要看气质。”这则留言立马获得广大网友点赞，被大量转发，火速引爆网络媒体。有人说，在过分追求“颜值”的时代，“主要看气质”的流行有“正能量”意义。

标签：文化

### 一、洪荒之力

“洪荒”本指混沌、蒙昧的状态，借指远古时代。传说，在天地初开之时，曾经有过一次大洪水，几乎毁灭了整个世界，大地一片荒芜。“洪荒之力”即指如天地初开之时足以毁灭世界的自然之力。在2015年热播的仙侠剧《花千骨》中“洪荒之力”频频出现，指的是法力无边且无法控制的妖神之力。2016年8月8日，在里约奥运会女子100米仰泳半决赛中，中国选手傅园慧晋级决赛。她在接受媒体采访时说：“我已经用了洪荒之力啦。”“洪荒之力”一夜走红，多用来形容超乎想象的巨大能量。

### 二、吃瓜群众

现实生活中，人们常一边嗑着瓜子，一边听人闲聊。在网络论坛中，人们发帖讨论问题，后面往往有一堆人排队跟帖，或发表意见，或不着边际地闲扯。如果只看热闹，不发言，便称“吃瓜子”。为了加快输入速度，“吃瓜子”后被简称为“吃瓜”。“群众”即普通人。不发言只围观的普通网民，则被称为“吃瓜群众”。现在使用范围扩大，凡是对某议题不了解或有意保持沉默的围观者，都可称“吃瓜群众”，不限于网络论坛中的网民。

### 三、工匠精神

“工匠精神”本指手工艺人对产品精雕细琢、追求极致的理念，即对生产的每道工序，对产品的每个细节，都精益求精，力求完美。2016年3月5日，李克强总理在《政府工作报告》中说，鼓励企业开展个性化定制、柔性化生产，培育精益求精的“工匠精神”。“工匠精神”一词迅速流行开来，成为制造行业的热词。随后，不仅制造行业，各行各业都提倡“工匠精神”。于是，使用范围扩展，任何行业、任何人“精益求精，力求完美”的精神，都可称“工匠精神”。

### 四、小目标

现在流行的“小目标”出自万达集团董事长王健林之口。他在《鲁豫有约大咖一日行》中谈到“很多年轻人想当首富”的话题时表示：“想做世界首富，这个奋斗的方向是对的，但是最好先定一个能达到的小目标，比如我先挣它1个亿。”就“世界首富”这个“大目标”来说，“1个亿”确实是一个“小目标”，王健林的话并无不妥；对普通人来说，“1个亿”几乎是遥不可及的“天文数字”，于是网友纷纷吐槽：“这个目标确实有点小”（反语）……“小目标”立马在网上走红，其意思正好与本义相反，指的是普通人难以达到的“大目标”，多含讥讽和自嘲的意味。

## 五、一言不合就XX

汉语中一直有“一言不合”的说法，义为一句话说得不投机，后面紧跟由此造成的后果，如“一言不合就翻脸”。现在流行的“一言不合就XX”，和2015年末举行的世界电子竞技大赛有关。这次大赛主办方的办赛态度极其粗疏，引起广大观众的强烈不满。一些网民便在其百度贴吧中贴出低俗内容以表示愤怒，导致贴吧被封。网络上把传播低俗内容的行为叫作“开车”。于是有人评论道：“现在的年轻人，一言不合就开车。”“一言不合就XX”，便在网上流行开来。随着使用的泛化，“一言不合”和后面的XX失去了语义上的逻辑联系，仅仅表示“突然”“任性”“动不动”的意思，如“老师一言不合就表扬学生”“单位一言不合就发奖金”“投资房地产，一言不合就赚得钵满盆满”“股市一言不合就下跌”等等。

## 六、友谊的小船，说翻就翻

“友谊的小船，说翻就翻”，意思是朋友之间的友谊很脆弱，随时都有打碎的可能。2016年3月31日，一个叫“喃东尼”的画师在其微博上贴出了一组名为《友谊的小船说翻就翻》的漫画，内容为：两个生动可爱的小企鹅是对好朋友，坐在同一条小船上，因为一个小原因（如一方变瘦），小船立马弄翻。这组漫画很快受到追捧，网民纷纷配上新的文字，掀起了一场“翻船体”造句大赛，如：“你要吃重庆火锅吗？我怕辣。友谊的小船，说翻就翻！”稍后，“说翻就翻”的搭配对象进一步拓展，凡是“搞砸”的事情都可以搭配，如“油价的小船，说翻就翻”“婚姻的小船，说翻就翻”“师生关系的小船，说翻就翻”等等。有人说，“友谊的小船”可能来自“友谊”的英文单词friendship，将其中的词缀ship作“船”解，friendship就可以理解成“友谊的船”。经典美剧《老友记》的台词中就有这种用法：“什么船是永不沉没的？友谊的船（a friendship）。”

## 七、供给侧

2015年11月10日，在中央财经领导小组第十一次会议上习近平总书记提出了“供给侧结构性改革”：“在适度扩大总需求的同时，着力加强供给侧结构性改革，着力提高供给体系质量和效率，增强经济持续增长动力。”在2016年“两会”上，“供给侧改革”又被反复提及，成为新闻媒体中的高频词。“供给”与“需求”相对；“供给侧”与“需求侧”犹如一枚硬币的两个面。“供给侧”即“供给一边”。进入21世纪以后，我国经济进入“新常态”，光靠“需求侧”拉动，难以实现经济的平稳增长，必须在“供给侧”改革发力。“供给侧”改革指明了全面深化改革的方向，已在各行各业展开，前景令人期待。

## 八、葛优躺

2016年7月25日，一组“葛优躺”表情包开始出现在网络上。其图片源自曾热播的大型家庭情景剧《我爱我家》，葛优饰演的“二混子”去别人家蹭吃蹭喝，像一摊烂泥瘫坐在沙发上，神态“妙趣横生”。随着表情包的走红，网民们便把极其懒散的瘫坐姿势称作“葛优躺”。有人说，如今生活节奏越来越快，生活压力越来越大，内心累积了越来越多的“负面情绪”，“葛优躺”式的“颓废”迎合了人们放空一切、降低焦虑、释放压力的心理需求。这正是“葛优躺”一词流行的心理背景。

