



HAL
open science

Written code-switching in the notes of second-language learners in bilingual classroom environments

Olga Grebeshkova

► **To cite this version:**

Olga Grebeshkova. Written code-switching in the notes of second-language learners in bilingual classroom environments. Linguistics. Université de Lyon, 2017. English. NNT : 2017LYSE2022 . tel-01628221

HAL Id: tel-01628221

<https://theses.hal.science/tel-01628221>

Submitted on 3 Nov 2017

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.



UNIVERSITÉ
LUMIÈRE
LYON 2

N° d'ordre NNT : 2017LYSE2022

THESE de DOCTORAT DE L'UNIVERSITÉ DE LYON

Opérée au sein de

L'UNIVERSITÉ LUMIÈRE LYON 2

École Doctorale : ED 484 Lettres, Langues, Linguistique et Arts

Discipline : Études anglophones

Soutenue publiquement le 21 janvier 2017, par :

Olga GREBESHKOVA

**Written code-switching in the notes of second-language learners in bilingual classroom environments /
L'alternance codique dans les notes des apprenants de langue seconde en milieu bilingue universitaire.**

Devant le jury composé de :

Peter GRIGGS, Professeur des universités, Université Lumière Lyon 2, Président

Svetlana KOROLKOVA, Professeure Associée HDR, Université d'état de Volgograd,
Rapporteuse

Alexandra BAGASHEVA, Professeure Associée HDR, Université de Sofia, Rapporteuse

Jim WALKER, Maître de conférences, Université Lumière Lyon 2, Directeur de thèse

Contrat de diffusion

Ce document est diffusé sous le contrat *Creative Commons* « [Paternité – pas d'utilisation commerciale - pas de modification](#) » : vous êtes libre de le reproduire, de le distribuer et de le communiquer au public à condition d'en mentionner le nom de l'auteur et de ne pas le modifier, le transformer, l'adapter ni l'utiliser à des fins commerciales.

Résumé

Introduction

La globalisation, le réseau internet qui donne accès à une multitude de textes multilingues, les exigences croissantes du monde de travail et parfois des raisons d'ordre plus personnel font que les uns et les autres se trouvent de plus en plus souvent au cœur d'un environnement multilingue (que l'on soit immigré, un professionnel qui travaille à l'étranger, un étudiant effectuant un séjour de mobilité, un enfant de couple mixte, etc.). Par la même occasion, nous devons de manière plus régulière nous rapprocher les uns des autres et nous adapter à des conditions changeantes, ce qui explique une explosion dans le nombre de personnes qui apprennent une seconde langue, par exemple.

L'alternance codique est un phénomène linguistique qui est fréquent, voire omniprésent, dans un environnement, comme celui que l'on se définit, où au moins deux langues sont en contact. Des personnes bi-/multilingues peuvent produire de multiples alternances codiques à l'intérieur d'une phrase, comme entre deux phrases ou deux tours de conversations. Ce phénomène est de mieux en mieux compris, et de nombreux modèles et approches ont été proposés pour rendre compte de ce phénomène d'alternance codique dans le discours.

L'une des préoccupations des chercheurs en sociolinguistique en générale, et en l'étude de l'alternance codique en particulier, est la manière dont les données bilingues sont collectées. Selon Labov, la linguistique doit chercher à étudier la langue parlée dans des conditions « naturelles » (Labov 1972, 205). Pour cette raison, beaucoup d'études ont été axées sur l'alternance codique dans la communication orale de la vie réelle. Par conséquent, ces différents modèles, basés sur des considérations grammaticales, phonologiques et sociales forts divergentes, ont été testés sur des dizaines de paires de langues, parfois très proches, typologiquement parlant, et parfois très éloignées. Il y a eu un véritable foisonnement d'études, au cours des années.

Mais l'alternance codique n'est pas qu'un phénomène oral. Elle existe bel et bien dans des formes écrites. Avec le progrès technologique, beaucoup de données multilingues écrites sont devenues accessibles pour les chercheurs sur Internet, ce qui a donné lieu à des études sur l'alternance codique dans la correspondance par courriel, les blogs, les journaux en ligne, etc. En outre, l'alternance codique occupe une place spéciale dans les œuvres littéraires, ce qui n'a pas manqué d'attirer

également l'attention de nombreux chercheurs. Différente à la fois de l'alternance codique dans les communications orales et écrites « naturelles », l'alternance codique littéraire a seulement récemment gagné une forme de reconnaissance.

Les travaux dont cette thèse rend compte ont pris leur source dans une observation intéressante que je fis pendant une conférence sur la psycholinguistique, me permettant de m'ouvrir à un champ d'études qui est, pour l'instant en tout cas, très sous-étudié : celui des données bilingues écrites « naturelles ». Celui qui donnait la conférence était un universitaire invité d'origine espagnole. Sa langue maternelle était l'espagnol, le diaporama qui servait de base à sa présentation était en anglais, mais comme nous étions en France et comme il avait déjà salué le public en français, il a décidé de poursuivre son discours en français. Même si le conférencier parlait français, parfois on entendait des phrases ou des mots dans sa langue maternelle. Par exemple, à un moment donné, quand il expliquait quelque chose, tout à coup l'écran avec le diaporama est devenu noir. Sa réaction spontanée a été : « Qu'est-ce qui se passe aquí? ». Au cours de sa présentation, les auditeurs, même ceux qui étaient unilingues, ont été immergés dans un milieu bilingue ou même trilingue. Et c'est là que j'observai une chose intéressante : les notes que je prenais étaient en anglais, français et russe. Je me demandais alors si la nature multilingue de la situation dans laquelle je me trouvais à ce moment-là pouvait expliquer ma propre alternance codique à l'écrit, et surtout pourquoi j'alternais ainsi, alors même que les notes que je prenais n'auraient pas d'autre lecteur que moi-même, à savoir l'auteur même.

Les notes bi/multilingues prises au cours d'une conférence représentent un type de pratique écrite qui diffère à la fois des textes littéraires et de la communication écrite. L'exemple ci-dessus démontre que les locuteurs bi/multilingues utilisent plusieurs langues non seulement pour communiquer des informations à quelqu'un, mais aussi pour noter quelques informations pour eux-mêmes. Une telle activité d'écriture peut être appelée écriture personnelle, parce qu'elle n'implique pas un lecteur externe, mais uniquement son auteur. Ainsi, en ce qui concerne l'écriture personnelle, je commençais à remarquer que les locuteurs bilingues appliquent l'alternance codique quand ils font des listes de tâches ou écrivent des journaux intimes, mes étudiants utilisent plusieurs langues pendant la prise de notes en classe de langue, j'écris en plusieurs langues en préparant mes cours. Ainsi, l'écriture personnelle semble être une riche source de données bilingues. Quelque répandue que soit cette activité d'écriture parmi les bilingues, l'alternance codique dans l'écriture personnelle reste nettement sous-étudiée. Le seul travail sur le journal intime bilingue a été réalisé par Montes-Alcalá (2000), où l'auteur a examiné l'alternance codique dans son propre journal personnel.

Une fois mon attention tournée vers l'écriture personnelle bilingue, je visais à recueillir un corpus qui pourrait offrir la possibilité de faire une analyse qualitative et quantitative de l'alternance codique. Pour plusieurs raisons (dont il sera question dans la deuxième partie de ce travail), les brouillons d'examen des étudiants ont été sélectionnés pour composer trois sous-corpus, à savoir des textes produits par des apprenants français, russes et bulgares en langue seconde. Par conséquent, cette thèse présentera les recherches sur l'alternance codique dans l'écriture personnelle académique et donnera une description multimodale de ce phénomène. Elle s'attèlera également à situer l'alternance codique dans l'écriture personnelle académique sur l'arène bilingue.

La première partie de cette thèse donne un aperçu approfondi de l'état de la littérature dans les domaines idoines, à savoir le bilinguisme, l'alternance codique et la prise de notes. Le chapitre 1 est ainsi consacré à la question du bilinguisme. En plus de passer en revue les définitions et la terminologie employée dans la littérature sur la question, ce chapitre décrit plusieurs typologies de bilinguisme pour plus tard situer les participants du corpus sur le continuum du bilinguisme.

Partie I

Chapitre 1. Bilinguisme

Le bilinguisme est un vaste champ qui fait appel à divers domaines d'études, tels que les neurosciences et de la psychologie des bilingues, l'intelligence artificielle, des études en contact entre les langues, l'éducation bilingue, etc. Le bilinguisme en tant que tel dévoile des questions et des polémiques intéressantes. Ce chapitre va tenter de répondre aux questions suivantes : Qui est une personne bilingue ? Comment les bilingues peuvent différer les uns des autres ? Quels sont les différents niveaux de bilinguisme ?

1.1. Définition

Donnant une image assez générale d'une personne bilingue, les définitions les plus souvent rencontrées dans la littérature s'associe généralement à l'un ou à l'autre des deux pôles d'un spectre, à savoir les maximalistes et les minimalistes. Cette tendance persiste dans des définitions plus détaillées d'un locuteur bilingue. Une interprétation maximaliste pourrait être trouvée dans le dictionnaire Webster de

1961 selon laquelle une personne bilingue est décrite comme « ayant ou utilisant deux langues avec l'aisance caractéristique d'un locuteur natif ; une personne qui utilise deux langues en particulier habituellement et avec une maîtrise semblable à celle d'un locuteur natif »¹ (Hamers et Blanc 1989: 6). L'entrée du dictionnaire Webster (<http://www.webster-dictionary.org/definition/bilingual>) d'aujourd'hui (2015) fait référence à un bilingue comme « une personne qui parle deux langues couramment »². Au pôle opposé John Macnamara propose une définition minimaliste en utilisant le terme « bilingue » pour « les personnes qui possèdent au moins une des compétences linguistiques [parler, écrire, écouter ou lire], même à un degré minimal dans leur langue seconde »³ (Macnamara 1967 : 59–60). Toutes ces définitions manquent de précision.

Les deux Mackey (1962/2000) et Baetens Beardsmore (1982) évoquent la notion de relativisme, montrant l'impossibilité de fournir une définition universelle du bilinguisme. Ils donnent au moins deux raisons qui représente le relativisme.

Tout d'abord, ces définitions ne suggèrent pas la limitation stricte à un certain nombre de langues concernées. Le bilinguisme, et cela sans égard à l'étymologie du terme, devrait être considéré comme un terme qui embrasse tout, y compris les cas de multi/plurilinguisme. « Nous allons donc considérer le bilinguisme comme l'utilisation alternée de deux ou plusieurs langues par le même individu »⁴ (Mackey 1962/2000: 27).

Deuxièmement, étant donné que le bilinguisme est perçu comme un concept relatif, il représente un continuum (Hamers et Blanc, 1989). Les bilingues peuvent varier dans leur compétence ; quel est le degré de la maîtrise la langue de la personne ? Les bilingues peuvent utiliser les langues pour des fonctions différentes ; dans quel but la personne utilise-t-elle la langue ? Les bilingues peuvent connaître les mêmes langues, mais les raisons et le détail des commutations peuvent différer ; chaque cas d'alternance codique est différent.

Certains chercheurs (Weinreich 1953/1970; Mackey 1962/2000; Baetens Beardsmore 1982; Hamers et Blanc 1989) préfèrent travailler avec des typologies ou les dimensions d'un phénomène, qui déterminent les limites d'un domaine particulier

¹ La traduction est de moi, comme celles qui suivent. La citation originale est “having or using two languages especially as spoken with the fluency characteristic of a native speaker; a person using two languages especially habitually and with control like that of a native speaker”

² “a person who speaks two languages fluently”.

³ “persons who possess at least one of the language skills [speaking, writing, listening or reading] even to a minimal degree in their second language”

⁴ “We shall therefore consider bilingualism as the alternate use of two or more languages by the same individual”

de recherche et qui se concentrent sur un type particulier de phénomène étudié, plutôt que de se contenter de définitions. La question des typologies est l'objet de la section suivante.

