



N° d'ordre NNT : 2019LYSE2039

THESE de DOCTORAT DE L'UNIVERSITÉ DE LYON

Opérée au sein de

L'UNIVERSITÉ LUMIÈRE LYON 2

École Doctorale : ED 476 Neurosciences et Cognition

Discipline : Sciences cognitives

Spécialité : Linguistique

Soutenue publiquement le 20 juin 2019, par :

Léa DE CAROLIS

Experimental assessment of sound symbolism and evolutionary considerations.

De la pluralité des principes de justice aux compromis.

Devant le jury composé de :

Sharon PEPERKAMP, Directrice de Recherche, C.N.R.S., Présidente

Ioana CHITORAN, Professeur des universités, Université Paris Diderot, Rapporteur

Padraic MONAGHAN, Professeur d'université, Université d'Amsterdam, Rapporteur

Luca NOBILE, Maître de conférences, Université de Bourgogne, Examineur

François PELLEGRINO, Directeur de Recherche, C.N.R.S., Co-Directeur de thèse

Christophe COUPÉ, Assistant Professor, Université de Hong Kong, Co-Directeur de thèse

Introduction

En linguistique, le terme de motivation, ou Cratylisme¹, correspond à l'idée selon laquelle il existerait un lien direct, fondamental entre des mots et les concepts auxquels ils réfèrent. Cette idée est en opposition avec l'arbitrarité du signe qui définit la relation entre un signifiant et un signifié comme relevant de la convention (Saussure, 1916). Par exemple, la différence de signification entre les mots français « pain » et « bain » ne provient pas de la différence de voisement² elle-même : la signification de « bain » n'est pas plus *voisée* que celle de « pain ». Il n'y a donc pas de significations associées aux segments³ [p] et [b] et leurs significations respectives sont arbitraires (Nuckolls, 1999).

Dans la littérature s'intéressant aux relations motivées, deux termes sont majoritairement utilisés : l'iconicité (e.g. Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015) et le symbolisme sonore (e.g. Imai, Kita, Nagumo, & Okada, 2008). Iconicité correspond à la relation naturelle motivée, comme c'est le cas pour les onomatopées. Le terme de symbolisme sonore est plus inclusif mais reste ambigu au regard de la terminologie de Peirce (1931). En effet, le mot « symbole » est inclus dans l'expression symbolisme sonore et, parallèlement, fait référence à une relation arbitraire dans la terminologie de Peirce. Cependant, on peut considérer que l'expression symbolisme sonore fait plutôt référence à une iconicité relative plutôt qu'absolue. En effet, Gasser, Sethuraman et Hockema (2011) ont proposé une dichotomie entre iconicité absolue (e.g. l'onomatopée) et l'iconicité relative (e.g. le contraste entre les voyelles [i] et [a] représente un contraste de magnitude) (Lockwood & Dingemanse, 2015). Une catégorie de mots, en particulier, peut correspondre à l'iconicité relative, les *idéophones* (appelés aussi *expressives* ou *mimétiques*⁴). Les mots de cette catégorie « évoquent » leurs référents via leur compositions segmentales et à travers des analogies. Par exemple, en japonais, 'buruburu' décrit quelque chose de tremblant (Gomi, 1989). Selon Lockwood et Dingemanse (2015) les idéophones dépendent de la culture, à un certain degré, ce qui les rend, du moins en partie, conventionnels.

¹ Ce terme provient du Cratyle de Platon dans lequel deux positions contrastives au sujet du langage sont juxtaposées, l'arbitrarité et la motivation.

² [p] et [b] s'articulent de la même manière, à la différence que [b] est accompagné de la vibration des cordes vocales.

³ Le terme « segment » est préféré dans cette thèse étant donné que le niveau auquel apparaissent les associations motivées n'est pas réellement défini, à savoir phonétique ou phonologique.

⁴ Le terme *idéophones* est utilisé dans cette thèse en tant que générique pour ce type de mot ; le mot *mimétiques* se réfère plus précisément aux idéophones japonais (qui incluent aussi les onomatopées)

Étant donné la difficulté à déterminer la nature de la relation et le manque d'uniformité entre les auteurs, le mot générique *motivé* sera préféré aux termes *iconique* et *symbolique* afin de ne pas prendre position, en particulier quand la nature de la relation est ambiguë.

Le langage est sans aucun doute arbitraire mais il existe parallèlement de nombreuses évidences de relations motivées. Tandis que l'arbitrarité permet une communication efficace, la motivation a pu représenter un élément clé dans l'émergence et l'évolution du langage. En effet, Imai et al. (2015) ont proposé plus particulièrement que les mots motivés seraient des « fossiles » de stades plus anciens dans l'évolution du langage au cours desquels les relations motivées faciliteraient le développement rapide d'un lexique commun dans les protolangues. En effet, il est peu probable que le langage soit apparu tel quel dans toute sa complexité et une hypothèse courante au sujet de l'évolution du langage propose un scénario en deux étapes selon lequel le système doublement articulé proviendrait d'un système plus basique (Tallerman, 2011). Ainsi, le stade précédent de notre système de communication a pu nécessiter des relations directes et iconiques entre signifiants et signifiés afin que la convention apparaisse et se diffuse entre locuteurs. Plus précisément, l'hypothèse est que les productions motivées auraient facilité l'accord entre les locuteurs, via la régularité de ces productions. Par exemple, produire systématiquement le même cri entre individus pour un stimulus menaçant donné permettrait à la convention d'émerger et après multiples répétitions, le cri seul suffirait à déclencher la réponse de fuite adéquate. Au contraire, si différents individus utilisaient différents cris pour un même stimulus, la convention serait compromise. La motivation peut ainsi sous-tendre la systématité.

Cependant, les relations motivées existent seulement via un *interprétant*, i.e. un être humain et son set de représentations cognitives. Afin que les individus se mettent d'accord entre eux, des représentations mentales communes sont nécessaires. Cela pose la question de la nature des correspondances trans-modales sous-tendant la relation entre signifiants et signifiés. A ce sujet, de nombreuses études expérimentales apportent des évidences à propos des associations partagées entre locuteurs de différentes langues et à différentes âges (voir section 1.2 et 2.5). La littérature explore aussi les mécanismes cognitifs en jeu dans les associations motivées, faisant partie d'un phénomène cognitif plus général, à savoir les correspondances trans-modales (voir section 3).

Ce préambule avait pour objectifs de clarifier la terminologie et de contextualiser la motivation par rapport à la théorie des signes, l'émergence du langage et la recherche plus générale en sciences cognitives. Par conséquent, l'introduction poursuivra avec une première section dans laquelle les différents phénomènes au sein des langues seront présentés, ainsi que les approches

expérimentales, plus précisément des études psycholinguistiques, mettant en évidence les associations chez les individus. La seconde section portera sur les hypothèses proposées au sujet des origines des différentes associations motivées, telles que l'origine gestuelle et le codage de la taille, ainsi que des approches émotionnelles. La troisième section présentera les mécanismes cognitifs possiblement impliqués. Enfin, la quatrième section exposera les limitations méthodologiques rencontrées dans la littérature et introduira les trois études conduites au cours de cette thèse ainsi que les objectifs visés.

1. Des preuves de motivations

a. Les manifestations dans les langues

Trois catégories de mots existent dans les langues et contredisent l'arbitrarité du signe : les onomatopées, les idéophones et les phonasthèmes.

L'onomatopée est un mot qui imite un son auquel il se réfère (Tanz, 1971) tel que les cris d'animaux. Par exemple, le cri du coq est « cocorico » en français, « kokekokkō » en japonais, « chicchirichi » en italien. Ces onomatopées se ressemblent et leur variabilité s'explique par des inventaires de sons différents entre les langues, ainsi que des règles phonotactiques différentes, etc.

Dans la seconde catégorie, l'idéophone a la particularité d'évoquer son référent et peut être défini comme un « mot sensoriel saisissant » (Dingemanse et al., 2015). Par exemple, en japonais « koro » est un objet léger qui roule une fois et « goro », un objet lourd qui roule une fois (Imai et al., 2008). Une différence de poids s'exprime donc par une différence de voisement. Les mots « korokoro » et « gorogoro » indiquent quant à eux la répétition de l'événement à travers la réduplication syllabique (bien que la réduplication soit chose courante dans les mimétiques de manière générale). Selon Iwasaki, Vinson et Vigliocco (2007), il existe un continuum de motivation dans les mimétiques : les *giseigo* ont une ressemblance directe, iconique, avec leurs référents, alors que les *gitaigo* ont une relation symbolique avec leurs référents⁵.