## 九、套路

汉语中本有“套路”一词，指编制成套的武术动作，如少林拳套路；也指成系统的技术、方式、方法等，如改革新套路。2016年网络流行语中的“套路”旧瓶装新酒，其“套”不是“成套”“成系统”的意思，而是指“圈套”“老套”。“套路”泛指经过精心编制的、用来迷惑人的说法或做法，甚至诡计、陷阱等，如“骗子行骗的套路就是恐吓，让人心生恐惧，从而听其摆布”。有时也指人们早已习惯的约定俗成的处理各种事情的手段、方式、方法等，如“记者采访的套路大家都很熟悉”。

## 十、蓝瘦，香菇

“蓝瘦，香菇”是“难受，想哭”的谐音。2016年10月，广西南宁一小哥失恋后录了一视频：“难受，想哭，本来今天高高兴兴，你为什么要说这种话？难受，想哭……”由于带有浓厚的当地口音，“难受，想哭”听上去成了“蓝瘦，香菇”。这一谐音既有调侃味又有形象感，立马受到年轻人的追捧，频频出现在微博、公众号中。一些大学的食堂，还跟风推出了“蓝瘦香菇”炒菜（含炒西蓝花、瘦肉、香菇）。随后还有人抢注了蓝瘦香菇实业有限公司。“蓝瘦，香菇”的流行，迎合了年轻人在表达上的游戏化心理。

2017

荐 2017年十大流行语 (2017-12-19 16:19:36)

+ 转载 ▼

标签：文化



一、不忘初心 不忘初心，即不忘本意、不忘初衷。2016年7月1日，在庆祝中国共产党成立95周年大会上，习总书记指出：面向未来，面对挑战，全党同志一定要不忘初心、继续前进。2017年10月18日，中国共产党第十九次全国代表大会胜利召开，“不忘初心，牢记使命”成为大会主题的关键词。习总书记在报告中指出：中国共产党人的初心和使命，就是为中国人民谋幸福，为中华民族谋复兴。这个初心和使命是激励中国共产党人不断前进的根本动力。“不忘初心”由此迅速流行开来，广泛用于各行各业中。

二、砥砺前行 砥砺，本义为磨刀石，引申指磨炼。砥砺前行，即在磨炼中奋勇前进。自2017年6月起，为了迎接十九大召开，各地响应中央号召，开展了以“砥砺奋进的五年”为主题的宣传活动，用各种形式展现十八大以来的五年中我国在经济发展、民生保障、生态文明、科技进步、文化教育、区域发展等方面取得的成就，激励广大干部群众坚定中国特色社会主义道路自信、理论自信、制度自信、文化自信，扎实推进各项工作，以优异成绩迎接党的十九大胜利召开。“砥砺前行”准确地概括了以往五年的成就，生动地描绘了中华民族的形象，迅速流行于各类媒体上。

三、共享 共享，义为共同享有、共同享用。共享单车的出现，解决了人们从住宅到公交站台、地铁等“最后一公里”的问题，得到了人们的认可。“共享”是互联网技术、移动支付技术日益成熟条件下的新的经济模式。“共享充电宝”“共享雨伞”“共享篮球”等等一系列与“共享”有关的产物纷纷进入了人们的生活。有报告显示，至2020年中国“共享经济”市场交易规模将占GDP的10%以上。“共享”一词的流行，体现了社会的想象力和创造力，展示了社会管理和人际关系的巨大变化。

四、有温度 2017年5月8日，中国共产党上海市第十一次代表大会隆重开幕。大会报告中指出，上海未来五年要成为人文之城：建筑是可以阅读的，街区是适合漫步的，公园是最宜休憩的，城市始终是有温度的。“有温度”这个有着细腻感觉的词语，此后逐渐流行开来。它的意思是让人感受到热度、暖意，含义由物理层面扩展到了人文层面，和我们党的为人民服务的宗旨高度契合。“有温度”的流行反映了人们对城市化建设中的人性化要求。

五、流量 “流量”本义是单位时间内通过河、渠或管道某一横截面的流体的量，或是通过道路的车辆、人员等的数量。在互联网时代，“流量”也指在一定时间内网站的访问量，以及手机等移动终端上网所耗费的字节数。明星大腕显然能让网站访问量以及手机等移动终端上网流量暴增。从2016年年底、2017年年初开始，娱乐圈频频用“流量”来形容那些粉丝多、人气高、影响广、商业价值大的明星。这里面既有市场层面的客观描述，又隐含着公众的文化评价，是一个耐人寻味的词语。

六、可能XXX假XXX 2017年1月，网络上出现一组调侃期末考试成绩不理想的表情包。此表情包由“我可能复习了假书”“我可能拿到了假试卷”“老师可能划了假重点”等句子配上各种苦恼表情符号组成，迅速在大学生群体中传播开来。“可能XXX假XXX”的句式便在此基础上形成，并快速传遍朋友圈和微博。“假”本指不真实的、伪造的等义，是一种客观存在；但在这个句式中，“假”却不是这个意思，而是指使用效果不佳、违反常理、不合预期，是一种主观感受。人们用对“假”的埋怨、指责来调节自己的情绪。

七、油腻 “油腻”本指“含油过多”。2017年10月24日，有网友发了一条微博谈“中年男性去油腻步骤”，立即引起网友热议，“油腻”因此引爆网络。随后，冯唐、高晓松等在各自的微博上发布博文，谈如何避免成为一个油腻的中年猥琐男。“油腻”的走红，反射出人们在物质生活条件提高后，对如何追求健康生活方式的反思。现在“油腻”的含义进一步扩大，油腔滑调、世故圆滑、不修边幅、邋遢不堪等等，都可称“油腻”。

八、尬 “尬”常用在双声联绵词“尴尬”中，一般不单用。“尴尬”指处境困难，不易处理，或行为、态度不正常、不自然。台湾有“尬舞”等词，“尬”作“比、斗”解，“尬舞”的意思是斗舞、比舞。“尬舞”等传入大陆后，人们接受了这种词语形式，却仍把“尬”理解成“尴尬”，“尬舞”便成为“尴尬地舞蹈”的意思。由此“尬”独立使用，成为新的流行用法，如“尬聊”“尬唱”“尬煮”“尬谈”等等，其中的“尬”均是“尴尬”的意思。