1.2. Typologies

1.2.1. La typologie de Hugo Baetens Beardsmore

Une des typologies les plus largement reconnues est celle proposée par Hugo Baetens Beardsmore. La typologie de Baetens Beardsmore est une classification très riche. Il rassemble un grand nombre de différents types de bilinguisme qui sont basés sur des facteurs tels que le temps d'acquisition du langage (précoce ou tardive, le bilinguisme consécutif ou successif), l'environnement de l'acquisition du langage (bilinguisme naturel ou primaire et secondaire), les compétences de la personne (gradation de ambilingualisme, équilibrisme, le bilinguisme « non-fluent » à semilinguisme, réceptif et productif), organisation mentale (subordonné, composé et coordonné), l'effet de L2 sur L1 (additif et soustractif), l'utilisation ou la mise en œuvre du bilinguisme (ascendant et récessif), etc. (Pour plus de détails, voir Baetens Beardsmore 1982).

Cependant, cette typologie a quelques faiblesses. Tout d'abord, il n'y a pas de structure interne. Le classement représente une sorte de liste où différents types de bilinguisme sont présentés plus ou moins au hasard. Certains types de bilinguisme sont des sous-catégories des autres, et en même temps, ils sont liés à l'une des catégories suivantes ; le lecteur est livré à lui-même pour faire ces connexions. Il n'y a pas de division des critères selon lesquels l'auteur distingue différents types de bilinguisme. Cette forme de représentation rend la compréhension de la typologie plus difficile, surtout en tenant compte du nombre des types de bilinguisme représentés. Deuxièmement, certaines explications sont dépassées par des études actuelles et la situation désormais mondialisée de contacts linguistiques.

1.2.2. La typologie de Hamers et Blanc

Dès le début, ces auteurs établissent une distinction entre bilingualité et le bilinguisme. « La bilingualité est l'état psychologique d'une personne qui a accès à plus d'un code linguistique comme moyen de communication sociale »⁵ (Hamers et Blanc 1989: 6). Le bilinguisme, quant à lui, est présenté comme un concept plus

⁵ La traduction faite par l'auteur de la thèse. La citation originale est "Bilinguality is the psychological state of an individual who has access to more than one linguistic code as a means of social communication"

général qui inclut celle de bilinguisme, « mais se réfère également à l'état d'une communauté linguistique dans laquelle deux langues sont en contact, avec pour résultat que les deux codes peuvent être utilisés dans la même interaction et qu'un certain nombre d'individus sont bilingues »⁶ (ibid). La bilinguisme et le bilinguisme tels qu'ils sont définis par Hamers et Blanc correspondent au bilinguisme individuel et sociétal, respectivement. Cependant, les auteurs vont plus loin en fournissant des classifications de chaque phénomène.

1.2.3. Conclusion pour typologies

Les typologies mentionnées ci-dessus décrivent le phénomène du bilinguisme d'une manière tout à fait explicite. Alors qu'il est plutôt difficile de déterminer les contours du bilinguisme par une simple définition, les typologies parviennent à expliquer les différents aspects et facettes de ce phénomène. Ils montrent qu'il n'y a jamais une situation univoque ; chaque cas de bilinguisme peut avoir ses caractéristiques particulières en raison de circonstances survenues dans une communauté ou une expérience bilingue personnelle. Bien sûr, ces deux approches ne sont ni exclusives ni unique.

Cependant, cette thèse n'a pas l'intention d'examiner toutes les définitions et typologies de bilinguisme ; au contraire, elle vise à donner une vue d'ensemble du phénomène du bilinguisme avec ses facettes multidimensionnelles et problématiques en même temps. Les grandes lignes du bilinguisme présentées dans ce chapitre mettent l'accent sur l'importance du relativisme ; à la fois le bilinguisme individuel et sociétal ou bilinguisme et le bilinguisme représentent un continuum. Chaque personne bilingue vit une expérience différente et fait face à un environnement unique où elle se familiarise avec la deuxième, troisième, etc. langue. Et chaque individu trouve sa place dans le continuum de bilinguisme/bilinguisme. Les arguments fournis dans ce chapitre permettent donc d'estimer que les apprenants de la langue seconde, qui sont les sujets de la recherche doctorale présente, sont bilingues et bilingues. Cette question sera examinée dans la deuxième partie, chapitre 4, qui donne plus de détails sur les participants dans l'étude.

⁶ La traduction faite par l'auteur de la thèse. La citation originale est "but refers equally to the state of a linguistic community in which two languages are in contact with the result that two codes can be used in the same interaction and that a number of individuals are bilingual"

Ce chapitre étudie également la question de la terminologie dans le domaine du bilinguisme, en examinant les termes en usage comme « langue maternelle », « première langue », « langue seconde », « langue dominante »⁷, etc.

Le chapitre 2, quant à lui, examine le phénomène de l'alternance codique à la fois dans ses formes orales et écrites. Section 2.1. donne un aperçu historique de l'alternance codique parlé ainsi que le cadre théorique, qui comprend des modèles grammaticaux, psycholinguistique et sociales de l'alternance codique parlé, et le reste du chapitre explore d'autres facettes du phénomène.

Chapitre 2. Alternance codique

2.1.1. Études fondamentales

2.1.1.1. Aperçu historique de l'alternance codique

Le phénomène de l'alternance codique est connu depuis le début du XXe siècle. En 1913, le linguiste français Jules Ronjat publie ses observations de son fils bilingue - *Le développement du langage observé chez un enfant bilingue*. Ce sont les premières données enregistrées reconnaissables sur le bilinguisme. Cependant, l'essor des études sur l'alternance codique, en linguistique, est assez récent. Au début du XXe siècle, ce sont les psychologues qui ont montré le plus grand intérêt dans le bilinguisme, tentant de décrire le lien entre le bilinguisme et les compétences mentales et les défauts des locuteurs. Ensuite, une place importante dans la recherche sur l'alternance codique est occupée par Uriel Weinreich et son œuvre *Languages in Contact* (1953). Il s'y intéresse aux effets des contacts linguistiques sur les langues. Il décrit trois types de bilinguisme : coordonné, composé et subordonné. « La description par Weinreich des codes de commutation suggère que les individus bilingues possèdent deux variétés linguistiques distinctes, qu'ils emploient (idéalement) à des occasions distinctes »⁸ (Nilep 2006: 5). L'utilisation fréquente de l'alternance codique fut considérée comme un défaut, à comparer aux productions du « bilingue idéal ».

Dans les premiers travaux, l'alternance codique fut caractérisée comme aléatoire, mais maintenant beaucoup de linguistes en viennent à la conclusion qu'il s'agit d'un phénomène régi par des règles ; cependant, il existe encore des divergences sur la nature des contraintes. Dans la seconde moitié du XXe siècle, les

⁷ "mother tongue", "first language", "second language", "dominant language", etc.

⁸ La traduction faite par l'auteur de la thèse. La citation originale est "Weinreich's description of switching codes suggested that bilingual individuals possess two separate linguistic varieties, which (ideally) they employ on separate occasions"

perspectives sur l'alternance codique changent. Une des œuvres influentes sur l'alternance codique parut en 1972 par Blom et Gumperz. Ils décrivent l'utilisation du bokmål et du ranamål à Hemnesberget, montrant l'importance des facteurs sociaux sur l'utilisation, et plus spécialement la commutation entre ces deux dialectes norvégiens. Blom et Gumperz introduisirent deux types de l'alternance codique ; *métaphorique*, lorsque les locuteurs changent de langue pour produire des effets communicatifs sur l'auditeur, et l'alternance codique *situationnelle*, quand il y a un changement d'une langue en raison d'un changement dans le contexte social.

En 1982 Gumperz développa un autre concept, celui de « l'alternance codique conversationnelle » qui comprend plusieurs fonctions conversationnelles telles que la citation, la spécification du destinataire, l'interjection, la répétition, la qualification du message, et la personnification vs. l'objectivation. Sa recherche inspira beaucoup de linguistes qui ont exploré et développé ses idées par la suite.

Une autre contribution importante à l'étude de l'alternance codique en termes de motivations sociales fut celle de Myers-Scotton, qui a construit sur l'alternance codique situationnelle et métaphorique de Gumperz son propre modèle. En 1983, elle a publié le livre *The negotiation of identities in conversation: a theory of markedness and code choice*, où elle propose le Markedness Model. Selon Myers-Scotton, les locuteurs choisissent les langues qu'ils utilisent pour indiquer leurs rôles sociaux, qu'elle appelle un ensemble de droits et d'obligations. Depuis 1970, il y a eu de nombreuses études démontrant diverses approches, modèles et contraintes de l'alternance codique. La section 2.1.2. fournira une présentation plus complète des principales théories de l'alternance codique.

2.1.1.2. Terminologie et définition

Ce chapitre examine également la question de la terminologie en explorant la différence dans l'orthographe du terme « code-switching »⁹, son origine et sa distinction de « borrowing », « interference » and « code-mixing ».

2.1.2. Cadre théorique.

Cette section esquisse un cadre théorique avec les aspects grammaticaux, psycholinguistiques et sociaux de l'alternance codique, illustrant différentes contraintes grammaticales : le modèle du Matrix Language Frame élaboré par Myers-Scotton, le principe psycholinguistique de « triggering » introduit par

⁹ Ici le terme anglais "code-switching" est employé à la place de son équivalent français "alternance codique" pour éviter la confusion et être explicite concernant le contenu de la section.

Michael Clyne, le « we-code » et « they-code » de Gumperz, la distinction entre l'alternance codique situationnelle et métaphorique faite par Gumperz et son collègue Jan-Petter Blom, le modèle de « Markedness » proposé par Carol Myers-Scotton.

2.2. L'alternance codique dans l'écriture

Alors que nous avons tendance à associer l'alternance codique à la seule langue parlée, la section 2.2. se focalise sur l'alternance codique dans l'écriture et démontre les domaines dans lesquels ces mêmes concepts ont été mis en œuvre par la recherche précédente, tels que la production des enfants bilingues et les apprenants d'une langue seconde, la littérature et la communication électronique.

Lorsque l'intérêt pour l'alternance codique écrite surgit, la plupart des chercheurs ont essayé d'appliquer des méthodes développées pour l'alternance codique parlée à leurs corpus écrits, en somme pour voir si les tendances que l'on observe dans l'alternance codique écrite reflètent celles de l'orale. Toutefois, si certaines pratiques d'écriture telles que la correspondance, les scripts de jeu, les communications en ligne, etc. pourraient sans doute être facilement étudiées avec les mêmes outils que l'alternance codique orale, d'autres types d'écriture qui ne sont pas du type communicationnel, ou en tout cas d'un type de communication dans laquelle l'identité de l'interlocuteur est inconnue ou multiple - par exemple, dans la publicité, les affiches, les panneaux signalétiques, etc. - nécessitent une autre approche mettant en évidence la valeur de ces genres littéraires particuliers. Le cadre multimodal proposé par Mark Sebba (2012) est le seul modèle élaboré spécifiquement pour les textes multilingues en tenant compte de leur diversité.

Quand on parle de l'alternance codique écrite, on a tendance à oublier le côté multidisciplinaire de ce phénomène qui intéresse non seulement la sociolinguistique, mais aussi les sciences de l'éducation et la psychologie. Ce chapitre met en lumière les différentes perspectives de l'alternance codique écrite, mettant l'accent sur les cinq thèmes suivants : la lecture/écriture des enfants bilingues, l'écriture en L2, l'alternance codique dans la littérature et l'écriture électronique, et enfin, le chapitre explique la seule approche existante pour l'analyse des textes multilingues.

Chapitre 3. Notes

Chapitre 3 traite de la notion de notes. Le processus de prise de note est examiné comme une activité cognitive.