La troisième catégorie, le phonasthème, est une séquence de segments associée à une signification particulière dans une langue donnée. Par exemple, le cluster [gl] apparaît dans des mots relatifs à la lumière en anglais (e.g. glow, glitter, glare). Selon Dingemanse et al. (2015), ces distributions régulières spécifiques à une langue représentent de la *systematicité* plutôt que

⁵ D'autres types de mimétiques existent mais ne seront pas abordés dans cette thèse.

de l'iconicité qui, étant récurrente entre les langues, repose quant à elle sur des analogies perceptives.

La relation motivée, dans le cas des idéophones, n'est pas toujours évidente (contrairement aux onomatopées) et certains phonasthèmes peuvent sembler motivés. Une manière de déterminer la possible origine motivée d'un mot est de rechercher des similarités entre les langues ou locuteurs de langues. Par exemple, si différentes langues ou différents locuteurs distinguent la magnitude via des différences de voisement, cela serait en faveur d'une origine motivée pour les mimétiques « koro » et « goro ».

La partie suivante s'intéresse ainsi aux régularités au sein des langues ainsi qu'aux similitudes entre langues et locuteurs. Les deux objectifs sont les suivants : 1) exposer les similarités entre les langues et les régularités au sein des langues à travers des analyses linguistiques ; 2) évaluer expérimentalement les associations chez des locuteurs de différentes langues. Dans les deux cas, une variété de concepts sont abordés (taille, émotions...).

b. Approches écologiques

i. Analyses de lexiques

Différentes études se sont penchées sur les fréquences particulières de segments en fonction du champ sémantique (taille, distance, émotions, noms d'animaux ou divers concepts.)

Ohala (1997) a rapporté des évidences d'associations entre segments et concepts de taille dans différentes langues (ewe, yoruba, espagnol, grec, anglais, irlandais et français). Les mots exprimant une petite taille tendent à contenir des voyelles hautes antérieures comme [i] et des consonnes non-voisées comme [t] tandis que ceux exprimant une grande taille contiennent davantage de voyelles basses postérieures comme [u] et des consonnes voisées comme [b].

Tanz (1971) s'est intéressée quant à elle aux mots signifiant « ici » et « là » dans différentes langues et rapporte des éléments similaires à Ohala : davantage de [i] pour les courtes distances et [a] et [o] pour les plus grandes distances. Elle propose également des parallèles avec d'autres phénomènes tels que le temps (plus de [i] et [ɪ] dans les verbes anglais pour le présent que pour le passé) et la distance sociale (e.g. « anata » et « kimi » signifient tout deux « toi » en japonais, le premier étant plus formel et distant). Cela s'expliquerait par une analogie articulatoire : la prononciation de [i] est plus resserrée que [a], laissant moins d'espace entre la langue et le palais.

Haynie, Bower et LaPalombara (2014) se sont intéressés également aux mots exprimant les tailles et distance (« ici, là ; près, loin ; petit, grand ; fin, gros ») dans 120 langues australiennes et mettent en évidence également des patterns similaires aux précédentes études.

Fónagy (1961) a analysé le contenu segmental de poèmes qui avaient été évalués au préalable par des participants comme étant agressifs ou doux. Il a observé des fréquences particulières en fonction de ces jugements : les poèmes agressifs contiennent plus de consonnes non-voisées comme [k] et [t] et les doux davantage de sonantes comme [l] et [m]. Fónagy (1983) a aussi récolté des jugements à propos de segments et ces jugements révèlent que [i] est petit, agile, gentil et doux, alors que [u] est corpulent, obtus, triste, amer, fort et sombre, par exemple.

Adelman, Estes et Cossu (2018) ont analysé des centaines de mots de cinq langues (anglais, espagnol, néerlandais, allemand et polonais) sur leurs compositions phonémiques en fonction de leurs valences émotionnelles. Ils ont obtenu des effets significatifs pour toutes les langues ($p < .001$) avec des tailles d'effet variables en fonction des langues (1.40-4.32%). Une grande proportion de segments sont porteurs de valence (21 à 45% en fonction des langues), ce qui s'explique davantage par les catégories phonémiques que par des différences phonétiques. Par exemple, en anglais, [f] apparaît dans des mots positifs et [s] dans des mots négatifs : la valence ne s'explique donc pas par leur mode d'articulation commun (fricative). De plus, les valences des segments diffèrent d'une langue à l'autre (e.g. [d] est positif en espagnol et négatif en anglais).

Blasi, Hammarström, Stadler et Christiansen (2016) ont conduit une étude auprès de 62% des langues du monde, couvrant 85% des familles linguistiques. Ils ont recherché des fréquences d'occurrence de segments particulières dans les concepts de la liste de Swadesh (constituant de vocabulaire basique d'une langue). Ils ont observé des segments significativement susceptibles d'apparaître dans certains concepts (e.g. [l] pour « langue ») et d'autres significativement improbables (e.g. [k] pour « langue »). Cette étude démontre que certains concepts ont tendance à contenir ou non des segments spécifiques, à travers le monde et indépendamment des familles de langues et de la dispersion géographique.

Berlin (1994) s'est intéressé à la composition phonémique de mots désignant des oiseaux et des poissons en Huambisa – une langue parlée au Pérou. Il a observé des patterns de fréquences de segments particuliers : les noms d'oiseaux contiennent davantage de [i] et les noms de poissons davantage de [a]. Il y a aussi des différences en fonction de la position dans le mot : les syllabes finales des noms d'oiseaux contiennent davantage d'obstruents alors que celles des noms de poissons contiennent davantage de nasales et continues. Un autre élément est intéressant : il existe un pattern interne au sein de ces deux classes biologiques en fonction de la taille (e.g. les noms d'oiseaux de grande taille contiennent davantage de [a] et les noms de petits poissons contiennent davantage de [i]).

En plus de ces analyses de lexique, Berlin a conduit une expérimentation au cours de laquelle il a présenté à des étudiants ces noms d'oiseaux et de poissons par paire, chaque paire contenant un exemplaire de chaque (e.g. « chichikía » vs. « katan », le premier étant l'oiseau). Les participants devaient deviner lequel référait à un oiseau, sachant que le second correspondait à un poisson. La performance a atteint 54% de réussite, ce qui était significativement supérieur à la chance.

Ce dernier élément permet de démontrer qu'il est possible de deviner la signification d'un mot – du moins lorsqu'il est présenté dans une paire et en connaissant la signification du second mot – sur la base de sa composition segmentale. Nous pouvons nous demander quelle performance aurait été obtenue si les participants n'avaient pas eu connaissance de la signification du second mot, la présence d'un contraste pouvant sous-tendre le choix. Les sections suivantes présentent des études psycholinguistiques exploitant ce même type d'expérimentation consistant à présenter des mots étrangers ou pseudo-mots par paire.

ii. Expérimentations avec des mots et incidence du milieu linguistique ou culturel

Brown, Black et Horowitz (1955) ont conduit une étude auprès d'anglophone auxquels ils ont présenté des mots étrangers par paire, issus de trois langues différentes : le chinois, le tchèque et l'hindi. Les participants étaient capables de deviner de ces antonymes, dénotant des expériences sensibles (e.g. « froid » et « chaud »), qu'ils soient présentés uniquement à l'écrit, ou prononcés, et ce pour les trois langues. De la même manière, Kunihiro (1971) a obtenu des performances supérieures à la chance avec des antonymes japonais présentés par paire auprès d'américains, sous forme écrite (en alphabet latin), ou orale, avec ou sans intonation.

Imai et al. (2008) ont conduit une expérimentation transculturelle auprès de locuteurs anglais et japonais auxquels ils ont présenté des pseudo-mimétiques (créés sur la base des analyses de Hamano en 1998). Ces mimétiques ont été fabriqués de manière à véhiculer des informations sur deux dimensions (la vitesse et le poids) et à correspondre à des clips vidéo présentant des manières de marcher. Par exemple, « batobato » correspond à l'action de courir avec des pas lourds ([t] exprimant les coups et [b] le poids). La tâche était soit d'évaluer le matching entre le clip et le pseudo-mimétique à l'aide d'une échelle de 1 à 7 (jugement) ; soit de sélectionner une action parmi deux, correspondant à un pseudo-mimétique donné (choix forcé). Les jugements étaient similaires pour les anglais et japonais, bien que les tailles d'effet étaient plus élevées et les valeurs de p plus faibles chez ces derniers ($d = 0.60, p < .05$ vs. $d = 6.05, p < .001$, respectivement). Pour la tâche de choix forcé, les japonais ont constamment choisi le clip attendu (100%) et les anglais ont fait des choix dans le même sens mais dans une moindre

mesure (64%). Ces différences pourraient refléter des différences culturelles résultant en une plus grande sensibilité chez les japonais que chez les anglais, possiblement dû à la grande proportion de mots motivés existant en japonais.