九、怼 怼，读作duì，义为怨恨，引申指凶狠。在河南方言中，“怼”音duǐ，表示收拾、冲撞。作为网络流行词的“怼”源自综艺节目《真正男子汉》。在该节目的第二季中，某嘉宾被“收拾”了一顿，满怀委屈地问教官：“这就叫被怼吗？”教官便探讨了一番“小怼小进步，大怼大进步，不怼不进步”的“理论”。“怼”因此渐渐流行起来，人们用它来表示故意找茬、反抗、反对等意思。后来其义扩大，进一步衍生出“比拼”“竞争”等含义。

十、打call “打call”一词源自外语，call在英文中是“喊、叫”的意思。“打call”最开始的意思是演唱会现场观众疯狂喊叫，后指跟随音乐的节奏，用挥动荧光棒等方式与台上的表演者互动。进入我国后，其适用范围逐步扩大开来，也用来表达对某人、某事、某物的支持和赞成。

## 6b : Etymologie et définition des sélections selon le Rapport 2014 sur le langage d'Internet en Chine

表 2014 网络流行语词源

网络流行语	词源来源
萌萌哒	萌萌哒是“萌萌的”的意思，由网络热词“么么哒”受到萌系文化影响演变而来，起始于豆瓣小组，是“该吃药了”的含义，即医治“中二病”（比喻日本青春期的少年过于自以为是等特别流行的俗语）之意，后由于新堂爱的动态图片而流行起来。
逗比	词汇发源于 2013 年，广泛进入网络使用是在 2014 年，根据合成词原则可以理解为：逗逗的+比（相仿俚语）。简单的说，就是说某个人很逗，有点犯傻，有点可爱。如果对陌生人说，指别人是“傻叉”的意思多。如果是好朋友之间，就是调侃和玩笑的意思多。
我也是醉了	是一种对无奈、郁闷、无语情绪的轻微表达方式。通常表示对人物或事物，无法理解、无法交流。多可与“无语”、“无法理解”、“无力吐槽”换用。我也是醉了最初源自武汉地区，网络出处为东来东往歌曲《我也是醉了》。
且行且珍惜	3 月 31 日，马伊琍微博回复文章出轨事件中写到“恋爱虽易，婚姻不易，且行且珍惜。”马伊琍“且行且珍惜”大度容忍了丈夫的出轨行为，众网友纷纷回头。一时间且行且珍惜流传于网络。
涨姿势	“长知识”的谐音，意指让人长见识了，开眼界了。属于网络用语，流行于微博、论坛等交流互动社区。许多人用的是拼音输入法也会将谐音前置，例如说压力大，为“鸭梨大”。
心塞	心塞，网络用语，“心坎梗塞”的简称，意指一种情绪：心理堵得慌，难受，说不出来的痛苦，周围有不顺心的事让你心里很不舒服。或者表示对某件事情很无语。
暖男	被誉为新型好男人的代名词，暖男（Sunshine Boy）本意指的是像熟日阳光那样，能给人温暖感觉的男子。小清新强调外在形象，而同系列的暖男却更强调内在。同时也称顾家暖男。指那些顾家、爱家，懂得照顾老婆，爱护家人，能给家人和朋友的温暖的阳光男人。
买买买	调侃微博红人王思聪和爸爸王健林的对话：王思聪：爸，这个……王健林：买买买！王思聪：爸，这个……王健林：买买买！
也是蛮拼的	网络用语，表示“挺努力”或“努力了，没有成功”出自网络剧《白衣校花与大长腿》热帖，然后通过《爸爸去哪儿 2》曹格的使用被熟知，并广泛传播。
逼格	名词，意为装逼的格调，指一个人装腔作势的格调。
画面太美我不敢看	出自蔡依林的《布拉格广场》，常见于论坛、博客、微信、贴吧，多用于调侃和自嘲。
小鲜肉	用于形容男演员，指年轻、帅气、有肌肉的新生代男偶像。一般是指年龄在 12-25 岁之间的性格纯良，感情经历单纯，没有太多的情感经验，并且长相俊俏的男生，也有引申义指年龄在 12 到 18 岁之间的性格纯良，感情经历单纯，没有太多的情感经验，并且长相美丽的女生。
那么问题来了	最早来源于蓝翔技校的广告语——挖掘机，哪家强？中国山东找蓝翔。2014 年年网评论区的会员张乐的留言：那么问题来了，学挖掘机到底哪家强？其实就是网友娱乐精神的体现。由于句式简单，朗朗上口，一时间红遍大江南北。（也有网民指出这是网络推