3.1. La prise de note

La prise de notes occupe une place importante dans la vie des étudiants. Voilà pourquoi elle est d'un si grand intérêt pour les éducateurs. Chaque pays a son propre système éducatif avec ses propres caractéristiques. En tenant compte de ces différences, les chercheurs de différents milieux universitaires et de différents systèmes éducatifs ont pu montrer l'influence qu'ils peuvent avoir sur des habitudes et des pratiques de prise de notes. Le processus de la prise de notes peut être défini entre autres facteurs comme une activité cognitive, ce qui implique l'enregistrement sous une forme écrite, provenant d'une source orale ou écrite, des informations que les preneurs de notes estiment importantes (Piolat 2006: 7). Le mot « cognitive » joue un rôle important dans cette définition. L'effort cognitif utilisé dans la prise de note ne doit pas être sous-estimé. Cette activité ne consiste pas seulement à écrire ce qui est dit ou écrit. « La prise de notes est une activité complexe qui exige la compréhension et la sélection des informations et qui inclut des processus de production écrites »¹⁰ (Piolat et al. 2005: 291). Les preneurs de notes doivent sélectionner des informations tout en écoutant, car ils ne disposent pas de temps à perdre avant et pendant la notation des informations.

3.2. Critique génétique

Après avoir regardé la prise de notes documentaires (à partir d'un texte lu ou discours écouté), au début de ce chapitre, l'accent se déplace vers un autre type de notes personnelles et qui est un deuxième « compagnon » des étudiants, à savoir les brouillons. Une proposition intéressante a été faite en 1972 par Jean Bellemin-Noël, qui a inventé le terme « avant-texte » :

l'ensemble constitué par les brouillons, les manuscrits, les épreuves, les « variantes », vu sous l'angle de ce qui précède matériellement un ouvrage quand celui-ci est traité comme un texte, et qui peut faire système avec lui (Bellemin-Noël cité dans Grésillon 1994: 108).

L'approche qui étudie de telles séries de documents est connue sous le nom de la critique génétique. Elle n'a rien de commun avec la biologie et de la génétique, au contraire, elle traite de la genèse des œuvres littéraires. L'objet de la critique génétique est un texte comme étant en « mouvement » (Grésillon 2008: 23). Jean-Louis Lebrave (1983) explique la dualité dans les positions que les lecteurs peuvent adopter à l'égard des avant-textes. Pour la majorité des gens, faire face à une copie

¹⁰ La traduction est de moi. La citation originale est "Note-taking is a complex activity that requires comprehension and selection of information and written production processes"

signifie de la lire. Cependant, quand on commence à réaliser cette activité habituelle, on comprend qu'elle est tout à fait différente de la lecture d'un texte ; c'est au cours de cette pseudo-lecture que l'on génère des informations génétiques quasi-spontanément conduisant à des hypothèses sur la manière dont le texte a été produit. Dans le même temps, nous sommes tous des producteurs de brouillons ; cette position d'écrivain nous fait sentir une appartenance énigmatique aux processus génétiques. Ainsi, Lebrave a proposé un modèle d'analyse de la production de textes basé sur ces deux aspects : être à la fois lecteur et écrivain. Fondamentalement, la lecture est une tentative d'organiser un avant-texte dans une double perspective. Tout d'abord, on lit un texte contenant des corrections et des substitutions de l'auteur qu'on a l'intention de mettre dans un ordre chronologique ; ce processus peut être compliqué par un grand nombre de corrections, comme par exemple les avant-textes des auteurs tels que Heine, Proust ou Flaubert, qui sont composés presque exclusivement de ratures. Deuxièmement, on essaye d'établir un ordre évaluatif dans un avant-texte ; allant au-delà de la conventionalité culturelle de la perfection du texte, le système évaluatif et en même temps interprétatif pourrait être une tentative d'expliquer à nous-mêmes pourquoi il y a tant de corrections dans les avant-textes. Ayant procédé du point de vue de la perspective d'un écrivain à celle d'un lecteur, Lebrave met en évidence la différence entre production orale et écrite ; à savoir qu'il est toujours possible de revenir et de rectifier, de supprimer ou d'enrichir, etc. le texte qui est déjà écrit, tandis que dans un discours oral, on peut seulement insérer un commentaire méta-linguistique (« non », « Je suis désolé, je me suis mal exprimé », « Je vais me corriger », etc.). Le fait qui détermine cette différence et l'écriture en général, est que « l'écrit produit quelque chose qui se trouve, non pas dans le temps, mais dans l'espace, et qui est un objet matériel manipulable à deux dimensions » (Lebrave 1983 : 18). Les deux dimensions, que Lebrave mentionne, sont un niveau textuel et graphique d'analyse. Lebrave explique que le niveau textuel fournit une sorte de matériel pour la lecture dans son sens général ; ce niveau correspond à la description des analyses des avant-textes de Grésillon, qui peuvent comprendre la langue (paraphraser, représentation du temps, de la syntaxe, le discours, etc.), la narration, les thèmes, la psychanalyse (l'inconscient du texte) et la contribution de l'auteur vs la contribution de l'époque historique (Grésillon 1994). En ce qui concerne le niveau graphique, il approche des avant-textes des arts visuels avec l'orientation sur l'activité d'écriture de son auteur et non pas sur le jugement d'un brouillon par des lecteurs. L'écriture est considérée comme une activité visuelle, car elle est guidée par le regard, et aussi comme une activité spatiale parce qu'un auteur organise un texte, autrement dit un tracé graphique, dans un espace (Olive et al., 2010: 35) .

3.3. Une étude avec un double corpus

Un dossier génétique assez original a été recueilli et examiné par Annie Piolat et Marie-Laure Barbier (2007). Elles ont réussi à regrouper deux types d'avant-texte : la prise de notes documentaires et des brouillons rédactionnels, qui précèdent le processus d'écriture. Avec cette étude, Piolat et Barbier font remarquer un nouveau champ d'écriture personnelle, qui est riche dans sa diversité des matériaux et qui est encore peu étudié. La direction ouverte par ces auteurs pourrait être très prolifique pour d'autres sciences telles que les sciences de l'éducation, la psychologie cognitive, linguistique, psycholinguistique, etc.

3.4. La prise de note des interprètes

Admettant l'importance de la prise de notes des interprètes dans un domaine général de la prise de notes, la section 3.4. fournit une brève description de cette question.

Après avoir examiné la notion de notes, ses formes et ses contextes ce chapitre finalise la première partie de ce travail et apporte une contribution essentielle à la compréhension complète des brouillons des étudiants, qui constituent le corpus de cette thèse.

Partie II

A la suite de l'état de l'art, qui a été représenté ci-dessus, la partie II se concentrera sur l'étude originale menée dans le cadre de cette thèse de doctorat. Elle examinera l'écriture bilingue, notamment l'alternance codique dans l'écriture personnelle académique. Comme le montre la partie I, le phénomène de l'alternance codique a été bien étudié depuis le début du XXe siècle. Cependant, l'intersection de l'écriture bilingue et la critique génétique n'a jamais fait jamais l'objet de l'intérêt qu'elle mérite. Est-ce simplement un domaine de recherche, qui attend son temps ? J'adhère à cette idée. Les multiples raisons du choix du thème de cette étude, qui seront présentées au chapitre 4, vont avancer des arguments en faveur de son intérêt et son importance.

Le premier chapitre de la partie II explore les raisons du choix du sujet et la méthodologie employée, aussi bien dans les premiers stades de la collecte des

données, les décisions prises quant à la transcription du corpus et la présentation des résultats.

Chapitre 4. Questions méthodologiques

4.1. Le choix du sujet

Comme nous venons de le voir, la majorité de la recherche jusqu'à présent a été axée sur l'alternance codique dans les communications orales et écrites, ce qui implique au moins deux participants : un locuteur/écrivain et un auditeur/lecteur. Très peu de travaux ont porté sur ce qu'on appelle l'écriture personnelle (Montes-Alcalá 2000). La spécificité de cette forme linguistique est qu'il n'y a qu'une seule personne impliquée, parce que l'écrivain joue le rôle du lecteur ; l'écriture personnelle ressemble à un monologue avec soi-même.

C'est un terrain potentiellement riche pour la recherche. Ce type d'écriture est inhérent à chaque individu lettré. C'est une sorte d'un espace personnel où les auteurs se sentent libres de s'exprimer sans aucune instruction ou jugement, qu'il soit moral ou formel. Cet espace présente un grand intérêt pour les linguistes. L'écriture personnelle peut jouer un rôle d'intermédiaire entre le fonctionnement du cerveau et la production finale, ce qui se reflète dans des exemples tels que les autocorrections qui sont présentes et acceptables dans l'écriture personnelle et tous les types d'ébauche qu'on fait en tant qu'esquisse d'une œuvre finale ou d'un plus grand projet. Ainsi, l'écriture personnelle des individus bilingues peut apporter une contribution précieuse à la psycholinguistique, fournissant des données tangibles et naturalistes pour l'analyse de la production bilingue. En outre, les individus forment des communautés et des sociétés. Dans le même temps, ces structures sociales exercent une influence sur les individus, influence qui est incorporée dans les normes sociales et morales, la culture, la politique, etc. Par conséquent, l'individu reflète la communauté et/ou la société, à laquelle il appartient, dans la même manière que l'écriture personnelle reflète l'individu. En conséquence, les questions sociolinguistiques sont nombreuses : Combien l'écriture personnelle est-elle représentative de la communauté/société ? Quelles sont les variables sociales qui affectent l'écriture personnelle ? Est-ce que l'alternance codique dans l'écriture personnelle bilingue varie avec différentes paires de langues ? Aussi paradoxal que cela puisse paraître au premier abord, l'écriture personnelle, comme un sous-type de l'écriture en général, peut être analysée du point de vue sociolinguistique. « [...] [L] écriture représente une forme particulière de la performance linguistique qui est inévitablement influencée par les conditions sociolinguistiques de sa production, de sorte que les textes de toute nature constituent également un type de source

linguistique qui peut être d'intérêt en soi, vue de la perspective sociolinguistique »¹¹ (Schneider 2013: 170).

L'écriture personnelle bilingue représente tout à fait une source novatrice pour les études bilingues. On doit définir l'écriture personnelle pour permettre son application à la suite de la recherche en matière de bilinguisme. Le présent travail se référera, dorénavant, à l'écriture personnelle comme à des éléments écrits ou des séquences d'éléments écrits produits par un auteur et qui, initialement, ne sont pas destinés à des lecteurs externes.

4.2. *Corpus*

La nouveauté du travail sur l'alternance codique dans l'écriture personnelle implique la difficulté d'un corpus inexistant. En ce qui concerne tous les types possibles de l'écriture personnelle qui auraient pu être recueillis, les brouillons d'examen des étudiants ont été considérés comme les meilleures données disponibles pour atteindre les objectifs de ce travail.

Les brouillons d'examen ont été recueillis dans trois départements d'anglais respectifs : Université Lumière Lyon 2 (France), Université d'État de Volgograd (Russie) et Université St. Kliment Ohridski de Sofia (Bulgarie). Les étudiants français étaient soit en 3ème ou en 4ème année de leurs études. Les participants russes étaient des étudiants soit en 4ème ou en 5ème année. En ce qui concerne les données bulgares, elles diffèrent en deux aspects : d'abord, seulement 30% des participants étaient en 4ème année d'études, tandis que 70% étaient en leur 1ère année; d'autre part, seulement 44,5% des données sont constituées de brouillons, tandis que les 55,5% restants sont des copies au propre. En dépit du fait que la disparité des données bulgares va compliquer la comparaison entre les trois groupes, ces données méritent d'être étudiées toutes seules.