Iwasaki et al. (2007) ont comparé des locuteurs de l'anglais et du japonais sur des jugements à propos de mimétiques giseigo (dénotant des rires) et gitaigo (dénotant des manières de marcher) sur différentes dimensions (e.g. gracieux *vs.* vulgaire ; bruyant *vs.* silencieux). Cette étude a mis en évidence des associations communes entre anglais et japonais, bien que ces derniers présentent davantage d'associations, plus complexes et plus robustes.

Les mots réels issus de différentes langues constituent un matériel écologique pour étudier les associations chez les individus. Cependant, ces mots présentent des limites puisqu'ils ne découlent pas seulement de la motivation mais aussi de l'arbitrarité, de règles phonotactiques, etc. L'utilisation de pseudo-mots permet de proposer de multiples combinaisons de segments. Un paradigme en particulier, appelé 'bouba-kiki' a beaucoup été exploité dans la littérature et a permis d'apporter des éléments au sujet des influences culturelles et linguistiques.

c. Bouba-kiki – un paradigme à la recherche d'un phénomène universel

Dans l'expérience originale de Köhler (1947) ce sont les pseudo-mots « maluma » et « takete » qui ont été utilisés, et qui furent par la suite supplantés par « bouba » et « kiki ». Le principe de cette expérimentation est de présenter deux pseudo-mots et deux formes visuelles, une arrondie et une pointue. Les participants avaient pour instruction d'associer l'un des deux pseudo-mots à l'une des deux formes. Köhler rapporte que les participants associaient « maluma » (ou « bouba ») à la forme arrondie et « takete » (ou « kiki ») à la forme pointue. Tandis que Ramachandran et Hubbard (2001) estiment que 95% de la population produisent les mêmes associations⁶, une méta-analyse plus récente estime une prévalence de 84-94% (Styles & Gawne, 2017).

Différentes études ont été réalisées et ont révélé des associations similaires chez 1) des enfants de Tanzanie parlant le kitongwe (une langue bantou) (Davis, 1961), 2) des adultes himbas (en Namibie) parlant l'otjiherero (Bremner et al., 2013). Tandis que les premiers, avaient connaissance de l'alphabet latin (via l'apprentissage du swahili), les seconds ne pouvaient pas être impactés par la forme des lettres.

Deux autres études ont au contraire révélé une absence d'associations entre segments et formes : 1) Rogers et Ross (1975) ont testé des locuteurs du hunjara en Papouasie-Nouvelle

⁶ Aucune précision n'est reportée quant à la population évaluée (âge, culture, nombre de personnes, langue parlée etc.) ou aux conditions expérimentales.

Guinée ; 2) Styles et Gawne (2017) ont testé des locuteurs du syuba au Népal. À noter que Styles et Gawne (2017) se sont penchés sur les langues parlées par ces deux communautés et ont ainsi proposé que l'absence d'effet pouvait s'expliquer par une violation des langues de ces locuteurs (e.g. [l] dans « maluma » n'existe pas en hunjara). Ceci permet de mettre en évidence un biais majeur dans la plupart des études (biais non limité à l'étude des associations motivées) : elles sont conduites par des expérimentateurs et auprès de participants qualifiés de « WEIRD » (Henrich, Heine, & Norenzayan, 2010), c'est-à-dire des occidentaux, éduqués, faisant partie de pays riches, industrialisés et démocratiques, ce qui représente une source de biais majeure. Dans le cas des études sur la motivation, les segments utilisés et contrastés sont peu nombreux ([p, b, m, t, d, n, k, g, i, e, a, o, u]) mais sont les plus fréquents dans les langues du monde, mais ne respectent pas nécessairement les langues parlées par les populations testées. Ainsi, on ne peut pas déterminer si ces associations n'existent pas réellement chez ces deux populations, ou si elles requièrent d'autres stimuli pour être mises en évidence.

Une autre étude conduite par Chen, Huang, Woods et Spence (2016) révèlent également des différences en fonction de la culture au niveau des traits visuels : les participants américains sont davantage influencés par l'amplitude des branches des formes pointues, tandis que les taiwanais sont davantage influencés par le caractère pointu. Ces différences peuvent refléter des traitements différents, plus analytiques chez les premiers et plus holistiques chez les seconds. Au final, cela démontre qu'il peut être nécessaire d'adapter les formes en fonction de la population évaluée, ce qui est source de variation dans les associations.

d. Les traits phonétiques et segments impliqués dans les associations

En parallèle aux études s'intéressant aux différents traits visuels, d'autres études se penchent sur l'impact de traits phonétiques ou segments dans les associations bouba-kiki. Tarte (1974) a conclu à un plus grand impact de la part des voyelles (en utilisant une petite sélection de consonnes et de voyelles), et des études plus systématiques penchent plutôt pour un effet supérieur des consonnes (e.g. Nielsen & Rendall, 2011). Fort, Martin et Peperkamp (2015) ont capturé un tableau plus subtil : les consonnes impactent significativement les associations, et l'effet des voyelles diffère en fonction du contexte consonantique dans lequel elles apparaissent. La supériorité des consonnes ne s'explique pas par la position initiale dans le mot, puisqu'ils ont testé à la fois des structures CVCV et VCV, ce qui n'apportait pas de variation. Cette étude a mis en évidence un continuum de groupes de segments dans les associations : les occlusives sont associées aux formes pointues et les sonantes aux formes arrondies, les fricatives se situant entre les deux. Via une tâche de jugement, Knoeferle, Li, Maggioni et Spence (2017) ont mis

en évidence un gradient encore plus précis. Du plus pointu au plus arrondi, les consonnes associées sont les suivantes : occlusives non-voisées > fricatives voisées > occlusives voisées > nasales > fricatives non-voisées > liquides > glides⁷.

Tandis que les segments se situant aux extrémités sont fortement associés à certaines formes, ceux du milieu peuvent être associés aux deux formes en fonction du contraste segmental. Par exemple, une occlusive voisée peut être associée à une forme pointue ou arrondie en fonction de l'autre segment avec lequel il est contrasté, une sonante ou une occlusive non-voisée, respectivement. De plus, ces résultats rapportés montrent qu'un trait n'est pas suffisant pour prédire les associations. Par exemple, l'effet du voisement diffère en fonction du mode : les occlusives voisées sont plus associées à une forme arrondie que les fricatives voisées (Knofler et al., 2017). Cependant, Cuskley, Simmer et Kirby (2015) ont proposé que ces associations s'expliquaient par les formes visuelles des lettres plutôt que par les propriétés acoustiques des segments. Par exemple, 'g' (une occlusive voisée) serait associée à une forme arrondie car c'est une lettre courbée, alors que 'z' (une fricative voisée) serait associé à une forme pointue.

La prochaine partie se penche sur les explications proposées au sujet de l'existence de ces associations motivées.

2. Les explications possibles à l'existence d'associations motivées

Cuskley et al. (2015) ont démontré que la forme des lettres pouvait influencer les associations mais elle ne peut pas expliquer à elle seule leur existence. En effet, elles sont mises en évidence chez des aveugles (Bottini, Barilari, & Collignon, 2019), ainsi que chez les himbas, comme précédemment présenté. Au contraire, il semblerait plutôt que ce soit la forme des lettres qui aurait été influencée par les segments puisqu'il est possible de deviner quel caractère contient le son [i] ou [u] lorsque différents caractères de différentes traditions écrites sont présentés par paires (Turoman & Styles, 2017).

Au sujet des associations en lien avec la taille ou la magnitude, Ohala (1984) propose une explication appelée *le codage de la taille* (size coding hypothesis) selon laquelle on peut communiquer des impressions de taille à travers la fréquence fondamentale (les sons de basse fréquence donnant une impression de grande taille). Ce codage serait à l'origine des associations entre segments et taille, telle que [i] avec « petit » et [a] avec « grand ».

D'un point de vue articulatoire, Ramachandran et Hubbard (2001) ont proposé de leur côté que le pattern d'association au sujet de la taille s'expliquerait par l'imitation articulatoire de

⁷ Un continuum au sujet de la taille était aussi rapporté, du plus grand au plus petit : sonantes > fricatives voisées et occlusives voisées > fricatives non-voisées et occlusives non-voisées.

différentes tailles : l'articulation de [i] mimerait un petit espace entre les doigts, par rapport à l'articulation de [a]. Dans la même veine, Sakamoto et Watanabe (2018) ont mis en évidence des associations particulières entre sensations tactiles et segments et font le lien avec des analogies perceptivo-motrices (e.g. les consonnes nasales bilabiales et alvéolaires impliquent des tissus mous, c'est pourquoi elles seraient associées à des sensations telles que « doux », « élastiques », « collant » et « humide »).