	手进行的商业炒作。）
no zuo no die	中式英语，意思是你如果不干蠢事，蠢事就不会反过来害你（但如果你干了，百分之百摆出来还的）。zuo 是中文汉字，意思就是“找死（找蠢的）”。采用拼音混编英文的风格来表达网络热词“不作死就不会死”的意思。
脑洞大开	脑洞是从“脑补”衍生出来的词。因为有空洞，所以需要脑补。脑补通常是指在头脑中对某些情节进行脑内补充，“脑洞大开”即指脑补能力非常强，想象力非常强。
有钱就是任性	网友调侃一件诈骗事件而引用的网络流行语，并在各大网络平台渐火。本意是对于有钱人的做事风格具有嘲讽语义，而现今多出于好友间或者微博空间论坛上调侃用语。
时间都去哪儿了	《时间去哪儿了》原是歌曲，2014 年王铮亮首度受邀参加中央电视台春节联欢晚会，自弹自唱《时间去哪儿了》，歌词朴实细腻、深情真挚，触动观众的心。2 月 7 日，中国国家主席习近平在接受俄罗斯电视台专访时提到了这首歌，引发了以“时间去哪儿了”的全民讨论。
保证不打死你	“保证不打死你”则源自吴镇宇在《爸爸去哪儿 2》中的一句口头语：“宝宝你过来，爸爸保证不打死你”。
我读书少你别骗我	源于 2012 年马伯庸所出版的图书《我读书少，你可别骗我》书名。
习大大	陕北方言，指父亲的干弟、胞弟和堂弟。叔叔指比父亲小的非胞弟、堂弟、干弟之外的同辈。意思是叔叔。早先网友以“习大大”称呼习近平主席，表达自己对习近平主席亲近爱戴之情。
上天台	2014 年巴西世界杯期间网络新出现的词语。由于本届世界杯的赛事冷门不断，很多买彩票者受到结果刺激。网络上开始流传出买球输了，我们一起去天台的说法，意为买球输光了，一起去跳楼。这种说法绝大多数属于开玩笑，是本届世界杯开始出现的自嘲减压的新词汇。
什么仇什么怨	11 月 24 日，苏州梁公交车上，一位乘客随地乱扔瓜子壳被一名乘客制止，但该男子不听，并纠缠阻碍者长达 7 分钟，并不停地重复说“我跟你有什么仇什么怨”达 37 遍。事件被网友拍下并上传网络，“什么仇什么怨”迅速成为网络流行语。
你家里人知道吗	“你家里人知道吗”起源于调侃一个在网吧看新闻联播的网友，然后成了网络流行语。
挖掘机技术哪家强	本是蓝翔技校招生的广告语，最先在网易跟帖区与百度贴吧零星出现，后演变成网络流行语。通常的使用方法是，先一本正经地讲述一个故事，然后在故事的结尾引导到“挖掘机”“挖坑”“埋了”等相关词汇。
APEC 蓝	北京 APEC 会议期间，京津冀实施道路限行和污染企业停工等措施，未保证空气质量达到良好水平。网友形容此时天空的蓝色为“APEC 蓝”。

## 6c : Captures d'écran de certains articles de presse portant sur les nouvelles entrées des dictionnaires français

*Le Figaro*

### **Ambiancer, arty, spin-off, émoticône... Le Petit Larousse accepte 150 nouveaux mots**

La cuvée 2017 voit l'entrée en force des expressions gastronomiques ainsi qu'aux mots issus des nouvelles technologies. Parmi les 50 nouvelles personnalités : Vincent Lindon, Sandrine Kiberlain, les Monty Python et Élisabeth Badinter.

Par Mohammed Aïssaoui  
Publié le 11/05/2016 à 08:04,  
Mis à jour le 11/05/2016 à 15:43

Elle ne paraîtra que le 26 mai, mais on connaît déjà quelques mots de l'édition 2017 du fameux dictionnaire Le Petit Larousse. 150 nouveaux mots font leur entrée ainsi que 50 personnalités pour la partie «Noms propres».

Quelles sont les tendances de ce nouveau millésime? «Les expressions culinaires et gastronomiques font leur entrée en force, ainsi que les néologismes issus du monde des arts. Quant aux nouvelles technologies, elles offrent toujours leur lot de nouveautés», explique Carine Girac, directrice du département dictionnaires et encyclopédies chez Larousse. Et d'ajouter: «En revanche, on remarque qu'il y a moins de termes venant de l'écologie et de la psychologie.»

Des exemples? **Burrata**, mot italien, nom féminin (fromage à pâte filée proche de la mozzarella) ou **ciabatta** (mot italien également, qui signifie «savate», par analogie de forme, c'est un pain à l'huile d'olive, plat et rectangulaire, dont la mie très aérée est recouverte d'une fine croûte lisse) entrent dans Le Larousse.

<https://www.lefigaro.fr/livres/2016/05/11/03005-20160511ARTFIG00058-ambiancer-arty-spin-off-emoticone-le-petit-larousse-accepte-150-nouveaux-m...> 1/4

Comme **argouse** (baie orangée comestible), **phô** (mot vietnamien, bouillon de bœuf agrémenté de nouilles de riz), **wrap** (galette de blé ou de maïs) ou **yuzu** (mot japonais, arbuste du groupe des agrumes).

Le monde de l'art innove. On peut désormais dire «**arty**» (qui présente ou revendique une dimension artistique), «**cinégénique**» (dont l'image cinématographique produit un bel effet) ou **fanfiction** (nom féminin composé de fan et de fiction qui signifie récit proposé par un fan sur Internet, qui fait suite à une fiction préexistante -roman, manga, série, film, jeu vidéo - ou en constitue une variation). Le **Mook** (de magazine et book, le nom est pourtant déposé) est entré dans les mœurs et dans le dictionnaire. «**Rétrofuturisme**», «**spin-off**» peuvent désormais se dire sans que ce soit fautif. Souvent, c'est l'anglais qui l'emporte, mais au lieu de one-man-show ou pourra désormais écrire «seul en scène».

### **Téléverser ses photos ou ambiancer**

Côté sciences et informatique, **Fablab** (laboratoire de fabrication) est admis. Tout comme **flasher**, **open source**, **QR code**, **troll**, **émoticône** (avec un accent circonflexe), **déréférencement**, **défaçage** (ou **défacement**: action de détourner la présentation d'un site web, notamment sa page d'accueil, mésaventure arrivée à TV5 Monde) ou **téléverser** (transférer des données d'un ordinateur local vers un ordinateur distant, via un réseau de télécommunications, exemple: téléverser ses photos sur un site Web).

Est-ce un effet de notre époque? **Autocentré** (qui centre tout sur soi-même, rapporte tout à soi) est maintenant dans Le Petit Larousse. Idem pour **covoiturer**, **socialiser**, **zone de confort** ; **économie collaborative**, **zadiste** (ce mot était dans l'édition 2016 du Petit Robert), **loup solitaire**, **antisystème**, **complotisme**, **déclinisme**, **europheobe**, **exfiltrer**, **lobotomiser**... Les **REP** (acronyme de zone d'éducation prioritaire) remplacent les ZEP. Et le virus **Zika** fait son apparition dans le dictionnaire.

**S'enkyster** n'a rien à avoir avec un kyste: la définition que donne Le Larousse est «s'installer durablement, en parlant d'un phénomène indésirable, d'une situation pénible». On pensait que «**patate**» existait déjà, mais c'est le sens d'«être en forme, avoir de l'énergie» que retient le dictionnaire.