Étant un adepte, comme nous l'avons vu, du point de vue que le bilinguisme est un concept relatif, qui représente un continuum (Mackey 1962/2000; Baetens Beardsmore 1982; Hamers et Blanc 1989) – une notion qui fut examinée dans la Partie I – je considère les apprenants de la langue seconde, en leur 3ème, 4ème et 5ème année d'études dans le département d'anglais de leurs universités respectives, comme des bilingues aux fins de la présente étude.

¹¹La traduction est de moi. La citation originale est "[...] [W]riting represents a special form of linguistic performance that is unavoidably shaped by sociolinguistic conditions of its production, so texts of whatever kind also constitute a type of linguistic source that may be of interest in its own right from a sociolinguistic perspective"

Toutes les copies ont été recueillies après les examens, qui étaient basés sur un stimulus langagier. On entend par là, d'abord, que la langue de la réponse d'examen était imposée : les français et les étudiants bulgares devaient répondre en anglais, tandis que les étudiants russes avaient certaines tâches à rédiger en anglais et certaines en russe. Deuxièmement, les feuilles contenant les questions d'examen et les textes sources étaient en la langue en laquelle la réponse devait être donnée. Plus de détails sur les examens sont fournis dans la description des données pour chaque groupe linguistique. Le montant total des données collectées est 632 de copies (572 brouillons et 60 copies au propre). Sur ce total, seuls ont été gardés pour l'analyse à proprement parler les brouillons avec la présence de deux ou plusieurs langues, soit 167 brouillons et 12 copies au propre.

Toutes les données ont été collectées soit en format original, manuscrit, soit ont été photographiées si le document original ne pouvait pas être déplacé. Toutes les copies ont fait l'objet d'une lecture minutieuse. Cependant, celles qui étaient entièrement monolingues : anglais, français ou russe, ont été rejetées, et seulement celles dans lesquels plus d'une langue était présente, transcrites. La transcription est un outil d'analyse, ce qui permet un transfert d'une écriture manuscrite vers un espace numérique de travail, c'est-à-dire sur un écran et dans un programme de traitement de texte. Le système de transcription adopté dans cette recherche a cherché à reproduire, le plus fidèlement possible, les copies originales conservant des éléments grapho-spatiaux de l'auteur, les passages raturés, la grammaire et l'orthographe d'origine, etc. Si l'alternance codique apparaissait fréquemment dans la copie, elle était transcrite entièrement. S'il n'y avait que quelques occurrences d'alternance codique, seule l'alternance codique dans son contexte était transcrite.

4.3. Méthodologie

Outre ce travail sur une forme de corpus jusqu'alors inexistants, une autre nouveauté de la présente thèse tire son origine d'une autre difficulté, qui est l'inexistence d'une méthodologie communément acceptés pour l'analyse de ce type de données.

Pour comprendre la raison de l'alternance codique dans les brouillons, une approche multi-perspective est essentielle. La perspective pragmatique montre les fonctions pragmatiques de l'alternance codique ; cependant, celles-ci sont étroitement liées avec la perspective linguistique, qui permet l'examen du schéma de l'alternance codique, qui reflète souvent les fonctions pragmatiques. En outre, une perspective visuelle fournit des informations complémentaires, qui peuvent soit clarifier l'utilisation des fonctions pragmatiques ou indiquer une direction inaperçue.

Dans cette thèse, la perspective visuelle sera utilisée pour démontrer si l'alternance codique est visuellement marquée.

La méthodologie de la perspective pragmatique est basée sur les travaux de Montes-Alcalá (2000), où elle explora les fonctions socio-pragmatiques de l'alternance codique écrite espagnol-anglais. Après avoir adapté cette liste au corpus de l'écriture personnelle, sept fonctions ont été sélectionnées : la citation directe, la clarification et/ou l'élaboration, le commentaire explicatif parenthétique, l'alternance codique déclenchée, l'alternance codique stylistique, le besoin lexical et l'alternance codique sans but particulier. Les sept fonctions mentionnées ci-dessus ne couvrent pas tous les usages pragmatiques de l'alternance codique dans le corpus recueilli ici. Ainsi, sept autres fonctions pragmatiques ont été déterminées et ajoutées pour embrasser les caractéristiques spécifiques de l'écriture dans les brouillons de l'examen. Ces fonctions reflètent les caractéristiques structurales et sémantiques de la rédaction des brouillons, et sont l'encadrement, la structuration, l'alternance codique terminologique, l'économie du langage, la répétition, la liste de vocabulaire et la vérification de l'orthographe.

La troisième perspective est la perspective linguistique, qui examinera l'alternance codique au niveau textuel, ce qui n'est pas forcément chose aisée dans des textes déstructurés comme le sont les brouillons d'examens, mais qui présente néanmoins un intérêt certain, nous le verrons. L'alternance codique peut se produire dans un mot, dans une phrase et entre les phrases voisines.

Les trois chapitres suivants présentent la description du corpus suivant les perspectives visuelles, pragmatiques et linguistiques de l'alternance codique dans les données françaises, russes et bulgares.

Chapitre 5. Les données françaises

Les résultats démontrent que les données françaises sont riches en alternance codique. Elles présentent 13 fonctions pragmatiques parmi lesquelles l'alternance codique sans but particulier, l'alternance codique déclenchée, l'alternance codique stylistique, la clarification et/ou l'élaboration, la structuration et l'encadrement sont les plus fréquentes. Tous les trois types d'alternance codique : intra-/inter-phrastique et à l'intérieur d'un mot, sont illustrés dans les brouillons français, avec une répartition presque égale des types intra-/inter-phrastiques. Un examen plus approfondi des 6 fonctions pragmatiques les plus fréquemment employées à travers la perspective structurelle apporte un soutien à la distribution égale de l'alternance codique intra-/inter-phrastique. Les fonctions pragmatiques de structuration,

l'alternance codique sans but particulier et l'alternance codique déclenchée admettent l'alternance codique intra-/inter-phrastique ; l'alternance codique stylistique et la clarification et/ou l'élaboration suggèrent essentiellement l'alternance codique intraphrastique ; alors que l'encadrement présuppose essentiellement l'alternance codique inter-phrastique. La distribution de ces six fonctions laisse entendre également les motifs de l'alternance codique qui illustrent le fait que l'alternance codique est facilement utilisée par les étudiants français ; les participants changent de langue pour éviter l'interruption dans le processus de l'expression ; d'ailleurs ils appliquent l'alternance codique pour donner des informations plus précises ; et enfin l'alternance codique est utilisée pour structurer ces informations et fournir des marqueurs pour une utilisation plus fonctionnelle du brouillon.

En dépit d'une grande variété de fonctions et des types de l'alternance codique, cette alternance passe inaperçue pour le lecteur. Un peu plus d'un tiers des participants marque visuellement l'alternance codique (45 sur 107); alors que seulement un seul brouillon contient toutes les occurrences de l'alternance codique marquées, sans autre marquage visuel dans le brouillon ; 43 brouillons démontrent que l'alternance codique est partiellement ou entièrement marquée en parallèle avec d'autres contenus du brouillon. Ce ratio signifie que le marquage visuel est utilisé pour mettre l'accent sur des points particuliers de la teneur de la réponse d'examen plutôt que d'attirer l'attention du lecteur sur un changement de langue.

D'ailleurs, une petite enquête, menée dans le cadre de cette thèse, ne montre aucune relation entre l'alternance de codes dans les brouillons et les notes attribuées aux étudiants après la correction de leur copie au propre. La conclusion, que cette description des données françaises suggère, est que les étudiants utilisent l'alternance codique à la fois comme un outil méthodologique et une alternative pour l'auto-expression. Ils utilisent toutes les ressources disponibles, et l'alternance codique est l'une d'entre elles, à préparer le brouillon d'examen de manière efficace avec une structure interne et externe convenable et un contenu explicite.

Chapitre 6. Les données russes

Les données russes comprennent 93 brouillons, qui ont été recueillis après les examens finaux dans le département d'anglais de l'Université d'État de Volgograd (Russie) au cours de la période d'été 2014–2015. La description des données russes montre que malgré le nombre limité de brouillons collectés, les données russes contiennent une variété importante de fonctions pragmatiques (leur nombre s'élève jusqu'à 12). Leur distribution est progressive ; l'occurrence des fonctions est distribuée de 9 à 1. Cependant, du point de vue linguistique, la préférence est

attribuée à l'alternance codique intra-phrastique. L'alternance de codes dans les brouillons russes est plutôt d'un type d'insertion ; généralement un mot ou un groupe de mots est inséré dans un énoncé avec une langue matrice. Cette inférence est soutenue par le fait qu'il n'y a aucune occurrence de l'alternance codique entre la proposition principale et subordonnée, et que la majorité des alternances codiques est de type intra-phrastique.

La moitié des fonctions pragmatiques, dont la fréquence varie entre 9 et 6, est composée de l'économie de la langue, la clarification et/ou l'élaboration, l'alternance codique sans but particulier, la réitération, la liste de vocabulaire et le besoin lexicale. Seule la fonction de l'alternance codique sans but particulier comprend l'alternance intra-/inter-phrastique, tandis que les cinq autres fonctions impliquent la plupart du temps seulement l'alternance intra-phrastique. Ce fait reflète une quantité supérieure d'alternance codique intra-phrastique. Dans le même temps, ces cinq fonctions pragmatiques démontrent que l'alternance de codes constitue un outil pour la composition efficace d'un brouillon ; elle sert à gagner du temps et de l'espace, et d'augmenter la précision de l'information.

La perspective visuelle montre également que l'alternance codique n'est pas marquée en tant que telle, mais constitue une partie du contenu visuellement marqué. Les étudiants russes accentuent les éléments importants de leurs réponses à l'examen soit pour se rappeler les principaux aspects de la réponse ou de se rappeler certaines questions douteuses, qui devraient être revues avant/pendant la composition de la copie au propre. Les brouillons avec une alternance codique visuellement marquée constituent près de la moitié de tous les brouillons bilingues (12 sur 28). La répartition de toute ou une partie de l'alternance codique visuellement marquée (5/7) ainsi que d'autres contenus des brouillons étant visuellement marqués, suggère que l'alternance codique représente un outil dans la composition d'un brouillon d'examen ; à savoir un outil économique appliqué pour les trois dimensions : le temps, l'espace et de qualité.

Chapitre 7. Les données bulgares

Les données bulgares sont les moins homogènes de toutes. La moitié des données est composée des brouillons et une autre de copies au propre. Toutes les données ont été recueillies pendant l'année scolaire 2014–2015. Les résultats montrent que les données bulgares sont assez hétérogènes en raison de la variété des types de copies et l'année d'inscription à l'université. Alors que la majorité des copies au propre a été rédigée par les étudiants en 4e année, les brouillons étaient collectés, soit après un examen d'entrée soit un examen de linguistique des étudiants

en 1ère année ; le niveau de L2 à la fois des candidats et des étudiants en 1ère année peut affecter l'utilisation pragmatique de l'alternance codique. Pour cette raison, la description des données bulgares doit être considérée comme une étude préliminaire, qui donne un premier aperçu de l'alternance codique bulgare–anglais dans l'écriture personnelle académique ; cette étude devrait être développée et réalisée sur un corpus plus grand et plus homogène, qui prendra en compte les limites et les contraintes des données actuelles, pour présenter des conclusions définitives. Cependant, certaines observations méritent d'être mentionnées.

La répartition des fonctions pragmatiques et des types intra-/inter-phrastique de l'alternance codique est déséquilibrée. La fonction pragmatique la plus fréquemment utilisée est l'encadrement avec 43 des 53 occurrences, qui est suivie par l'alternance codique sans but particulier avec seulement 7 occurrences. Dans la mesure où la fonction pragmatique de l'encadrement suggère surtout l'utilisation de l'alternance codique inter-phrastique, le rapport de l'alternance intra-phrastique et inter-phrastique est biaisé en faveur de ce dernier type (4:38, et 2 brouillons contenaient les deux types). D'ailleurs, très peu d'étudiants marquent visuellement l'alternance de codes, mais si elle est marquée, l'accent est mis sur l'énoncé et non sur l'alternance codique en soi. Ces résultats permettent de déduire que les étudiants bulgares, qui participèrent à cette étude, utilisent l'anglais et le bulgare dans des buts différents : tandis que la langue anglaise est utilisée pour élaborer la réponse d'examen, la langue bulgare est utilisée pour fournir l'information externe, qui est sans rapport avec le contenu de la réponse d'examen. Chaque langue a sa propre utilisation délibérée.