Chuenwattanapranithi, Xu, Thipakorn et Maneewongvatana (2008) ainsi que Adelman et al. (2018) ont étudié les relations entre traits acoustiques ou segments et émotions. D'un point de vue évolutionnaire, ces derniers proposent une théorie adaptative selon laquelle une pression de sélection aurait favorisé les individus présentant les meilleures habiletés pour communiquer, plus précisément, les plus efficaces en perception et production de segments, ce qui leur aurait permis de communiquer à propos d'émotions, et ainsi des dangers potentiels. Plus l'émotion négative est rapidement reçue, plus la réponse adéquate peut être rapidement réalisée. Cela expliquerait le pattern qu'ils ont observé dans les lexiques de cinq langues (anglais, espagnol, allemand, néerlandais et polonais) : les premiers segments des mots ont tendance à prédire la valence de ces mots. De plus, les analyses à propos de l'anglais et de l'allemand révèlent également que les segments les plus rapides à prononcer ont tendance à être en début de mots négatifs. Ainsi, ces auteurs proposent que la motivation aurait sous-tendu l'émergence du langage, du moins en partie, à travers la sélection naturelle.

Les études chez les enfants permettent d'évaluer le caractère possiblement inné des associations motivées. Plusieurs études se sont intéressées aux associations motivées présentes chez les enfants, impliquant les tailles (Peña, Mehler, & Nespors, 2011) et les formes (Imai et al., 2015; Maurer, Pathman, & Mondloch, 2006; Spector & Maurer, 2013). La distinction entre associations congruentes et incongrues impliquant des formes et des segments (la congruence étant basée sur les études antérieures) serait présente dès l'âge de 4 mois (Ozturk, Krehm, & Vouloumanos, 2013). Cependant, les interprétations diffèrent en fonction des paradigmes et il est difficile d'établir réellement une préférence pour les associations congruentes. De plus, un nouveau-né de 4 mois peut déjà avoir une exposition à sa langue maternelle suffisante pour commencer à démontrer des associations trans-modales. Cela n'est donc pas suffisant pour déterminer le caractère inné de ces associations qui peuvent au contraire découler d'un apprentissage statistique ou sémantique. Cependant, cela n'exclut pas l'existence de biais cognitif sous-tendant leur émergence, facilitant éventuellement leur apprentissage.

A ce propos, Nygaard, Cook et Namy (2009) a proposé un paradigme d'apprentissage de mots japonais à des adultes anglophones. Il y avait trois conditions différentes : 1) ils apprenaient la

réelle signification (e.g. « kurai » signifie « sombre ») ; 2) ils apprenaient son antonyme (e.g. « kurai » associé à « clair » ; 3) ils apprenaient une signification non liée (e.g. « kurai » associé à « humide »). La performance d'apprentissage était meilleure dans les deux premières conditions par rapport à la dernière. Ainsi, la composition phonémique semble être en lien avec le champ sémantique, plutôt qu'avec la réelle signification. Imai et al. (2008) et Kantartzis, Imai, & Kita (2011) ont démontré que les mots motivés présentaient un avantage par rapport aux mots arbitraires chez les enfants de trois ans i.e. les enfants généralisaient mieux la signification pour les mimétiques congruents. Nielsen et Rendall (2012) ont montré que l'apprentissage d'associations entre formes pointues et arrondies et composition phonémique était supérieur lorsqu'il y avait une congruence dans les appariements. Monaghan, Christiansen, et Fitneva (2011) ont approfondi les avantages respectifs de la systématique (et non de la motivation, mais la motivation peut sous-tendre la systématique) et de l'arbitrarité, et ont mis en évidence l'avantage de la systématique pour l'apprentissage de catégorie, et de l'arbitrarité pour l'apprentissage d'exemplaires particuliers.

L'avantage que présentent les associations motivées dans l'apprentissage de mots ou de catégories pose la question des mécanismes cognitifs le sous-tendant. La partie suivante porte sur la nature cognitive de ces associations.

3. Mécanismes cognitifs potentiellement impliqués

Les associations motivées semblent appartenir à une plus grande famille d'associations, les correspondances trans-modales (Spence, 2011), qui permettent de faire le lien entre les différentes modalités d'un même événement (par exemple, lier le bruit d'une balle de baseball que l'on frappe avec son événement visuel, etc.). Il existe trois types de correspondances transmodales :

- 1) Les structurelles : elles sont sous-tendues par l'organisation cérébrale, le fonctionnement des connexions synaptiques ou par des mécanismes cognitifs communs). Par exemple, l'intensité peut être codée de la même manière ce qui expliquerait l'association entre un volume sonore élevé et une forte luminosité (Marks, 2004).
- 2) Les statistiques : elles découlent de l'exposition à l'environnement et permettent une représentation cohérente de divers inputs sensoriels provenant de sources redondantes (Ernst, 2007). Par exemple, toucher un objet produit une représentation haptique qui active une estimation visuelle via une représentation conceptuelle de la taille.

- 3) Les sémantiques : elles résultent de l'exposition et de l'apprentissage du langage et permettent de lier différentes modalités qui sont décrites par des termes linguistiques communs (Marks, 2004). Par exemple, une fréquence fondamentale et une position dans l'espace peuvent toutes deux être décrites en termes d'élévation.

Une autre explication provient de (Ramachandran & Hubbard, 2001b) : les associations motivées découleraient du même mécanisme neuronal que la synesthésie. La synesthésie est un phénomène automatique et involontaire dont environ 4% de la population ferait l'expérience (Simner et al., 2006). Elle consiste à percevoir une sensation secondaire (*concurrent*) dans une modalité suite à la stimulation d'une autre ou d'une même modalité (*inducteur*). La synesthésie la plus courante consiste à voir des lettres ou chiffres en couleur. Plutôt qu'utiliser le terme de synesthésie, Nikolić (2009) propose le terme d'*idéesthésie* car l'inducteur serait un concept plutôt qu'un trait de bas niveau. En effet, d'après l'étude de Dixon, Smilek, Duffy, Zanna et Merikle (2003), la couleur perçue par un synesthète ne dépend pas de la forme elle-même mais du concept auquel il renvoie : un rond évoquera deux couleurs différentes selon qu'il est présenté en tant que chiffre (un zéro) ou en tant que lettre (un « o »). Réciproquement, une même syllabe japonaise évoque la même couleur qu'elle soit écrite en Hiragana ou en Katakana (Asano & Yokosawa, 2011).

Les études au sujet des correspondances trans-modales permettent d'apporter des informations pertinentes au sujet des associations motivées. Ces études ont mis en évidence que 1) les correspondances se modulent facilement et rapidement (e.g. Ernst, 2007), 2) on peut obtenir des effets facilitateurs (de la part d'association congruente) ou d'interférences (de la part d'association incongrue), ou les deux (e.g. Marks, Ben-Artzi, & Lakatos, 2003), ou l'inverse (Kovic, Plunkett, & Westermann, 2010), en fonction du paradigme et du matériel, 3) les associations seraient relatives plutôt qu'absolues i.e. elles dépendent du contexte (e.g. Marks, 1987), 4) il n'est pas évident de déterminer si elles existent à un haut-niveau ou à un bas-niveau de processus cognitifs mais elles apparaissent dans des protocoles implicites (e.g. Hung, Styles, & Hsieh, 2017).

4. Contribution méthodologique et expérimentale à l'étude des associations motivées

a. Discussion à propos des résultats antérieurs

La tâche bouba-kiki est un paradigme qui a permis de mettre en évidence des associations similaires chez des locuteurs de différentes langues à différents âges, permettant ainsi de formuler des hypothèses au sujet de l'émergence du langage aux niveaux ontogénique et

phylogénique. Cependant, la majorité des études sont conduites auprès d'occidentaux avec des pseudo-mots compatibles avec les langues parlées par les expérimentateurs. Les quelques divergences dans les résultats pointent vers la nécessité de conformer le matériel aux populations étudiées. De plus, ce contraste conceptuel largement utilisé, entre formes arrondies et pointues, est relativement distant par rapport aux hypothèses d'émergence et d'évolution du langage. Il est peu probable que les premières communications aient porté sur ce type de formes, ce qui complique la formulation d'hypothèses à ces sujets sur la base des résultats d'expérience bouba-kiki. Par ailleurs, les paradigmes bouba-kiki diffèrent largement les uns des autres en fonction des études et ce à plusieurs niveaux : 1) les contrastes phonétiques ou segmentaux utilisés ; 2) les populations étudiées (culture, langue, âge...) ; 3) le type de présentation des stimuli. C'est sur ce dernier point que le projet de thèse s'est attardé. Le paradigme original proposé par Köhler est de type 2x2 : deux formes et deux pseudo-mots étaient présentés aux participants qui devaient choisir une des deux formes à associer à l'un des deux pseudo-mots. Deux contrastes étaient donc présents à la fois. Il est aussi possible de présenter une paire de formes avec un seul pseudo-mot, le choix portant ainsi sur les formes en l'absence d'un contraste phonétique (e.g. Fort et al., 2015). A l'inverse, certaines études présentent une seule forme et deux pseudo-mots, le choix portant ainsi sur le contraste phonétique (e.g. Chen et al., 2016). Enfin, certains paradigmes présentent des appariements de forme et pseudo-mot, sans aucun contraste, tel que dans les tâches d'apprentissage (Kovic et al., 2010). Cependant, dans ce dernier cas de figure, les contrastes existent néanmoins dans la tâche entre les différents essais.