<https://www.lefigaro.fr/livres/2016/05/11/03005-20160511ARTFIG00058-ambiancer-arty-spin-off-emicone-le-petit-larousse-accepte-150-nouveaux-m...> 2/4



La francophonie a toujours donné naissance à de nouveaux termes. Exemples: **ambiancer** (issu d'Afrique), c'est-à-dire faire la fête, ou **sans-allure** (venu du Québec), «qui fait preuve d'un manque de discernement ou de savoir-vivre, qui se conduit de manière stupide ou discourtoise». Voilà un mot qui ne manquera pas d'occasions d'être utilisé.

### **Les nouvelles personnalités de l'édition 2017**

Plus de cinquante personnalités entrent dans le Panthéon du célèbre dictionnaire. C'est une immense reconnaissance. «On intègre une personne qui a œuvré pour la culture et le savoir, avec l'idée de pérennité», souligne Carine Girac. Impossible de toutes les citer. Il y a l'artiste chinois **Ai Weiwei**, la lauréate du Prix Nobel de littérature **Svetlana Aleksievitch** (elle entre avec cette orthographe), **Élisabeth Badinter**, **Brian De Palma**. À 27 ans, le cinéaste canadien **Xavier Dolan** doit faire partie des plus jeunes dans cette catégorie des noms propres, et pour longtemps, car il n'est pas dans l'habitude du Petit Larousse de retirer des noms. **Robert Guédiguian**, **Sandrine Kiberlain**, **Vincent Lindon**, **Benjamin Millepied**, **Benoît Peeters** (écrivain français et cinéaste de BD), l'économiste **Thomas Piketty**, le sociologue **Michel Wieviorka**, le couturier **Hubert de Givenchy** rejoignent le prestigieux dictionnaire. La troupe du **Splendid** et celle des **Monty Python** les accompagnent. Côté sports: **Cristiano Ronaldo** et le coureur automobile **Lewis Hamilton** peuvent être fiers d'être parmi eux.

### **«Mieux que Google»**

Cette édition fête le bicentenaire de Pierre Larousse né en 1817. Jean d'Ormesson, qui a accepté d'être le parrain du Petit Larousse 2017, lui rend un vibrant hommage. «Vous savez ce qu'il y a de plus extraordinaire chez Pierre Larousse, c'est cette capacité de travail stupéfiante, il a été un grand grammairien, un grand lexicographe et l'auteur d'une œuvre prodigieuse qu'il a accompli en une vingtaine d'années, il était d'une origine très simple, il était passé par l'école primaire et évidemment je crois que, à propos de Larousse, on peut parler de génie», dit-il. L'académicien apprécie également l'ouvrage: «C'est un dictionnaire encyclopédique, si vous vous servez bien du Larousse, oserais-je le dire, vous avez

à peine besoin de Google. Ce petit dictionnaire est un trésor pour quiconque veut se renseigner rapidement et facilement. C'est un instrument de travail, c'est une source de renseignements et c'est un lieu de promenade intellectuelle.»



## La «beuh» entre dans «Le Petit Robert» et les «bolos» dans «Le Petit Larousse»

**DICTIONNAIRES** Les éditions 2016 de ces ouvrages de référence intègrent des dizaines de nouvelles définitions...

F.R. | Publié le 18/05/15 à 08h36 — Mis à jour le 18/05/15 à 08h44

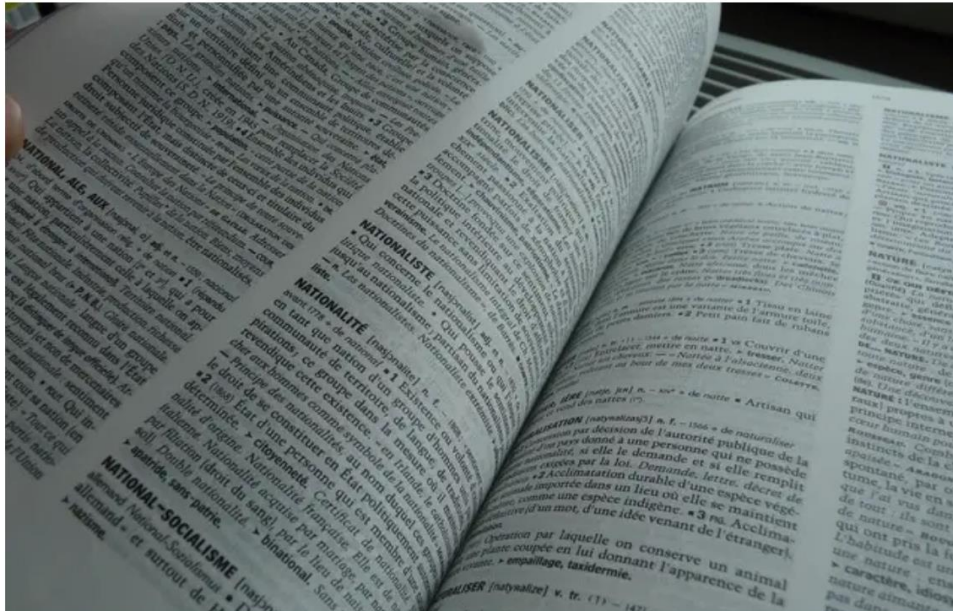


Illustration dictionnaire. — 20 Minutes

*Le Petit Robert* 2016 se met au vert. Les mots « covoiturer », « écoconduite » et « recyclerie » font leur entrée dans la nouvelle édition du dictionnaire qui sera en vente dès jeudi, [révèle Le Parisien ce lundi](http://www.leparisien.fr/societe/les-zadistes-occupent-le-dico-18-05-2015-4780421.php) (<http://www.leparisien.fr/societe/les-zadistes-occupent-le-dico-18-05-2015-4780421.php>). Les « zadistes », ces militants qui s'installent dans les « ZAD » (« Zones à défendre ») [comme ce fut le cas à Notre-Dame-des-Landes](https://www.20minutes.fr/nantes/1473587-20141103-dame-landes-apres-fut-le-cas-a-notre-dame-des-landes) (<https://www.20minutes.fr/nantes/1473587-20141103-dame-landes-apres-fut-le-cas-a-notre-dame-des-landes>

[débordements-week-end-faut-expulser-zad](#)), sont accueillis à bras ouverts par le *Petit Robert*, ainsi que les « faucheurs volontaires » et les « décroissants ».