Partie III

La partie descriptive est suivie par les discussions analytiques dans la partie III.

Chapitre 8. Comparaison

Le chapitre 8 établit une comparaison interne des résultats des sous-corpus et revoit la possibilité de la comparaison externe.

8.1. Comparaison interne

Section 8.1 effectue la comparaison des données françaises, russes et bulgares sur la base des perspectives visuelles, pragmatiques et linguistiques et propose une comparaison fondée sur le principe de la relation entre la langue et le contenu proposé par Sebba (2012). Les résultats sont considérés à la fois individuellement et en relation avec d'autres résultats.

L'étude montre que les participants utilisent l'alternance codique dans leur écriture personnelle académique quel que soit le groupe linguistique auquel ils appartiennent. Chacun d'entre eux emploient l'alternance codique comme un outil pour rendre leur travail plus efficace ; l'alternance de codes est l'une des options ainsi que L1, l'organisation structurelle du texte, marquage visuel, etc., qui sont disponibles pour les étudiants. Les deux types de l'alternance codique, à savoir intra- et inter-phrastique, sont présents dans le corpus. Cependant, la distribution de variables est fort différente dans les trois sous-corpus, laissant supposer que les participants français, russes et bulgares utilisent l'alternance codique pour des raisons différentes. Toutes les trois perspectives (visuelle, pragmatique et linguistique) ainsi que les relations de la langue du contenu, produisirent les résultats hétérogènes. Alors que tous les participants, comme mentionné ci-dessus, utilisent l'alternance codique comme un outil d'efficacité, il faut se rappeler que l'efficacité est multi-facettes et ici chaque groupe linguistique met l'accent sur un aspect particulier de l'efficacité – ce qui est le plus approprié pour leur rédaction des brouillons d'examen.

8.2. Comparaison externe

Section 8.2 présente une comparaison générale de l'écriture personnelle académique avec la communication orale et la littérature sur la base des fonctions pragmatiques utilisées. Les fonctions pragmatiques ne sont pas comparées directement en raison de leur distribution considérablement variée à la fois dans le discours oral en raison de différents contextes sociaux (Gumperz 1982; Myers-Scotton 1999; Gardner-Chloros 2009; Walters 2005) et dans les textes littéraires à cause des multiples genres littéraires et l'origine de l'œuvre littéraire (Jonsson 2005; Jonsson 2010; Montes-Alcalá 2012). Cette variation est également observée dans le corpus de la thèse actuelle, probablement à cause de la différence dans le degré de la maîtrise de la langue et le contexte culturel des étudiants. La comparaison générale des fonctions pragmatiques dans cette section conclut que les brouillons d'examen représentent une forme d'écriture oralisée.

Chapitre 9. Discussion

Chapitre 9 entreprend une discussion dialectique et indique à la fois les contributions scientifiques et pratiques de la présente recherche.

Conclusion

L'état de l'art présenté dans ce travail a dévoilé une délimitation théorique de l'écriture personnelle bilingue. Après avoir exploré les notions générales du bilinguisme, l'alternance de codes et les notes, on a vu qu'au carrefour entre les trois se trouvent une forme de production bilingue écrite naturalistique, qui reste largement sous-étudiée pour le moment.

Visant à mener à la fois la recherche qualitative et quantitative sur l'alternance codique dans l'écriture personnelle, cette thèse fournit une description multimodale complète de l'alternance codique dans l'écriture personnelle académique. La nouveauté de cette recherche réside à la fois dans son corpus et sa méthodologie. C'est la première étude à recueillir un corpus de brouillons universitaires bilingues. Après la description des données basée sur les perspectives visuelles, pragmatiques et linguistiques, les résultats ont été examinés dans leur interrelation les uns aux autres en soulignant le caractère composite de l'alternance codique. L'alternance de codes représente un outil, que les étudiants utilisent habilement dans leur écriture personnelle académique. Les raisons d'utilisation diffèrent non seulement entre les individus mais aussi entre les sociétés. En ce qui concerne les limitations et les contraintes du présent corpus, des recherches supplémentaires sont nécessaires pour attester ou de contester les résultats. Néanmoins, l'incohérence des résultats met en évidence l'influence de facteurs externes sur l'écriture personnelle. Par conséquent, cette étude ouvre deux perspectives sociolinguistiques. Tout d'abord, d'explorer l'écriture personnelle en tant que témoignage sociolinguistique ; de déterminer les facteurs sociaux qui influent sur l'alternance codique et qui la rendent différente dans l'ensemble des pays/cultures/communautés. Deuxièmement, d'explorer l'écriture personnelle pour avoir un aperçu de la société reçu du côté de ce que j'appellerais volontiers le côté "personnel" caché de la société.

Ce travail a placé l'alternance codique dans l'écriture personnelle académique sur la même ligne que l'alternance codique orale, écrite et littéraire, à la fois en insistant sur ses caractéristiques sociales et apportant la reconnaissance de l'alternance codique dans l'écriture personnelle qu'elle mérite. En outre, cette étude fit valoir la nécessité d'une approche interdisciplinaire pour fournir une

compréhension globale du phénomène de l'alternance codique et de son rôle sur l'arène linguistique.

Université Lumière Lyon 2
UFR des Langues
Département d'Etudes du Monde Anglophone

Written code-switching in the notes of second-language learners in bilingual classroom environments.

Olga Grebeshkova

2017

Thèse

pour obtenir le grade de Docteur de l'Université Lumière Lyon 2
Discipline : Etudes Anglophones - Sociolinguistique

Composition du Jury :

Alexandra Bagasheva, Université Saint-Clément d'Ohrid de Sofia (Bulgarie)	Rapporteuse
Maria Georgieva, Université Saint-Clément d'Ohrid de Sofia (Bulgarie)	Directrice de recherche
Peter Griggs, Université Lumière Lyon 2 (France)	Président du Jury
Svetlana Korolkova, Université d'Etat de Volgograd (Russie)	Rapporteuse
Jim Walker, Université Lumière Lyon 2 (France)	Directeur de recherche

Acknowledgements

The last four years have led to the accomplishment of this PhD thesis, which would not be the same without the help and the support of the people around me. I would like to thank you with all my heart and tell you that I am grateful to have known you in my life.

Thank you James Walker, for being always present from the very first steps of this work, till its last full stop. I am grateful for your omnipresent support.

Thank you Maria Georgieva, for having accepted the co-direction and thus permitting me to start this PhD research.

Много благодаря на Александра Багашева. Owing to your help the collection of the Bulgarian sub-corpora was possible. I am also grateful for your warm welcoming in Sofia.

Большое спасибо Валентине Литвиновой, Алтынай Утешевой и Светлане Корольковой за сотрудничество и за возможность сбора корпуса русских студенческих черновики.

I am grateful to all my friends who sent me the words of support and who believed in me.

Огромное спасибо моей семье, которая позволяла мне заниматься и которая всегда подпитывает меня своей любовью и заботой. Я вас очень люблю.

There are never too many thanks. This work contains your intentions, your thoughts and your warm words. Let me THANK YOU for that.

Abstract in English

The phenomenon of code-switching has been known and researched since the beginning of the 20th century. The psycholinguistic interest in spoken code-switching was followed by socio-pragmatic perspectives in the second half of the 20th century; then towards the end of the century, researchers' attention was drawn to code-switching in writing. Even though language alternation has been well studied since, there are still some lacunae. For example, the intersection of bilingual writing and genetic criticism presents a source of naturalistic bilingual written data which remains significantly under-studied. The present thesis attempts to fill this gap by providing a qualitative and quantitative analysis of second-language learners' examination drafts. The multimodal approach encompasses the visual, the pragmatic and the linguistic perspectives. The findings show that code-switching represents a tool which students skillfully use in their academic personal writing. However, the purposes of its use differ not only among individuals but also among societies/communities. This emphasizes the sociolinguistic value of bilingual personal writing. This work also recognizes code-switching in personal writing as an independent type of code-switching, which can be placed in line with the oral, written and literary code-switching and which merits to be studied as such.

Abstract en français

Le phénomène de l'alternance codique a été connu et étudié depuis le début du XXe siècle. L'intérêt psycholinguistique à l'alternance codique orale a été suivi par des perspectives socio-pragmatiques dans la seconde moitié du XXe siècle ; puis vers la fin du siècle, l'attention des chercheurs a été attirée sur l'alternance codique écrite. Même si l'alternance codique a été bien étudié depuis, il y a encore quelques lacunes. Par exemple, l'intersection de l'écriture bilingue et la critique génétique présente une source de données naturalistes bilingues écrites qui restent fortement peu étudiées. Cette thèse tente de combler cette lacune en fournissant une analyse qualitative et quantitative des brouillons d'examen des apprenants de langue seconde. L'approche multimodale englobe la perspective visuelle, pragmatique et linguistique. Les résultats montrent que l'alternance codique représente un outil que les élèves utilisent habilement dans leur écriture personnelle académique. Néanmoins, les objectifs de son utilisation diffèrent non seulement entre les individus mais aussi entre les sociétés / communautés. Cela souligne la valeur sociolinguistique de l'écriture personnelle bilingue. Ce travail démontre également l'alternance codique dans l'écriture personnelle comme un type indépendant de l'alternance codique, qui peut prendre une place égale à côté de l'alternance codique orale, écrite et littéraire et qui mérite d'être étudié en tant que tel.

Table of contents

Introduction	5
Part I State of the art	9
Chapter 1. Bilingualism	11
1.1. Definition	11
1.2. Typologies	13
1.2.1. Hugo Beatens Beardsmore's typology	13
1.2.2. Hamers and Blanc's typology	16
1.2.3. Conclusion to Typologies	22
1.3. Terminology	23
Chapter 2. Language alternation	29
2.1. Spoken code-switching	30
2.1.1. Foundational studies	30
2.1.1.1. Historical overview of code-switching	30
2.1.1.2. Terminology and definitions	31
2.1.2. Theoretical framework	36
2.1.2.1. Grammatical aspects of code-switching	37
The free morpheme constraint and the equivalence constraint	37
The clitic constraint	39
The government constraint	40
The Matrix Language Frame model	40
2.1.2.2. Psycholinguistic model of triggering	42
2.1.2.3. Social aspects of code-switching	44
We-code vs they-code	45
Situational vs metaphorical/conversational code-switching	45
Discourse-related vs participant-related language alternation	47
Markedness model	47
2.2. Language alternation in writing	49
2.2.1. Psychological and educational perspective	50
2.2.1.1. Literacy of bilingual children	50
2.2.1.2. The use of L1 in second-language composing	52
2.2.2. Sociolinguistic perspective	57
2.2.2.1. Code-switching in literature	57