Il est difficile de démêler les impacts respectifs des différences expérimentales et peu d'études se sont intéressées à l'influence des paramètres méthodologiques. Une étude conduite par Aveyard (2012) a cependant fourni des indications à ce sujet. Deux conditions ont été proposées aux participants : une dans laquelle un pseudo-mot était présenté avec deux formes, une seconde dans lequel un pseudo-mot était présenté avec quatre formes (une forme cible, un distracteur de la même catégorie, deux distracteurs de l'autre catégorie). Dans la première condition, l'apprentissage incident était plus efficace pour les appariements congruents dès le premier block d'apprentissage – les appariements incongrus nécessitant davantage de temps, la progression s'effectuait au fil des blocks. Dans la seconde condition, les participants étaient plus lents et moins efficaces dans leur apprentissage et l'apprentissage apparaissant uniquement pour les appariements congruents. Cette étude montre que les paramètres expérimentaux peuvent impacter la mise en évidence d'associations.

De manière similaire, les paradigmes pourraient impliquer différents processus cognitifs en fonction du type de stimulus sur lequel porte le choix. En effet, choisir entre deux formes pour un pseudo-mot ou entre deux pseudo-mots pour une forme donnée devraient impliquer différents traitements cognitifs. De plus, une tâche d'association implique également différents processus par rapport à une tâche de jugement à propos d'un appariement.

b. Apport des présentes études

Puisque les formes de type bouba-kiki sont à distance de scénarii au sujet de l'émergence et de l'évolution du langage, des stimuli à propos d'émotions, de taille et d'animaux pourraient être plus pertinents. La première étude a ainsi utilisé des animaux en guise de matériel conceptuel à la place de formes, afin d'adopter une approche plus écologique. Les animaux ont permis d'évaluer différents contrastes : la dangerosité, la répulsion, la taille et la classe biologique (oiseaux *vs.* poissons). Nos hypothèses se sont basées sur différentes études à propos, 1) des émotions (Adelman et al., 2018 ; Fónagy, 1961, 1983), 2) de la taille (Berlin, 1994 ; Blasi et al., 2016 ; Haynie et al., 2014 ; Iwasaki et al., 2007 ; Knoeferle et al., 2017 ; Ohala, 1997 ; Sapir, 1929 ; Vainio, Tiainen, Tiippana, Rantala, & Vainio, 2017) et de la classe biologique (Berlin, 1994 ; Blasi et al., 2016).

La seconde étude s'est intéressée à l'impact de la variabilité des paradigmes expérimentaux – et plus précisément les différents types de présentation – dans la continuité de la précédente étude, i.e. avec les mêmes types de stimuli, au sujet d'animaux. Avec la comparaison de quatre types de présentation (2x2, 2x1, 1x2 et 1x1), cette étude visait à évaluer l'importance des contrastes. La population (des locuteurs du français) et les matériels conceptuel et phonétique étaient identiques à travers les quatre paradigmes, afin de clairement déterminer l'influence de la présence d'un contraste phonétique (1x2), conceptuel (2x1), des deux (2x2) et leur absence (1x1).

La troisième est basée sur l'étude de Westbury (2005) qui a proposé une tâche de décision lexicale à des participants. Les mots et pseudo-mots étaient présentés dans des cadres pointus et arrondis et les résultats ont mis en évidence une interaction entre la forme des cadres et la composition phonémique des pseudo-mots. Ainsi, les pseudo-mots composés d'occlusives non-voisées étaient plus rapidement catégorisés comme pseudo-mot lorsqu'ils étaient présentés dans un cadre pointu, et de manière similaire, les pseudo-mots composés de sonantes étaient plus rapidement catégorisés lorsqu'ils étaient présentés dans un cadre arrondi. Nous avons utilisé le même paradigme mais avons ajouté un autre paramètre : la police d'écriture des chaînes de lettres. Ceci permettait d'évaluer l'hypothèse de Cuskley et al. (2015) selon laquelle la forme

des lettres était à l'origine des associations. Ainsi, la police était soit angulaire soit courbée. Cette étude permettait donc d'évaluer à la fois l'existence de l'effet bouba-kiki avec un paradigme implicite de bas-niveau et la possible influence de la forme des lettres.

Les trois questions auxquelles cette thèse se proposait de répondre sont les suivantes :

- 1) Quelles associations motivées peuvent émerger à partir d'une approche écologique, en utilisant des concepts et des dimensions potentiellement pertinents pour des interactions primaires d'un point de vue évolutionnaire ?
- 2) Les différences de présentation influencent-elles la mise en évidence d'associations, et plus spécifiquement, dépendent-elles de la présence ou de l'absence de contrastes linguistiques ou conceptuels ?
- 3) Le priming supraliminal d'une forme influence-t-il le traitement d'un pseudo-mot en fonction de sa composition phonémique ou graphémique dans une tâche de décision lexicale ?

Expérimentations

Etude 1 : De Carolis, L., Marsico, E. & Coupé, C., 2017, Evolutionary roots of sound symbolism. Association tasks of animal properties with phonetic features, *Language & Communication*, 54, pp. 21-35

Résumé

En contradiction avec l'arbitrarité du signe de Saussure, le symbolisme sonore – l'association systématique et motivée de sons avec des significations – est fortement présent à travers les langues. Cela aurait pu permettre à nos ancêtres de développer un système de communication, qui aurait ensuite évolué vers des signes conventionnels. Nous avons testé des associations entre segments / traits phonétiques et une variété d'attributs à propos d'animaux (taille, dangerosité...) chez des participants francophones. Une première expérimentation – utilisant des images d'animaux – n'a révélé aucune association statistiquement significative, tandis qu'une seconde – utilisant des labels (e.g. « un animal petit » vs. « un animal grand ») – en a révélé plusieurs. Cette différence montre ainsi l'impact de différences méthodologiques et met en lumière que les photographies d'animaux compromettent la mise en évidence d'associations par leur multidimensionnalité. Une analyse au niveau du lexique des noms d'animaux en français a révélé que la présence de ces associations chez les participants ne s'explique pas par des fréquences particulières de ces segments ou traits phonétiques en fonction des attributs

d'animaux. Nous discutons ces résultats à la lumière des scénarii au sujet de l'évolution et de l'émergence du langage.

Mots clés : non-arbitrarité, multimodalité, symbolisme sonore, origine du langage, tâche d'association, noms d'animaux

Etude 2 : De Carolis, L., & Coupé, C., submitted to Cognition in May 2019, Phonetic and conceptual contrasts in the assessment of sound symbolic associations: comparing protocols and inferring cognitive processes

Résumé

Le paradigme le plus connu dans l'étude des associations motivées est 'bouba-kiki' qui consiste en l'association d'un pseudo-mot avec une forme visuelle, ronde ou pointue. Il existe de grandes variations à travers les expérimentations au niveau des paramètres expérimentaux et des populations étudiées telles que : la langue des participants (e.g., français, anglais, néerlandais), leur âge (adultes, enfants), la composition segmentale des pseudo-mots (de multiples voyelles et consonnes), les contrastes visuels ou conceptuels (majoritairement des formes, mais aussi des oiseaux et poissons dans l'étude de Berlin en 1994). La tâche à effectuer diffère aussi en fonction des études : un choix entre deux pseudo-mots et deux formes (2x2), un choix entre deux pseudo-mots pour une forme (1x2), un choix entre deux formes pour un pseudo-mot (2x1), et enfin un jugement (par exemple) entre un pseudo-mot et une forme (1x1).