### **Zadiste, caricature, laïcité... Votez pour le mot de l'année 2015 avec «20 Minutes»**

Les anglicismes n'ont pas non plus été snobés : le « [bitcoin](#) » (<https://www.20minutes.fr/dossier/bitcoin>) (monnaie électronique), le « [captcha](#) » (<https://www.20minutes.fr/high-tech/1494227-20141203-captcha-textes-illisibles-recopier-presque-fini>) ou le « big data » ont été validés, même si, pour ce dernier, le terme français de « mégadonnées » reste recommandé.

### **« Partir en cacahuète »**

Côté familiarités, les « baltringues » figurent parmi les nouveaux entrants. De même que les expressions « prendre cher » (ne pas être ménagé), « tendu comme un string » (être très angoissé, à cran) ou « envoyer des scuds » (critiquer quelqu'un, en plaisantant ou non, sans faire dans la demi-mesure). L'entrée de la « beuh » (qui désigne la marijuana) dans le dico en fera peut être tousser plus d'un qui pourrait bien « partir en cacahuète » (perdre son sang-froid ou, dans le cas d'une situation, dégénérer).

*Le Petit Larousse*, qui en est à sa 160e édition, fait savoir ce lundi que 150 nouveaux mots et expressions intègrent le millésime 2016. Le qualificatif de « bolos » et le verbe « mémériser », ainsi que les « tutos » (les « [tutoriels](#) » ([https://www.20minutes.fr/dossier/tutoriel\\_beaute](https://www.20minutes.fr/dossier/tutoriel_beaute))) et les « selfies » ([cela ne plaira pas à Catherine Deneuve](#) (<https://www.20minutes.fr/people/1604775-20150510-catherine-deneuve-selfies-horreur>)) sont parmi les heureux élus, [annonce Europe 1](#) (<http://www.europe1.fr/societe/selfie-tuto-bolos-decouvrez-les-nouveaux-mots-du-larousse-942150>).

# Les youtubeurs et la twittosphère arrivent dans le dico

Les mots du Net font cette fois encore une entrée massive dans la version 2017 du Petit Robert, en librairie le 19 mai. Ils vont pouvoir s'enjailler et vivre ensemble dans cette bible de notre langue.

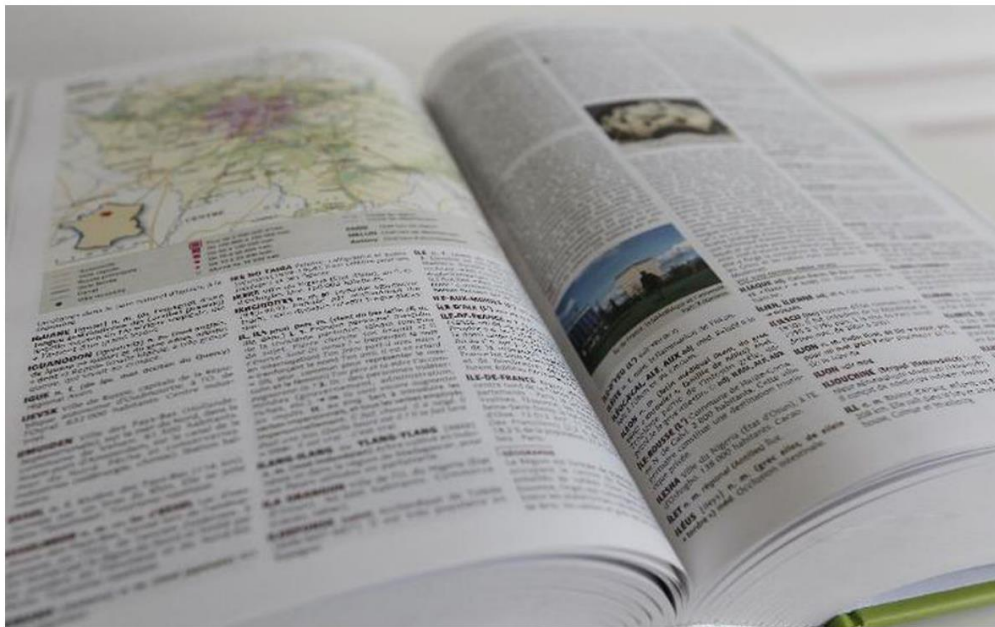


ILLUSTRATION.  
LP/MATTHIEU DE MARTIGNAC

Par **Vincent Mongaillard**

Le 12 mai 2016 à 00h00

Les mots nouveaux sont arrivés. Ils sont quelque 150 noms, verbes, adjectifs,

<https://www.leparisien.fr/societe/les-youtubeurs-et-la-twittosphere-arrivent-dans-le-dico-12-05-2016-5788229.php>

1/7

anglicismes... à s'inviter dans la cuvée 2017 du dictionnaire le Petit Robert en librairie à partir du 19 mai. « Et il n'y a aucun sortant ! On ne rajoute pas de pages mais on joue sur la maquette, sur l'interlignage », dévoile-t-on chez l'éditeur. Voici, en exclusivité, un florilège d'entrants au panthéon des mots.

## Sur les écrans

Les stars de la Toile comme EnjoyPhoenix ou Norman l'ont popularisé, le Petit Robert l'a couronné. Le substantif **youtubeur** (euse), « personne qui publie ses propres vidéos sur le site YouTube », est désormais doté d'une définition officielle. De quoi faire causer sur la **twittosphère**, cette « communauté des personnes qui postent des tweets et de celles qui les lisent ». Ce sont souvent les mêmes qui adorent **geeker**, « passer du temps sur son ordinateur ». Et qui raffolent de l'**emoji**, cette « petite image utilisée dans un message électronique pour exprimer une émotion, représenter un personnage, une action... » Ce nom masculin, marque déposée, est issu du japonais : « e » qui signifie dessin et « moji », lettre. Gare, tout de même, aux excès qui peuvent conduire à la **nomophobie**, cette « dépendance extrême au téléphone portable ». Un smartphone qui s'avérera tout de même fort utile au moment de réserver, via une application, son hôtel ou son taxi. Car aujourd'hui, la plupart des secteurs d'activité de services se font **ubériser**, c'est-à-dire « déstabiliser et transformer avec un modèle économique innovant tirant parti des nouvelles technologies ».