2.2.2.2. Code-switching in electronic writing.....	61
2.2.2.3. Theoretical framework for multilingual texts.....	63
Chapter 3. Notes.....	67
3.1. Note-taking.....	67
3.2. Genetic criticism	74
3.3. A study with a dual corpus.....	77
3.4. Interpreters' note-taking.....	79
3.5. Conclusion.....	81
Part II Data description and results	83
Chapter 4. Methodological matters.....	85
4.1. The choice of the subject.....	85
4.2. Corpus	87
4.3. Methodology	98
Chapter 5. French data	107
5.1. Visual perspective	108
5.2. Pragmatic perspective	112
5.2.1. Pragmatic functions	112
5.2.2. Summary table.....	133
5.3. Linguistic perspective	134
5.3.1 Intra-sentential code-switching.....	134
5.3.2. Inter-sentential code-switching.....	136
5.3.3. Word-internal switching.....	139
5.3.4. Summary table.....	140
5.4. Grading.....	141
5.5. Summary and conclusion	142
Chapter 6. Russian data.....	145
6.1. Visual perspective	146
6.2. Pragmatic perspective	150
6.2.1. Pragmatic functions	150
6.2.2. Summary table.....	166
6.3. Linguistic perspective	167
6.3.1. Intra-sentential code-switching.....	167
6.3.2. Inter-sentential code-switching.....	170
6.3.3. Word-internal code-switching.....	171
6.3.4. Summary table.....	172
6.4. Summary and conclusion	172

Chapter 7. Bulgarian data	175
7.1. Visual perspective	177
7.2. Pragmatic perspective	179
7.2.1. Pragmatic functions	179
7.2.2. Summary table	185
7.3. Linguistic perspective	186
7.3.1. Intra-sentential code-switching	186
7.3.2. Inter-sentential code-switching	187
7.3.3. Word-internal code-switching	189
7.3.4. Summary table	189
7.4. Summary and conclusion	189
Part III Comparison and discussion	191
Chapter 8. Comparison	193
8.1. Internal comparison	193
8.2. External comparison	199
Chapter 9 Discussion	205
9.1. Discussion	205
9.2. Contribution of the study	209
Conclusion	211
References	213
Web resources	229
Résumé	231
Appendices	253
Appendix 1	255
1.1. French samples	256
1.2. Russian samples	276
1.3. Bulgarian samples	293
1.3.1. Fair copies	293
1.3.2. Drafts	300
Appendix 2	305

Introduction

“Johannes Wagner, who was born and raised in Germany and moved to Denmark 30 years ago, often thinks in German, speaks Danish and writes in English” (Auer & Li Wei 2007: 392). This kind of switching between languages for different purposes, which is second nature for those people who have more than one language at their disposal, lies at the heart of this thesis.

With globalization and its concomitant online access to a huge amount of information in different languages, greater demands on the labour market, a wider array of personal interests, etc., people either find themselves immersed in multilingual environments (for example immigrants, professionals working abroad, foreign students, children of mixed marriages, etc.) or they seek to come closer and deliberately try to adapt themselves to these new conditions (for example learners of a second language). One feature of these increased multilingual environments, whether natural or learned, is an increase both in the prevalence of, and an interest in code-switching.

Code-switching is a sociolinguistic phenomenon that can be found in environments where at least two languages or language varieties are in contact. Bi-/multilinguals switch fluently between the languages they speak. Sometimes the switch can occur between sentences or within an utterance. I am almost sure that every bi-/multilingual speaker has faced, at least once in their life, a situation where they do not know which language is better to use or indeed are unaware that the sentence they have started is in one language, but its ending is in another one. As a speaker of Russian, English and French, I have found myself many times in situations of this kind. This explains my interest in bilingualism and especially in code-switching. Questions which affect one personally, such as this, provide motivation and encouragement when undertaking research. Johannes Wagner, quotation about whom began this chapter, and who himself uses three different languages in his everyday life, is one of the many linguists interested in second language conversation and learning. Another well-known personality is Li Wei, a British linguist of Chinese origin. Li Wei is a professor of applied linguistics and is famous for his publications mainly devoted to bilingualism and multilingualism. In his works a considerable part is dedicated to code-switching. These authors examine code-switching from within and from without: as a scientist and as a participant at the same time. The same is true of me.

Code-switching has been of interest for me for several years. After my first year in France I noticed that I had some difficulties finding the right word while speaking with my family, who are totally Russian speaking. However, I felt at ease in communication with my Russian friends who also studied in France. This can be easily explained by the use of code-switching. With my friends I was switching Russian and French, knowing that they would understand me in both languages; however, I could not do the same with my family. I had to find a Russian equivalent for each word or phrase that I was used to employing without translation with friends. After a year in France speaking “pure” Russian seemed challenging for me.

There is a considerable body of research on code-switching. Many approaches and models have been developed to explain and detail cases of code-switching like mine. One of the concerns of scholars is the manner in which the bilingual data are collected; according to Labov linguists should study the language spoken in “natural” conditions (Labov 1972: 205). For that reason, a lot of studies have been focused on the code-switching in real-life oral communication. Different grammatical and social models and rules have been tested on dozens of language pairs; the results of similar and even very different languages have been compared. A great number of studies have been performed.

With technological progress, a lot of written multilingual data has become accessible for researchers on the World Wide Web, which gave rise to the studies on code-switching in email correspondence, blogs, online journals, etc. Furthermore, code-switching in literary works also occupies a special place in the canon of research. Different both from code-switching in oral and written communications, literary code-switching has only recently gained the recognition it perhaps deserves.

An interesting observation that I made during a conference on psycholinguistics which I attended opened up to me the idea that naturalistic bilingual written data could be an interesting research topic, one that is, for the present, under-studied. The speaker at the conference was an invited academic of Spanish origin. His native language is Spanish, the PowerPoint presentation of his work was in English, but as we were in France and as he had already greeted the audience in French, he decided to continue his talk in French. Even though he was speaking French, sometimes one could hear some phrases or words in his native language. For example, when he was explaining something, suddenly the screen with the PowerPoint presentation turned black; his reaction was: “Qu’est-ce qui se passe *aquí*?”. During his presentation listeners, even those who were monolingual, were immersed into a bilingual or even trilingual setting. Such a medium favours the

application of several languages. And it was at this moment that I observed an interesting thing: the notes that I was taking were in English, French, and Russian.

The bi-/multilingual notes taken during a conference represent a type of written practice which differs both from literary texts and written communication. The example shown above demonstrates that bi-/multilingual speakers use several languages not only to communicate information for someone, as the Spanish lecturer was doing, but also to note down some information for themselves, as it appeared I was doing. Such a writing activity can be called personal writing, because it is not intended for an external reader—only for its author. Thus, turning to personal writing, I started to discern that bilingual speakers code-switch when they make to-do lists or write private diaries, my students code-switch while note-taking in language classes, I code-switch while preparing my lessons. Hence, personal writing appears to be a rich source of bilingual data. However, despite the fact that this writing activity is widespread among bilinguals, code-switching in personal writing remains dramatically under-studied. The only work on the bilingual private diary was performed by Montes-Alcalá (2000), where the author examined code-switching in her own personal journal.

Since my attention was now focused on bilingual personal writing, I needed to collect a corpus which could provide the possibility to make a qualitative and quantitative analysis of code-switching. Due to several reasons (which will be discussed in Part II), students' examination drafts were selected to compose the three sub-corpora, namely the samples of the French, Russian and Bulgarian second-language learners. Hence, this PhD thesis will present my research on code-switching in academic personal writing and give a multimodal description of this phenomenon. It will also attempt to explain the reasons for code-switching in the examination drafts. It will lastly situate code-switching in academic personal writing on the bilingual stage.

The first part of the thesis will give a thorough literary overview. Chapter 1 will be devoted to the issue of bilingualism; besides dealing with definitions and terminology, this chapter will describe several typologies of bilingualism to situate later the participants of the corpus in the bilingual continuum. Chapter 2 will examine the phenomenon of language alternation both in its oral and written forms. Section 2.1. will provide a historical overview of spoken code-switching research as well as the theoretical framework, which includes grammatical, psycholinguistic and social models of oral code-switching. Section 2.2. will be focused on language alternation in writing regarding such fields of its application as the productions of bilingual children and second-language learners, literature and electronic

communication. It will provide a multimodal frame of analysis for bilingual texts as well. Chapter 3 will deal with the notion of notes. The process of note-taking will be first examined as a cognitive activity. Then the emphasis will be on the genetic criticism. And finally different types of notes will be depicted: students' notes and interpreters' notes. This chapter will be essential to contribute a full comprehension of the students' samples, which comprise the corpus of this thesis.

Part II will present the empirical evidence of the research. Chapter 4 will justify the choice of the subject and introduce the novelty of the corpus and the methodology. Chapters 5, 6 and 7 will examine the French, the Russian and the Bulgarian sub-corpora, respectively; it will provide a detailed description of the data from the visual, pragmatic and linguistic perspectives.

The descriptive part will be followed by the analytical discussions in Part III. Chapter 8 will draw an internal comparison of the sub-corpora's results and follow the possible external comparison. Then Chapter 9 will develop a dialectical discussion and state the contributions of the present research.

Finally, the conclusion will summarize the findings and draw perspectives for future research.

Part I

State of the art

Chapter 1. Bilingualism

Bilingualism is a vast field which is central to a variety of disciplines such as neuroscience and psychology, artificial intelligence, studies in language contact, bilingual education, etc. Bilingualism as such hides interesting issues and polemics. This chapter will try to answer such questions as: who is a bilingual? How can bilinguals differ from each other? What are the different levels of bilingualism? By clarifying the above-mentioned issues this chapter will explain why the participants of the present research, who are second-language learners of English, can and will be considered bilingual speakers.

1.1. Definition

It is not an easy task to define bilingualism. It was long deemed as an equal knowledge and usage of two languages. Li Wei says that “[t]he word “bilingual” primarily describes someone with the possession of two languages” (Li Wei 2000a: 7). Hamers and Blanc also give a "popular" characteristic of a bilingual person adding one detail with regard to the degree of competence saying that “[i]n the popular view being bilingual equals being able to speak two languages perfectly” (Hamers and Blanc 1989: 6).

Giving a rather general picture of a bilingual person these definitions bring about two opposite points of view representing a maximalist and minimalist viewpoints. A maximalist interpretation could be found in the Webster Dictionary of 1961 according to which a bilingual person is described as “having or using two languages especially as spoken with the fluency characteristic of a native speaker; a person using two languages especially habitually and with control like that of a native speaker” (Webster Dictionary, as cited in Hamers and Blanc 1989: 6). The entry of the Webster’s dictionary of today (2016) refers to a bilingual as “a person who speaks two languages fluently”. The reference to the native speaker has disappeared following the tendency of the recent studies where bilingual speakers are not compared to the monolingual or native speakers any longer; at the same time the latter definition does not reflect the development of the typologies¹ where the degree of bilingualism is measured with regard to the linguistic, cultural, psychological, etc. factors of bilingual speakers. This definition remains rather

¹ The typologies of bilingualism will be discussed in the section 1.2.

general and maximalist. At the opposite pole John Macnamara proposes a minimalist definition using the term “bilingual” for “persons who possess at least one of the language skills [speaking, writing, listening or reading] even to a minimal degree in their second language” (Macnamara 1967: 59–60).

All these definitions lack precision. They do not clarify what it means “to speak two languages perfectly” or “fluently”, “with control like that of a native speaker”; neither do they determine what is meant by “minimal degree” of the language skills. Should those people who speak more than two languages be excluded from the notion of a bilingual speaker? Is the level of language proficiency the only criterion for designating a bilingual?

Such authors as Mackey (1962/2000), Macnamara (1967) and Baetens Beardsmore (1982) amplify their definitions by evoking the notion of continuum and thus showing the impossibility of providing a universal definition of bilingualism. Bilingualism represents a continuum or a series of continua which can be inconsistent amongst bilingual individuals. Bilinguals may vary in their proficiency: how well does the person master the language? Bilinguals may have different functions of language use: for what purpose does the person use the language? Bilinguals may know the same languages but switching facilities may differ: what is the quantity of alternation between the languages they use? There are many other linguistic and non-linguistic dimensions, such as the age of acquisition, the cognitive organisation, cultural identity, etc., which are applied to compare bilinguals. If bilingualism is regarded as a continuum it implies that bilingualism is a relative notion (Mackey 1962/2000; Baetens Beardsmore 1982). Relativism is also emphasized by the fact that there is no strict limitation of a number of languages involved. Bilingualism should be, thus, considered as an all-embracing term including cases of multi-/plurilingualism. “We shall therefore consider bilingualism as the alternate use of two or more languages by the same individual” (Mackey 1962/2000: 27).