Cette étude vise à évaluer l'influence de la présence ou de l'absence des deux types de contrastes, celui des segments et celui des concepts, en comparant quatre possibilités de présentation. Les concepts sont des attributs d'animaux – classe biologique (oiseaux vs. poissons), taille, dangerosité et répulsion – et dix hypothèses trouvées dans la littérature au sujet des associations motivées sont évaluées, e.g. l'association grand et [a] et entre petit et [i]. Les segments, les concepts et le profil des participants étaient identiques entre les quatre protocoles. Les résultats ont révélé des différences entre les protocoles, plus précisément différents nombres d'hypothèses statistiquement vérifiées et des tailles d'effet plus ou moins élevées. Chacune des dix hypothèses était confirmée au moins une fois à travers les quatre protocoles et aucune contradiction avec ces hypothèses n'est apparue. Certaines associations étaient systématiquement confirmées, telle que celle précédemment énoncée (taille et voyelles). Au niveau des tailles d'effet, la présence des deux contrastes (2x2) a conduit aux plus grandes valeurs de tailles d'effet, suivi du 1x2. Cela pourrait suggérer que le contraste segmental facilite la détection d'association motivée. Au global, les différences obtenues indiquent l'intérêt à étudier les processus cognitifs en jeu dans l'élaboration d'associations motivées.

Mots clés : symbolisme sonore, contrastes phonétique et conceptuel, tâches d'association et de jugement, méthodologie, protocoles

Etude 3 : De Carolis, L., Marsico, E., Arnaud, V. & Coupé, C., 2018, Assessing sound symbolism: Investigating phonetic forms, visual shapes and letter fonts in an implicit bouba-kiki experimental paradigm, PLoS ONE, 13:12, e0208874

Résumé

En général, dans les tâches d'associations bouba-kiki, il est demandé à un participant de déterminer la meilleure association entre une forme parmi deux – une arrondie et une pointue – et entre deux pseudo-mots – e.g. bouba et kiki. De nombreuses études rapportent que les formes pointues sont associées à kiki et les arrondies à bouba. Cette tâche est très répandue dans l'étude des associations motivées. Cependant, les tâches d'association sont très explicites dans le sens où elles mettent en avant le contraste visuel ou phonétique et la tâche à effectuer i.e. établir un lien trans-modal entre des stimuli de différentes natures. Des études récentes ont proposé comme explication que la ressemblance visuelle entre formes et lettres sous-tendait ces associations, du moins chez les sujets lettrés. Dans cette étude, nous présentons un paradigme plus implicite pour tester l'effet bouba-kiki, à savoir une tâche de décision lexicale avec des mots et pseudo-mots présentés dans des cadres arrondis et pointus. Les chaînes de lettres étaient présentées avec deux polices différentes – une angulaire et une courbée – afin d'évaluer l'impact d'éventuels biais graphémiques. Les temps de réponse ont été évalués avec des modèles GAMLSS qui offrent de nombreuses options en termes de distributions de probabilités. La distribution Gamma généralisée s'est révélée être la plus appropriée. Aucun effet d'associations motivées n'est apparu mais une interaction entre formes pointues et lettres angulaires est apparue, engendrant des temps de réponse plus rapides. Nous discutons ces résultats en rapport avec les effets de saillance visuelle de la part de formes pointues, le priming, des activations cérébrales, la synesthésie et l'idéesthésie.

Discussion

La première des études visait à évaluer que les animaux pouvaient représenter un matériel pertinent pour éliciter des associations motivées, en se basant sur l'idée selon laquelle ils avaient pu représenter des éléments importants pour des premiers échanges (en tant que potentielles menaces et sources de nourriture). Ils représentent ainsi un matériel écologique en comparaison aux formes arrondie et pointue (bouba-kiki) et une manière indirecte d'évaluer des reliquats des premières communications. Plusieurs contrôles ont été appliqués afin de limiter les sources de

variations parasites : 1) les photos étaient présentées en nuances de gris, 2) les autres modalités étaient équilibrées (e.g. un contraste de taille avec deux animaux équivalents en termes de répulsion, dangerosité et de même classe biologique), 3) les tailles de présentation respectaient les représentations mentales de ces tailles⁸. De plus, les modalités contrastées étaient présentées aléatoirement pour masquer dans une moindre mesure ce qui était étudié et pour rendre la tâche plus implicite. De nombreux segments ont été testés ([i, a, u, y, p, b, t, d, k, g, f, v, s, z, ʃ, ʒ, ʁ, m, n, l]) sous la forme VCVC avec une réduplication syllabique (e.g. [ipip]). Après correction pour tests multiples, aucune association n'est apparue significative. Deux explications possibles ont alors été proposées : soit le type de présentation (2x1) n'était pas approprié, soit les images d'animaux sont trop complexes pour mettre en avant un contraste en particulier.

Une seconde expérimentation a alors été mise en place afin de déterminer laquelle des deux explications était valide. Le même protocole (2x1) en format papier-crayon fut proposé à des participants mais en utilisant des labels à la place des images (e.g. « un animal petit »). Dans ces circonstances, sept hypothèses parmi neuf furent significatives après correction. Ainsi, nous avons pu conclure que ce type de présentation permet bien de mettre en évidence des associations et ce sont probablement les images qui étaient trop complexes, malgré les contrôles appliqués. Plus généralement, cette étude a permis de démontrer que les caractéristiques d'animaux pouvaient être pertinentes dans l'étude des associations motivées, et permettaient d'étudier des aspects émotionnels en plus de la taille et des classes biologiques.

La seconde étude était axée sur l'impact des différences méthodologiques et plus particulièrement les différentes présentations possibles des stimuli (2x2, 1x2, 2x1, 1x1), tout en conservant le même matériel et la même population. Dans la lignée de la seconde expérimentation de l'étude précédente, des labels d'animaux ont été utilisés. Dix associations ont été testées avec les mêmes contrastes conceptuels que précédemment (classe biologique, taille, dangerosité et répulsion). Afin de pallier les problèmes méthodologiques de la seconde expérience de l'étude précédente, des ajustements ont été adoptés : 1) à la place d'une tâche papier-crayon, la tâche était informatisée, permettant le contrôle des temps de présentation et de réponse, de randomiser les essais afin de modestement « masquer » l'objet de l'étude et de mesurer les temps de réponse ; 2) les pseudo-mots étaient présentés à l'oral plutôt qu'à l'écrit ; 3) les analyses portaient sur les voyelles, les consonnes ainsi que l'interaction entre les deux. Les résultats semblent indiquer une supériorité d'un contraste segmental (plus d'associations

⁸ Les derniers points ont été rendus possibles par des jugements collectés au préalable.

significatives et des tailles d'effet plus importantes en 2x2 et 1x2). Cependant, cette conclusion est à modérer : dans le cas du 1x2, le contraste conceptuel est en quelque sorte présent (« un animal petit » évoque inévitablement son opposé, « un animal grand »). Ainsi, la « supériorité » du 1x2 par rapport au 1x1 et au 2x1 s'explique peut-être par la présence des deux contrastes, ce qui le rend similaire au 2x2. Quant à la différence entre le 2x2 et le 1x2 – le 2x2 étant plus « efficace » que le 1x2 – elle s'expliquerait par le choix forcé pour un label dans le cas du 1x2. On observe également des différences dans les associations par rapport à l'étude précédente, en particulier au sujet des oiseaux et poissons, ce qui pourrait indiquer l'impact de l'écrit par rapport à la présentation orale. En effet, davantage d'associations apparaissaient dans l'étude précédente, ces associations rejoignaient ce que Berlin (1994) a obtenu, alors qu'il présentait également ce stimuli à l'écrit. Cela peut ainsi poser la question des mécanismes sous-jacents et de l'impact de la modalité de présentation et plus précisément l'impact de la forme écrite.

L'étude suivante est de type 1x1 puisqu'elle consiste en une tâche de décision lexicale avec des stimuli présentés à l'écrit. Elle permet de tester l'effet bouba-kiki de manière implicite (en accord avec la critique émise par Nielsen et Rendall (2012), selon laquelle les effets sont surestimés à cause de la transparence des tâches). Elle permet également d'évaluer l'impact de la forme écrite, en accord avec l'explication proposée par Cuskley et al. (2015).

Cette troisième étude est une extension de celle conduite par Westbury (2005) qui a également proposé une tâche de décision lexicale avec des mots et pseudo-mots qui différaient sur leurs compositions phonémiques. Il a ainsi obtenu des temps de réponse plus rapide pour les pseudo-mots composés d'occlusives non-voisées présentés dans des cadres pointus et pour les pseudo-mots composés de sonantes dans des cadres arrondis. Nous avons également testé ceci (ainsi que des pseudo-mots composés d'occlusives voisées) en ajoutant comme paramètre la police d'écriture : les chaînes de lettres étaient soit angulaires (*Agency FB*) soit courbées (*Gabriola*). Nous avons appliqué de nombreux contrôles sur les pseudo-mots et mots (nombre de lettres et de phonèmes, nombre de voisins, structures syllabiques, etc.)