## A table

La gastronomie venue d'ailleurs alimente les nouveautés. Au menu, un **pad thaï**, « plat traditionnel thaïlandais composé de nouilles de riz, d'œufs, de crevettes ou de viande, agrémenté de soja, de cacahouètes... » servi avec un **piquillo**, « piment doux d'une variété produite dans le Pays basque espagnol ». Les végétariens opteront plutôt pour les graines germées d' **alfalfa**, « luzerne cultivée, riche en protéines et en calcium », qui ne remplit pas l'estomac du **viandard**, « personne qui aime la viande, qui en mange beaucoup ».

## Aux terrasses

Après un bon repas, les couche-tard foncent **s'enjailler**, comprenez « faire la fête » en « langage des jeunes ». Ce verbe, issu de l'argot de Côte d'Ivoire, est une déformation de l'anglais to enjoy, s'amuser. Il est synonyme de faire la **chouille**, régionalisme de Lorraine qui a conquis l'Hexagone. En se rassemblant aux terrasses des quartiers métissés, les noctambules contribuent au **vivre-ensemble**, « fait de vivre harmonieusement entre concitoyens, dans l'égalité, le respect et la solidarité ». Après une soirée alcoolisée, à l'heure où démarre, à la télé ou la radio, la journée du **matinalier** (« personne qui anime une émission matinale »), il est recommandé

d'éviter la virée en **fixie**, « bicyclette sans roue libre, au pignon solidaire de la roue arrière ». Il est plus sage d'évacuer les toxines le lendemain au guidon d'un **aquabike**, « sorte de vélo fixe que l'on immerge pour faire de l'exercice dans l'eau ».

**QUIZZ. A quelle décennie ces mots sont-ils entrés dans le Petit Robert ?**

A / France

FRANCE / LANGUE FRANÇAISE

## La beuh, les selfies et les bitcoins arrivent dans le dictionnaire



Publié le : 19/05/2015 - 21:07 - Modifié le : 19/05/2015 - 19:48



Célébration 2015 du Larousse avait intronisé, entre autres, 'vapoter' et 'zénitude' ... FRED DUFOUR / AFP

Texte par : Christophe Carmarans 6 mn

**Comme chaque année à la même époque, les éditeurs Larousse et Robert font connaître la liste des mots nouveaux qui font leur entrée dans la cuvée 2016 de leurs dictionnaires. Apparus dans le langage courant parfois depuis fort longtemps, ils sont néanmoins révélateurs de l'air du temps.**

Au fil des ans, c'est devenu un marronnier. Entendez par là, nous précise **Le Petit Robert** « un sujet rebattu qui reparaît régulièrement, en langage journalistique ». De quoi s'agit-il ? De l'annonce, tous les mois de mai, des mots nouveaux et des expressions nouvelles qui font leur apparition dans le dictionnaire, en l'occurrence dans **Le Petit Larousse** et **Le Petit Robert**, les deux ouvrages de référence les plus vendus en France.

Cet événement annuel permet aux publications en question de se faire un peu de publicité tout en officialisant des mots ou expressions souvent entrés dans le langage courant depuis plusieurs années. Les éditions Larousse et Robert prouvent aussi, en creux, que les linguistes sont parfois en retard de quelques longueurs sur le monde dans lequel nous vivons tous. Ainsi, **Le Petit Larousse** découvre cette année que la **sape** est « un vêtement élégant et à la dernière mode », ce que la plupart d'entre vous saviez sans doute, déjà, depuis plusieurs décennies.

Décalage linguistique

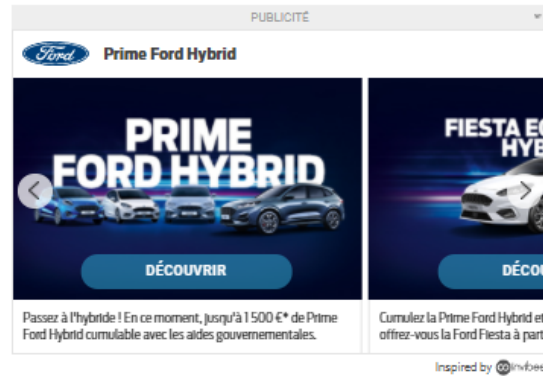


Protégeons le jeune public

CSA CSA



Le Petit Robert n'est pas en reste quand il nous révèle l'existence du **bitcoin**, cette « unité monétaire que l'on utilise sur internet » et qui a été créé en 2008. C'est-à-dire il y a quand même sept ans. Il faut cependant se montrer indulgent avec les éditeurs car ils se sont fixé des critères tout à fait respectables et compréhensibles pour consacrer un mot. Il faut qu'il soit « populaire », « souvent repris par les médias » et, par conséquent, « en rapport avec l'actualité sans risquer de disparaître rapidement ». Voilà d'où vient parfois le décalage avec le langage courant, et en particulier avec le langage médiatique.



Exemple type de vocable médiatisé, le mot **zadiste** qui est, hélas, entré définitivement dans le langage usuel après un événement dramatique survenu le 26 octobre dernier : la mort de Rémi Fraisse, ce manifestant de 21 ans tué après l'explosion d'une grenade lancée par un gendarme sur **le site du projet de barrage de Sivens**, dans le Tarn, un homicide involontaire qui a connu un retentissement considérable en France. Le site de Sivens était désigné par le mouvement écologiste comme une Zone à défendre, donc une ZAD, d'où le mot zadiste qui a été prononcé des milliers de fois cette semaine-là et qui fait donc, logiquement, son apparition dans le dictionnaire quelques mois plus tard.