Some scholars (Weinreich 1953/1970; Mackey 1962/2000; Baetens Beardsmore 1982; Hamers and Blanc 1989; Butler and Hakuta 2006; Edwards 2006) prefer to work with typologies or dimensions of a phenomenon which determine the boundaries of a particular area of research and which focus on a particular type of the studied phenomenon rather than content themselves with definitions. The issue of typologies will be discussed in the next section.

1.2. Typologies

1.2.1. Hugo Baetens Beardsmore's typology

One of the more widely acknowledged typologies is that proposed by Hugo Baetens Beardsmore. It is often quoted in other works. For example, Hamers and Blanc mention that “Baetens Beardsmore (1982) has listed some thirty-five definitions and typologies of bilingualism [...]” (Hamers and Blanc 1989: 7). Inter alia, Li Wei adduces the table with a great variety of terms of a bilingual speaker proposed by Baetens Beardsmore (Li Wei 2000a). Also Colin Baker recommends Baetens Beardsmore's *Bilingualism: Basic Principals* in his book *Foundations of Bilingual Education and Bilingualism* (1993) for further reading. Google Scholar shows a rate of 901 for the number of quotations of Baetens Beardsmore's second edition.

Baetens Beardsmore's typology is a very rich classification. It assembles a lot of different types of bilingualism which are based on such factors as time of language acquisition (early and late, consecutive and successive bilingualism), environment of language acquisition (natural or primary and secondary bilingualism), speaker proficiency (gradation from ambilingualism, equilingualism, non-fluent bilingualism to semilingualism; receptive and productive), mental organization (subordinate, compound and coordinate), the effect of L2 on L1 (additive and subtractive), the use or implementation of bilingualism (ascendant and recessive), etc. (For more detail see Baetens Beardsmore 1982).

However, this typology has a few weaknesses. First, there is no internal structure. The classification represents a kind of list where different types of bilingualism are presented more or less randomly. Some types of bilingualism are subcategories of others, and at the same time they are linked to one of the next categories; these connections readers are left to make on their own. This way of representation makes comprehension of the typology more difficult, especially taking into account the number of the represented types of bilingualism. Second, some explanations are no longer in line with present studies and globalized language contact situations.

For example, one of the first distinctions that Hugo Baetens Beardsmore (1982) makes is between *societal* and *individual bilingualism*. He does not give any definition; instead, the author uses a descriptive approach. He says that in the studies of societal bilingualism the emphasis is placed on the linguistic situation in a

community, the inter-relationships between “linguistic forces”, the contact of politics, economics, society, education, culture and language. According to these studies “[...] language plays an important role in a differential social distribution of positive and negative social values of both a material and symbolic nature” (Kjolseth, as cited in Baetens Beardsmore 1982: 4). Without giving any description of individual bilingualism he stresses that one of the most prominent differences between societal or group bilingual research and that of individual bilingualism is the case of multilingual countries or nations which consist mainly of monolingual individuals living close to each other. Baetens Beardsmore gives the examples of Belgium and Switzerland where societal bilingualism is rooted on the principle of territorial unilingualism (Baetens Beardsmore refers to William F. Mackey). He says that the majority of the inhabitants of these countries are monoglots. However, there are a number of bilingual individuals who represent mediators both between the different groups present in the country and between societal and individual bilingualism. Of course, we should take into account that since Baetens Beardsmore’s publication in 1982 the demographic situation in these countries has changed. For example, the Swiss Federal Statistical Office published the results of 2014 survey announcing that “[...] almost two-thirds (64%) of Switzerland’s population use more than one language at least once a week. 38% use two, and a quarter (26%) three or more.”² While the official statistics on bilingual speakers in Switzerland and Belgium is hardly available because the official websites give information only about the numbers of speakers of each language group, but do not inform about their bi-/multilingual abilities; there are some interesting figures easily accessible about another multilingual country—Canada. The 2011 census showed that 17.5% of the Canadian population reported being able to conduct a conversation both in English and French; and 20.6% of Canadians declared a mother tongue other than one of Canada’s official languages³. Boundaries between monolingual individuals in a multilingual country are blurring nowadays considering globalisation and migration policies. It has been more than 30 years since the first edition of *Bilingualism: Basic Principals* by Baetens Beardsmore was published; now some of his examples correspond more to idealized situations than to real ones and should probably be revised.

² In *First results of the language, religion and culture survey 2014 – Language* [online], retrieved 9/11/2016, from <https://www.bfs.admin.ch/bfs/en/home/statistics/population.assetdetail.1000408.html>

³ In *Linguistic Characteristics of Canadians* [online], retrieved July 2015, from <https://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-314-x/98-314-x2011001-eng.cfm>

Another example of a category of bilingualism which is rather ill-defined is that of diagonal bilingualism. Baetens Beardsmore shows three types of bilingualism (namely *horizontal*, *vertical* and *diagonal bilingualism*) mentioning that they belong to the typology proposed by Pohl (1965). According to Baetens Beardsmore, horizontal bilingualism describes the cases when two distinct languages have an equal status, for example Catalan and Spanish or Dutch and French in Brussels. Horizontal bilinguals could use their languages for different functions but they could also use the different languages in similar situations creating an overlapping of languages. Vertical bilingualism represents the situation of *diglossia* where a high and a low variety are counterposed in the society i.e.

DIGLOSSIA is a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation (Ferguson 1959/2000: 75).

Baetens Beardsmore gives the examples of Walloon Belgium with Walloon and French, Germanic Switzerland with Schwyzertütsch and German and Bali with Balinese and Indonesian. Diagonal bilingualism occurs in the situations where a dialect or non-standard language is used together with genetically unrelated standard language, for instance Louisiana French and English in Louisiana in the USA, Low German and French in German Belgium or Maori and English in Maori communities in New Zealand. Diagonal bilingualism seems to describe a category that comprises all the cases that could not be included either in the horizontal or vertical bilingualism. Besides, should Louisiana French be considered as a dialect or a localized standard language? In addition, Maori is neither a dialect, nor a non-standard language, but an indigenous language which has been recognized as one of the official languages of New Zealand since 1987.

This typology comprises some of the previously proposed types of bilingualism, for example those by Weinreich (1953/1970) and Pohl (1965). Interesting as Baetens Beardsmore's typology may be, it is somewhat dated, and I would like to turn my attention to a more recent one, which improves on Baetens Beardsmore's typology in a number of ways. In its turn, it includes some types of

bilingualism which were proposed by Weinreich and Baetens Beardsmore, but which are arranged quite differently.

1.2.2. Hamers and Blanc's typology

Before introducing the new typology, I would like to come back to the concept of bilingualism. It would be interesting to have a look at the following definition given by Aucamp (1926):

Bilingualism is the condition in which two living languages exist side by side in a country, each spoken by one national group, representing a fairly large proportion of the people (Aucamp, as cited in Baetens Beardsmore 1982: 2)

This definition differs from all the above-mentioned ones by its societal aspect. It corresponds to the societal type of bilingualism given by Baetens Beardsmore; however, Aucamp does not represent different types of bilingualism by the above-mentioned definition rather than bilingualism as such. The same approach could be found in Hamers and Blanc (1989). From the very beginning they distinguish between *bilinguality* and *bilingualism*. "Bilinguality is the psychological state of an individual who has access to more than one linguistic code as a means of social communication" (Hamers and Blanc 1989: 6). While bilingualism is represented as a more general concept which includes that of bilinguality "but refers equally to the state of a linguistic community in which two languages are in contact with the result that two codes can be used in the same interaction and that a number of individuals are bilingual" (Ibid). Bilinguality and bilingualism as defined by Hamers and Blanc correspond to individual and societal bilingualism respectively. However, the authors go further by providing classifications of each phenomenon.

Bilinguality is a continuum formed by different dimensions, which could sometimes correlate between each other. However, there is never one-to-one correspondence between the factors and the resulting type of bilinguality. A combination of dimensions determines a psychological state of a bilingual, which is very individual because each person lives their own bilingual experience. Here is the table of bilinguality's dimensions proposed by Hamers and Blanc in their book *Bilinguality and Bilingualism* (1989).

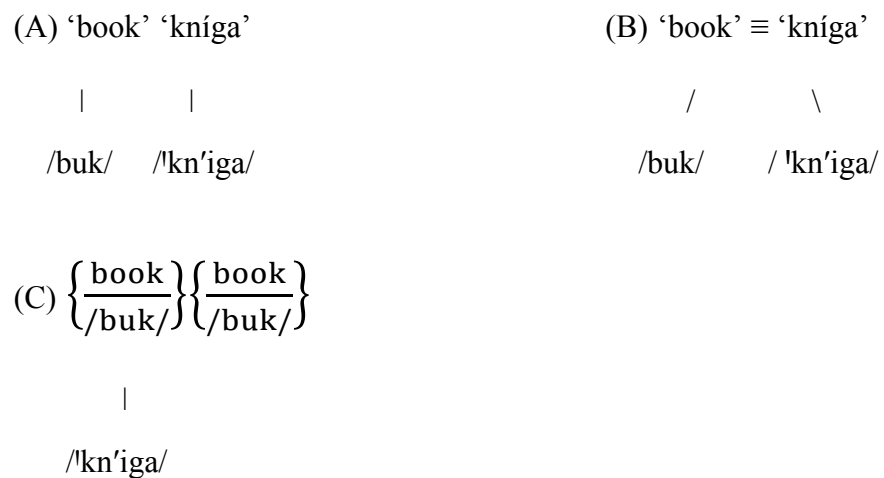
Summary table of psychological dimensions of bilinguality

Dimensions		Comments
A.	According to competence in both languages	$L_{A/1}$ competence = $L_{B/2}$ competence $L_{A/1}$ competence > or < $L_{B/2}$ competence
B.	According to cognitive organization	$L_{A/1}$ unit equivalent to $L_{B/2}$ unit = one conceptual unit $L_{A/1}$ unit = conceptual unit 1 $L_{B/2}$ equivalent = conceptual unit 2
C.	According to age of acquisition	$L_{B/2}$ acquired before age 10/11 L_A and L_B = mother tongues L_1 = mother tongue; L_2 acquired before 11 L_2 acquired between 11 and 17 L_2 acquired after 17
D.	According to presence of L_2 community in environment	presence of L_2 community absence of L_2 community
E.	According to the relative status of the two languages	$L_{A/1}$ and $L_{B/2}$ socially valorized → cognitive advantage L_2 valorized at expense L_1 → cognitive disadvantage
F.	According to group membership and cultural identity	double membership and bicultural identity $L_{A/1}$ membership and cultural identity $L_{B/2}$ membership and cultural identity ambiguous membership and anomic identity

The types of bilinguality which correspond to dimensions (see the table):

- B - According to cognitive organization (*compound/coordinate bilinguality*) and
- E - According to the relative status of two languages (*additive/subtractive bilinguality*)

also appeared in the Baetens Beardsmore's typology. The distinction according to cognitive organization belongs to Uriel Weinreich, who introduced it in his work *Languages in Contact* in 1953. Weinreich describes in a very accessible way three types of mental organization in bilinguals – namely *coordinate*, *compound* and *subordinate bilingualism* (Weinreich, 1953/1970):



Coordinate bilingualism is shown in the diagram A. It represents a situation where a person has a separate system of meaning for each language which means that a signifier (a physical form of a sign) in each language refers to a separate signified, or as Weinreich says, *semanteme* (the meaning expressed by the sign). The diagram B arranges the English signified 'book' and Russian 'kníga' (the examples as well as their visual representation are taken from Weinreich 1953/1970) into one system of meaning which has two signifiers; this type is called compound bilingualism. The last subordinate type of bilingualism (diagram C) characterizes a person who treats an L2 signifier through the L1 signifier and only the latter is related to the signified. Weinreich gives a good example of an English learner of Russian who may perceive the signified /'kn'íga/ not as the object, but firstly as the English word book.