Dans de telles conditions nous n'avons pas répliqué les résultats obtenus par Westbury (2005). Au contraire, nos résultats semblent davantage être en accord avec l'hypothèse de Cuskley et al. (2015) étant donné que nous avons obtenu une interaction entre police d'écriture et cadres. Cependant, les effets simples révèlent un effet facilitateur dans un seul cas : lorsque les pseudo-mots sont présentés avec une police angulaire dans un cadre pointu (et pas avec une police courbée dans un cadre arrondi). Cela relèverait potentiellement d'un effet de saillance visuelle. En effet les lignes perpendiculaires sont détectées plus rapidement que des traits arrondis

(Coelho, Cloete, & Wallis, 2010) et nous pouvons faire le lien avec l'hypothèse du recyclage neuronal proposé par Dehaene (2005) : l'aire spécialisée dans la forme visuelle des mots (VWFA) aurait été recyclée pour cette fonction car elle avait auparavant une fonction similaire – le traitement des jonctions et intersections tel que c'est le cas pour l'aire analogue chez les primates (Dehaene & Cohen, 2007). Ainsi, les études sur les associations motivées devraient tenir compte de ces effets de bas niveau qui peuvent influencer les résultats.

Conclusion

Cette thèse apporte différentes contributions méthodologiques dans l'étude des associations motivées. Premièrement, les caractéristiques d'animaux et leurs classes biologiques peuvent éliciter des associations et permettent ainsi l'évaluation d'aspects émotionnels, par exemple, et permettent d'évaluer indirectement des questions liées à l'émergence et à l'évolution du langage. Deuxièmement, la seconde étude a permis de mettre en évidence les différences que pouvaient induire les différents types de présentation et plus précisément l'impact de la présence et de l'absence des deux types de contrastes (segmental et conceptuel). Ainsi, le choix du type de présentation devrait être considéré avec attention. Troisièmement, la dernière étude a démontré que des biais perceptuels existent pour les formes – du moins les pointues – ce qui pouvait provoquer des effets de saillance visuelle et de priming. Ainsi, ces éléments devraient être pris en compte pour les futures études au sujet de bouba-kiki.

Il serait pertinent d'approfondir davantage la relation entre les associations motivées et 1) les correspondances trans-modales, 2) la synesthésie et l'idéesthésie (e.g. Cuskley, Dingemanse, Kirby, & Leeuwen, 2019), 3) les théories de la relativité linguistique et la cognition incarnée (e.g. Šetić & Domijan (2007)). Ces cadres théoriques ne s'excluent pas mutuellement mais représentent plutôt des approches complémentaires à l'étude des associations motivées. Ils permettent, au contraire, de comprendre davantage les mécanismes cognitifs permettant de d'unifier les expériences dans leur multimodalité, et d'investiguer indirectement l'évolution et l'émergence du langage. D'autres approches sont également complémentaires : 1) l'étude des troubles du langage (e.g. Meteyard, Stoppard, Snudden, Cappa, & Vigliocco, 2015), 2) l'imagerie cérébrale (e.g. Kanero, Imai, Okuda, Okada, & Matsuda, 2014), 3) l'étude des lexiques (Johansson & Carling, 2015).

Bien que l'arbitrarité soit une caractéristique fondamentale du langage, il y a de fortes preuves de motivation autant dans les lexiques que chez les individus. Et nous espérons avoir pu contribuer ainsi à mettre en lumière une facette du fonctionnement cérébral de notre espèce et de son évolution.

Références bibliographiques

- Adelman, J. S., Estes, Z., & Cossu, M. (2018). Emotional sound symbolism: Languages rapidly signal valence via phonemes. *Cognition*, *175*, 122–130.
- Asano, M., & Yokosawa, K. (2011). Synesthetic colors are elicited by sound quality in Japanese synesthetes. *Consciousness and Cognition*, *20*(4), 1816–1823. <https://doi.org/10.1016/j.concog.2011.05.012>
- Aveyard, M. E. (2012). Some consonants sound curvy: effects of sound symbolism on object recognition. *Memory & Cognition*, *40*(1), 83–92. <https://doi.org/10.3758/s13421-011-0139-3>
- Berlin, B. (1994). Evidence for pervasive synaesthetic sound symbolism is ethnozoological nomenclature. In *Sound Symbolism* (pp. 77–93).
- Blasi, D. E., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound – meaning association biases evidenced across thousands of languages. <https://doi.org/10.1073/pnas.1605782113>
- Bottini, R., Barilari, M., & Collignon, O. (2019). Sound symbolism in sighted and blind. The role of vision and orthography in sound-shape correspondences. *Cognition*, *185*(February), 62–70. <https://doi.org/10.1016/j.cognition.2019.01.006>
- Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). “Bouba” and “Kiki” in Namibia? A remote culture make similar shape-sound matches, but different shape-taste matches to Westerners. *Cognition*, *126*(2), 165–172. <https://doi.org/10.1016/j.cognition.2012.09.007>
- Brown, R. N., Black, A. H., & Horowitz, A. E. (1955). Phonetic symbolism in natural languages. *Journal of Abnormal and Social Psychology*, *50*, 388–393.
- Chen, Y. C., Huang, P. C., Woods, A., & Spence, C. (2016). When “bouba” equals “kiki”: Cultural commonalities and cultural differences in sound-shape correspondences. *Scientific Reports*, *6*(December 2015), 1–9. <https://doi.org/10.1038/srep26681>
- Chuenwattanapranithi, S., Xu, Y., Thipakorn, B., & Maneewongvatana, S. (2008). Encoding Emotions in Speech with the Size Code. A Perceptual Investigation. *Phonetica*, *65*(4), 210–230.
- Coelho, C. M., Cloete, S., & Wallis, G. (2010). The face-in-the-crowd effect: When angry faces are just cross(es). *Journal of Vision*, *10*(1), 1–14.
- Cuskley, C., Dingemans, M., Kirby, S., & Leeuwen, T. (2019). Cross-modal associations and synesthesia : Categorical perception and structure in vowel – color mappings in a large

- online sample. *Behavior Research Methods*, 1–25.
<https://doi.org/https://doi.org/10.3758/s13428-019-01203-7>
- Cuskley, C., Simmer, J., & Kirby, S. (2015). Phonological and orthographic influences in the bouba-kiki effect. *Psychological Research*. [https://doi.org/DOI 10.1007/s00426-015-0709-2](https://doi.org/DOI%2010.1007/s00426-015-0709-2)
- Davis, R. (1961). The fitness of names to drawings. A cross-cultural study in Tanganyika. *British Journal of Psychology*, 52(3), 259–268. <https://doi.org/10.1111/j.2044-8295.1961.tb00788.x>
- Dehaene, S. (2005). Evolution of human cortical circuits for reading and arithmetic: The ‘neuronal recycling’ hypothesis. In S. Dehaene, J.-R. Duhamel, M. Hauser, & G. Rizzolatti (Eds.), *From Monkey Brain to Human Brain*. Cambridge: MIT Press.
- Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2), 384–398. <https://doi.org/10.1016/j.neuron.2007.10.004>
- Dingemans, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Systematicity in Language. *Trends in Cognitive Sciences*, 19(10), 603–615. <https://doi.org/10.1016/j.tics.2015.07.013>
- Dixon, M. J., Smilek, D., Duffy, P. L., Zanna, M. P., & Merikle, P. M. (2003). The role of meaning in grapheme-colour synaesthesia. *Cortex*, 42, 243–252.
- Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, 7(5), 1–14. <https://doi.org/10.1167/7.5.7.Introduction>
- Fónagy, I. (1961). Communication in poetry. *Word*, 17, 194–218.
- Fónagy, I. (1983). *La vive voix. Essais de psycho-phonétique*. Paris: Payot.
- Fort, M., Martin, A., & Peperkamp, S. (2015). Consonants are More Important than Vowels in the Bouba-kiki Effect. *Language and Speech*, 58(2), 247–266. <https://doi.org/10.1177/0023830914534951>
- Gasser, M., Sethuraman, N., & Hockema, S. (2011). Iconicity in Expressives : An Empirical Investigation. *Empirical and Experimental Methods in Cognitive/Functional Research*, (July), 163–180.
- Gomi, T. (1989). *An illustrated dictionary of Japanese onomatopoeic expressions*. The Japan Times.
- Hamano, S. (1998). *The sound-symbolic system of Japanese*. Stanford, CA & Tokyo: CSLI & Kuroshio Publisher.
- Haynie, H., Bower, C., & LaPalombara, H. (2014). Sound Symbolism in the Languages of Australia. *PLoS ONE*, 9(4), e92852. <https://doi.org/10.1371/journal.pone.0092852>

- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 61–135. <https://doi.org/10.1017/S0140525X0999152X>
- Hung, S. M., Styles, S. J., & Hsieh, P. J. (2017). Can a Word Sound Like a Shape Before You Have Seen It? Sound-Shape Mapping Prior to Conscious Awareness. *Psychological Science*, 28(3), 263–275. <https://doi.org/10.1177/0956797616677313>
- Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, 109(1), 54–65. <https://doi.org/10.1016/j.cognition.2008.07.015>
- Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., & Kita, S. (2015). Sound symbolism facilitates word learning in 14-month-olds. *PLoS ONE*, 10(2). <https://doi.org/10.1371/journal.pone.0116494>
- Iwasaki, N., Vinson, D. P., & Vigliocco, G. (2007). What do English Speakers Know about gera-gera and yota-yota?: A Cross-linguistic Investigation of Mimetic Words for Laughing and Walking. *Japanese-Language Education around the Globe*, 17(6), 53–78.
- Johansson, N., & Carling, G. (2015). The de-iconization and rebuilding of iconicity in spatial deixis: An Indo-European case study. *Acta Linguistica Hafniensia*, 47(1), 4–32. <https://doi.org/10.1080/03740463.2015.1006830>
- Kanero, J., Imai, M., Okuda, J., Okada, H., & Matsuda, T. (2014). How sound symbolism is processed in the brain: A study on Japanese mimetic words. *PLoS ONE*, 9(5), 1–8. <https://doi.org/10.1371/journal.pone.0097905>
- Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese sound-symbolism facilitates word learning in English-speaking children. *Cognitive Science*, 35(3), 575–586. <https://doi.org/10.1111/j.1551-6709.2010.01169.x>
- Knoeferle, K., Li, J., Maggioni, E., & Spence, C. (2017). What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Scientific Reports*, 7(1), 1–11. <https://doi.org/10.1038/s41598-017-05965-y>
- Köhler, W. (1947). *Gestalt Psychology (2nd Edition)*. New York City: Liveright.
- Kovic, V., Plunkett, K., & Westermann, G. (2010). The shape of words in the brain. *Cognition*, 114(1), 19–28. <https://doi.org/10.1016/j.cognition.2009.08.016>
- Kunihira, S. (1971). Effects of the Expressive Voice on Phonetic Symbolism. *Journal of Verbal Learning and Verbal Behavior*, 10, 427–429. [https://doi.org/10.1016/S0022-5371\(71\)80042-7](https://doi.org/10.1016/S0022-5371(71)80042-7)
- Lockwood, G., & Dingemans, M. (2015). Iconicity in the lab: a review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Frontiers in*

- Psychology*, 6(August), 1–14. <https://doi.org/10.3389/fpsyg.2015.01624>
- Marks, L. E. (1987). On Cross-Modal Similarity: Auditory-Visual Interactions in Speeded Discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, 13(3), 384–394. <https://doi.org/10.1037/0096-1523.13.3.384>
- Marks, L. E. (2004). Cross-modal interactions in speeded classification. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The Handbook of Multisensory Processes*. (pp. 85–105). MIT Press. <https://doi.org/10.1063/1.3246835>
- Marks, L. E., Ben-Artzi, E., & Lakatos, S. (2003). Cross-modal interactions in auditory and visual discrimination. *International Journal of Psychophysiology*, 50, 125–145. <https://doi.org/10.1016/S0167-8760>
- Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: sound-shape correspondences in toddlers and adults. *Developmental Science*, 9(3), 316–322. <https://doi.org/10.1111/j.1467-7687.2006.00495.x>
- Meteyard, L., Stoppard, E., Snudden, D., Cappa, S. F., & Vigliocco, G. (2015). When semantics aids phonology: A processing advantage for iconic word forms in aphasia. *Neuropsychologia*, 76, 264–275. <https://doi.org/10.1016/j.neuropsychologia.2015.01.042>
- Monaghan, P., Christiansen, M. H., & Fitneva, S. A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, 140(3), 325–347. <https://doi.org/10.1037/a0022924>
- Nielsen, A., & Rendall, D. (2011). The sound of round: evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Expérimentale*, 65(2), 115–124. <https://doi.org/10.1037/a0022268>
- Nielsen, A., & Rendall, D. (2012). The source and magnitude of sound-symbolic biases in processing artificial word material and their implications for language learning and transmission. *Language and Cognition*, 4(2012), 115–125. <https://doi.org/10.1515/langcog-2012-0007>
- Nikolić, D. (2009). Is Synaesthesia Actually Ideesthesia? An Inquiry Into the Nature of the Phenomenon. In *Proceedings of the Third International Congress on Synaesthesia, Science & Art*. Granada, Spain.
- Nuckolls, J. B. (1999). The Case for Sound Symbolism. *Annual Reviews of Anthropology*, 28(1999), 225–252. <https://doi.org/10.1146/annurev.anthro.28.1.225>
- Nygaard, L. C., Cook, A. E., & Namy, L. L. (2009). Sound to meaning correspondences facilitate word learning, *I12*, 181–186. <https://doi.org/10.1016/j.cognition.2009.04.001>

- Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica*, 41, 1–16.
- Ohala, J. J. (1997). Sound Symbolism. *4th Seoul International Conference on Linguistics [SICOL]*, (Ching 1982), 98–103. <https://doi.org/10.1525/jlin.1996.6.1.109>
- Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: evidence for sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*, 114(2), 173–186. <https://doi.org/10.1016/j.jecp.2012.05.004>
- Peirce, C. S. (1931). *Elements of Logic. The Collected Papers of Charles Sanders Peirce, vol. 2.* (C. Hartstone & P. Weiss, Eds.). Cambridge, Mass.: Belknap Press of Harvard University Press. <https://doi.org/10.1038/135131a0>
- Peña, M., Mehler, J., & Nespore, M. (2011). The Role of Audiovisual Processing in Early Conceptual Development. *Psychological Science*, 22(11), 1419–1421. <https://doi.org/10.1177/0956797611421791>
- Ramachandran, V. S., & Hubbard, E. M. (2001a). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings. Biological Sciences / The Royal Society*, 268(1470), 979–983. <https://doi.org/10.1098/rspb.2001.1576>
- Ramachandran, V. S., & Hubbard, E. M. (2001b). Synaesthesia — A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, 8(12), 3–34.
- Rogers, S. K., & Ross, A. S. (1975). A cross-cultural test of the Maluma-Takete phenomenon. *Perception*, 4, 105–106.
- Sakamoto, M., & Watanabe, J. (2018). Bouba/Kiki in touch: Associations between tactile perceptual qualities and Japanese phonemes. *Frontiers in Psychology*, 9, 1–12. <https://doi.org/10.3389/fpsyg.2018.00295>
- Sapir, B. Y. E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, 12(3), 225–239.
- Saussure, F. de. (1916). *Cours de linguistique générale.* (C. Bally & A. Sechehaye, Eds.). Paris: Payot.
- Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes*, 22(2), 297–312. <https://doi.org/10.1080/01690960600732430>
- Simner, J., Mulvanna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., ... Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, 35(8), 1024–1033. <https://doi.org/10.1068/p5469>
- Spector, F., & Maurer, D. (2013). Early sound symbolism for vowel sounds. *I-Perception*, 4(4),

239–241. <https://doi.org/10.1068/i0535>

- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, and Psychophysics*, 73(4), 971–995. <https://doi.org/10.3758/s13414-010-0073-7>
- Styles, S. J., & Gawne, L. (2017). When Does Maluma/Takete Fail? Two Key Failures and a Meta-Analysis Suggest That Phonology and Phonotactics Matter. *I-Perception*, 8(4), 1–17. <https://doi.org/10.1177/2041669517724807>
- Tallerman, M. (2011). Protolanguage. In M. Tallerman & K. R. Gibson (Eds.), *The Oxford Handbook of Language Evolution*. Oxford: Oxford University Press.
- Tanz, C. (1971). Sound symbolism in words relating to proximity and distance. *Language and Speech*, 14(3), 266–276.
- Tarte, R. D. (1974). Phonetic symbolism in adult native speakers of czech. *Language and Speech*, 17(1), 87–94. <https://doi.org/10.1177/002383097401700109>
- Turoman, N., & Styles, S. J. (2017). Glyph guessing for ‘oo’ and ‘ee’: Spatial frequency information in sound symbolic matching for ancient and unfamiliar scripts. *Royal Society Open Science*, 4(9). <https://doi.org/10.1098/rsos.170882>
- Vainio, L., Tiainen, M., Tiippana, K., Rantala, A., & Vainio, M. (2017). Sharp and round shapes of seen objects have distinct influences on vowel and consonant articulation. *Psychological Research*, 81(4), 827–839. <https://doi.org/10.1007/s00426-016-0778-x>
- Westbury, C. (2005). Implicit sound symbolism in lexical access: evidence from an interference task. *Brain and Language*, 93(1), 10–19. <https://doi.org/10.1016/j.bandl.2004.07.006>