Cette année, le Larousse et le Robert ajoutent ainsi chacun environ 150 mots et expressions nouvelles à leur répertoire riche de plus de 60 000 entrées. Comme souvent, la technologie est une grande pourvoyeuse de vocables nouveaux. Au bitcoin cité plus haut viennent ainsi s'ajouter le **selfie** - ou **égoportrait comme le décrivent fort à propos les Québécois** – mais aussi le **big data** connu également dans le langage informatique sous le nom de **mégadonnées** et puis l'**open data**, ou donnée ouverte, qui est tout simplement une donnée numérique d'origine publique ou privée.

Nouveau aussi, le **captcha une association de lettres** permettant de différencier de manière automatisée un utilisateur humain d'un ordinateur et donc d'éviter l'envoi de spams. Et enfin le titre de **community manager**, ou gestionnaire de communauté, fonction qui consiste à animer et à fédérer des communautés sur internet. Nous en avons même un, et non des moindres, à rfi.fr.

Autre domaine créateur de néologismes, les transports et leur dérivés qui ont cette année mis au goût du jour les **particules fines** émises entre autres par les moteurs diesel et qui ont imposé en région parisienne la pratique, peu utilisée, de la **circulation alternée**, laquelle consiste à interdire la circulation à certaines catégories de véhicules en se fondant, par exemple, sur les plaques minéralogiques (pairs un jour, impairs un autre jour). Cette pollution intempestive a également favorisé le **covoiturage**, autrement dit l'utilisation conjointe et organisée d'un véhicule, voire même à l'utilisation d'un **gyropode ou transporteur personnel** qui n'est autre un véhicule électrique monoplace à deux roues.

## Newsletter

Recevez toute l'actualité internationale directement dans votre boîte mail

[Je m'abonne ▶](#)





Le duo français Daft Punk est passé à la postérité. AFP

La gastronomie (**bistronomie**), les sciences en général (**anthropocène**), mais aussi la politique (**dédiabolisation**) alimentent aussi régulièrement les dictionnaires en mots nouveaux et nous vous laisserons le soin d'assouvir votre curiosité en découvrant par vous-même l'origine et la signification de ces trois mots, si vous ne les connaissez pas déjà. Plus stupéfiante cette année, l'apparition du mot **beuh** dans le Petit Robert, mot d'argot très usuel qui désigne la marijuana ou encore celui de **bolos** qui dépeint une personne cérébralement défavorisée. Surprise aussi avec la reconnaissance de certaines expressions comme **tendu comme un string** ou **maquillée comme un camion volé** tellement imagées qu'elles ne nécessitent pas forcément, ici, d'explication détaillée.

Comme il ne faut oublier personne, on terminera cette revue de détail par l'arrivée dans les dictionnaires de nouveaux noms propres. Le Larousse accueille dans les pages de son édition 2016 le philosophe **Bernard Henri-Lévy**, le pâtissier **Pierre Hermé**, l'acteur **Michael Caine**, l'artiste **urbain Banksy** mais aussi la sonde **Rosetta** et les studios d'animation **Pixar**. De son côté, Le Robert fait passer à la postérité, pêle-mêle, le cybermilitant (ce mot n'existe pas encore mais tant pis, ce sera pour 2017) **Julian Assange**, l'auteur à succès **Dan Brown**, le duo électro **Daft Punk**, les cinéastes **Ashgar Farhadi** et **Michel Gondry**, le biathlète **Martin Fourcade**, la NSA, sans oublier évidemment **la Zumba** puisqu'il faut bien faire danser tous ces mots...

LIVRES FRANCE LANGUE FRANÇAISE FRANCOPHONIE LITTÉRATURE



**Titre :** Une étude discursive de la circulation des nouveautés langagières sur les réseaux sociaux numériques chinois et français : le cas de *Sina Weibo* et *Twitter*

**Mots clés :** Néologie, Néologisme, Réseaux sociaux numériques, discours numérique, analyse comparée chinois-français, linguistique d'Internet

**Résumé :**

Cette recherche s'intéresse à des productions langagières néologiques diffusées sur les dispositifs numériques de micro blogging qui connaissent souvent une circulation rapide et élargie. Elle prend pour objet des nouveautés langagières en chinois et en français sur une période de 9 années allant de 2009 à 2017 et reconnues comme telles par des sources institutionnelles et ou médiatiques. Les productions langagières sont appréhendées dans l'environnement technorelationnel des réseaux socio-numériques *Sina Weibo* (chinois) et de *Twitter* (version française). L'étude qui se veut à la fois linguistique et communicationnelle relève d'une analyse du discours numérique dans une

perspective écologique et comparée qui tient compte des spécificités des deux langues considérées. Cette perspective amène à explorer la notion de culture numérique participative et les usages des nouveaux dispositifs médiatiques.

L'analyse linguistique d'un corpus d'entrées néologiques et l'analyse discursive d'un corpus de tweets visent à observer si les nouvelles modalités d'expression et les pratiques des usagers influencent la créativité langagière et à voir comment leur circulation va de pair avec les évolutions de sens.

**Title:** A discursive study of newly invented language productions on the French and Chinese social networks: *Sina Weibo* and *Twitter*

**Keywords:** neology, neologism, social networks, digital discourse, Chinese-French comparative analysis, Internet linguistics

**Abstract:**

This research is interested in newly invented language productions and its spread on Social Networks of microblog. This kind of language creation is often propagated rapidly and extensively. It focuses on new language productions in Chinese and French from 2009 to 2017, collected from institutional and/or media sources. Those productions will be observed in the technorelational environment of the social networks *Sina Weibo* (Chinese) and *Twitter* (French version). The study, which is in both aspects of linguistics and communication, is based on an analysis of digital discourse in an ecological and comparative perspective that takes into

account the specifics of two languages. This perspective leads to explore the notion of participatory culture and the uses of new media devices. Meanwhile, the characteristics that define neology and neologism are going to be applied to discuss the nature of new language productions. Then, to figure out the impact of new communication medium and its user practices on language creativity, and the evolution of its meaning during its circulation on social networks, a comparative linguistics analysis of a corpus of neological entries and a discursive analysis of a corpus of tweets are going to be conducted.