It was already stipulated by Weinreich (1953/1970) that a bilingual could not be entirely of the coordinate or compound type. They could be more compound in one field and more coordinate in another one; boundaries may shift with time. For

example, a child who is brought up in a bilingual family is a compound bilingual, speaking daily their languages A and B while living in the country with the official language A. Then growing up, this person decides to study, for example, medicine and gets enrolled in medical school in the country with the official language B. Thus through their academic education, this person becomes a coordinate bilingual in the medical sphere. After the graduation and 10 years of medical practice in the country of their studies, this person returns to the country of origin and continues to work in the medical field, this time with the working environment being in the language A. With time this person becomes a compound bilingual in the medical field.

Hamers and Blanc keep only compound and coordinate types in their classification while excluding subordinate mental organization.

The second subdivision of bilinguality, which is common to Baetens Beardsmore's and Hamers & Blanc's typologies, is according to the relative status of two languages (point E). The division between *additive* and *subtractive* bilinguality was first proposed by Lambert (1973). Baetens Beardsmore resorts to the definition given by Lambert in 1974 of *additive* bilingualism "in which the second language brings to the speaker a set of cognitive and social abilities which do not negatively affect those that have been acquired in the first language but where the two linguistic and cultural entities involved in being bilingual combine in a complementary and enriching fashion" (Lambert, as cited in Baetens Beardsmore 1982: 19). Baetens Beardsmore gives the examples of Greek and Latin which were deemed indispensable for a well-educated elite in ancient Rome, or the contemporary situation in Singapore with positive status of both Mandarin and English. At the opposite side appears the notion of *subtractive bilingualism* "where the second language is acquired at the expense of the aptitudes already acquired in the first language and where, instead of producing complementarity between two linguistic and cultural systems, there is competition" (Baetens Beardsmore 1982: 19). Such situations could be found in societies in which one language is considered more prestigious socio-economically while the other language is depreciated. For example, when immigrant children are enrolled in a class with a monolingual education in the official language of the country, which differs from children's mother tongue. Or a more specific example may be given, for instance the case of Pakistani or Bangladeshi children in the UK.

Coming back to Hamers and Blanc's classification, the points (see the table)

- A - According to competence in both languages and
- C - According to age of acquisition

are partially shared by Baetens Beardsmore's typology; they take into consideration the same dimensions. However, the two typologies do not include the same types of bilinguality. For example, Hamers and Blanc are not interested in the question raised by Baetens Beardsmore—namely both in the degree of competences as a whole and in the degree of separate skills—but rather reduce the scope to the equality or dominance of one of the languages. That leads to *balanced* and *dominant bilinguality* respectively.

According to age of acquisition, Hamers and Blanc distinguish between *childhood*, *adolescent* and *adult bilinguality*. In childhood bilinguality, the acquisition of two or more languages happens before the age of 10 or 11 at the same time as the general development of a child. The child can develop two languages concurrently through informal unschooled learning. For example a child of a family where parents speak different languages acquires at a time two mother tongues⁴, which are labelled L_A and L_B in the above given table. This type of bilinguality is called (a) *simultaneous* or *infant bilinguality*. *Consecutive* childhood bilinguality (b) occurs when a child learns a second language early in childhood but after having acquired the minimal linguistic basis of the first one. In this case and in all other consecutive bilingual acquisition, Hamers and Blanc refer to the first acquired language as L_1 and L_2 to the second language. Consecutive childhood bilinguality may happen both informally or as a result of intentional learning; the former case may represent a situation with an immigrant family with children and the latter one may issue from certain bilingual educational programs.

The dimension of bilinguality according to the presence of L_2 community in the environment (point D) shows an interesting perspective on bilinguality, which was not explored by Hugo Baetens Beardsmore in his exhaustive typology. Hamers and Blanc introduce *endogenous* and *exogenous bilinguality*. “An endogenous language is one that is used as a mother tongue in a community and may or may not be used for institutional purposes, whereas an exogenous language is one used as an official, institutionalized language but has no speech community in the political entity using it officially” (Hamers and Blanc 1989: 11). The authors give an example of English and French in West, Central and East Africa. Children from such African countries speak one language at home and French or English at school as it is an exclusive language of institution. They will grow up with an exogenous bilinguality in the two languages.

⁴ Section 1.3. will examine the issue of terminology. And notably the term "mother tongue" will be discussed in detail.

Finally, Hamers and Blanc distinguish bilinguals according to group membership and cultural identity (point F). A bilingual person may identify him-/herself with two cultural groups whose languages they speak. In this case the person could be also called *bicultural*. A balanced biculturalism often goes together with a balanced bilinguality. Biculturalism is probably an analogue at the socio-affective level of additive bilinguality at the cognitive level. However, Baetens Beardsmore notes that bilingualism does not need to be accompanied by biculturalism (Baetens Beardsmore 1982: 20). This assumption coincides with other cases of bilinguality according to cultural identity proposed by Hamers and Blanc. In multilingual societies a person can be a fluent bilingual while identifying themselves only with one cultural group, thus, remaining *monocultural*. Due to some bilingual experience a person can come to leave their mother-tongue cultural identity and adapt cultural identity of the second-language community. In this case the person will come to be an *L₂ acculturated bilingual*. However, if a bilingual renounces their L₁ cultural identity and cannot manage to identify themselves with the second-language cultural group, the person becomes *deculturated*.

All these dimensions describe quite well the psychological state of bilinguals situating them in a continuum of bilinguality. Each bilingual individual can be viewed from different dimensions of their bilinguality at the same time. In their turn, these people can be part of a multilingual community, which has some bilingual characteristics. To exist, a bilingual community needs at least two languages spoken on a daily basis by some members of the community. Either the community consists of two monolingual groups speaking two different languages with a few bilinguals who speak both languages, or some people in each group speaking a lingua franca, the third language shared by each group. Hamers and Blanc represent bilingualism as a continuum as well. At the one end of the continuum there are two monolingual groups who use one language each for all purposes along with a small number of bilinguals. At another end of the continuum there is one group where quite many members speak both languages but for different functions. Hamers and Blanc distinguish three types of bilingualism:

1. Territorial bilingualism occurs when there are politically defined territories for each group. They stay mostly on their own territory and the official status of the other languages often depends on the country's policy. Territorial bilingualism can be found in such countries as Belgium, Switzerland, the former Soviet Union, Spain and Canada.

2. Another case of bilingualism can be found in the countries where along with the native languages of ethnic groups there is another language which is common to all the groups but which is native to none or few of them. The example is a lingua franca such as Swahili in Eastern and Central Africa or an exogenous language like French or English in some African countries which were politically superposed often due to a colonial past.
3. The third type of bilingualism can be found in diglossic societies where a substantial number of speakers use both languages in a complementary way due to a higher status of one of the languages. Bilinguals use the languages for different functions and domains. Examples of countries with diglossic bilingualism are Paraguay with the use of Spanish and Guarani and Haiti with French and Creole.

Hamers and Blanc regrouped various dimensions of bilinguality and bilingualism showing the multidimensional nature of these phenomena. Like Hugo Baetens Beardsmore, they took into account such disciplines as psychology, sociology and linguistics. However, they managed to emphasize that the phenomenon of bilingualism should be studied not only on the individual level but also on the societal and intergroup levels; both of which should be regarded as continua.

1.2.3. Conclusion to Typologies

The above-mentioned typologies describe the phenomenon of bilingualism in a quite explicit manner. While it is rather difficult to determine bilingualism in a definition, typologies manage to explain different aspects and facets of this phenomenon. They show that there is never one-to-one situation; each case of bilingualism can have its particular features due to circumstances which have occurred in a community or personal bilingual experience. Of course these two typologies are neither exclusive nor unique. There is, for example, a typology proposed by William F. Mackey (1962/2000) who says that since bilingualism is a relative concept it may be described regarding four intrinsic characteristics: degree, function, alternation and interference. The question of degree involves proficiency;

how well does the individual know each of the languages? The question of function concerns roles of languages in the individual's total pattern of behaviours; what does a bilingual speaker use their languages for? Alternation refers to the extent to which the speaker alternates between their languages. And the fourth question of interference focuses on the extent to which the bilingual succeeds in keeping their languages apart or whether they fuse them.

However, this PhD work does not intend to look over all the definitions and typologies of bilingualism. Rather, it aims at giving an overview of the phenomenon of bilingualism with its multidimensional and at the same time problematic issues. The outline of bilingualism presented in this chapter emphasizes the importance of relativism; both individual and societal bilingualism or bilinguality and bilingualism represent a continuum. Each bilingual individual lives different experiences and faces different environments where they become acquainted with second/third, etc. languages. And each such individual finds their place in the continuum of bilinguality/bilingualism. The arguments provided in this chapter thus allow me to claim that the second-language learners, who are the subjects of the present PhD research, are indeed bilingual. This issue will be looked at in Part II section 4.2. giving more detail on the participants of the research.

A matter of concern that has not been mentioned yet in the field of bilingualism is the question of terminology which will be discussed in the next section.

1.3. Terminology

When studying bilingualism, one has to employ terms referring to the languages which a bilingual uses. In 1969 Lieberman wrote the definition of "mother tongue" which was adopted by the United Nations: "language usually spoken in the individual's home in his early childhood, although not necessarily used by him at present" (Lieberman, as cited in Romaine 1989/1995: 19). The term "mother tongue" has often been used by linguists in this meaning; however, it has some popular connotations.

First of all, the word "mother tongue" brings into mind the notion of mothers who pass the knowledge of their language or several languages on to their children. However, in some communities in the Vaupes (area of Colombia) and some bordering regions of Brazil, it is the fathers who convey their language to the

children. Suzanne Romaine (1989/1995) explains that these groups are patrilineal. In these communities, marriages are exogamous, i.e. one must marry a person from another language group. Husband and wife may not speak each other's languages, but they understand one another being passive bilinguals and speaking their own languages. Their children may master both languages, but consider the father's language as the primary one.

Secondly, "mother tongue" is often thought as the first language to have been learnt. However, there can be counterexamples. Romaine refers to Malherbe who "notes that it is quite common to find white infants in South Africa who grow up with Zulu nannies and can speak Zulu before they can speak either English or Afrikaans" (Malherbe, as cited in Romaine 1989/1995: 20).

Quite often, "mother tongue" is associated with the language spoken at home. However, a Pakistani, who speaks Panjabi at home, will name Urdu, the official language of Pakistan, as their mother tongue (Saifullah-Khan, as cited in Romaine 1989/1995: 21).

Another belief concerns competence, namely that a "mother tongue" is the individual's best known language. However, a bilingual may master better that language in which they were schooled and claim another language as the mother tongue because they "feel[s] a stronger affective attachment" (Romaine 1989/1995: 22) to the language spoken at home.

To avoid misunderstandings with the word "mother tongue", such terms as "first language", "second language" or "native language" are often used. Nevertheless, Uriel Weinreich avoids them, claiming that from the structural point of view, the question of the order of the language acquisition is irrelevant. What is important is which language serves as the source or model and which is the recipient or replica in the process of interference. Thus he employs the terms "source language" and "recipient language" respectively (Weinreich 1953/1970: 74).

Sometimes we find the term "dominant language". This concept evokes the reference to high language competence of a bilingual individual. The dominant language is not necessary the first language learnt. However, in the situation where a speaker constantly uses several languages it could be difficult to determine the dominant one (Baetens Beardsmore 1982). At the same time François Grosjean applies the term "dominant language" or "majority language" for the language which is considered more prestigious than the other one and which is spoken by the group who dominates the country in political, cultural and economic spheres (Grosjean 1982: 120).

Another suggestion was made by Dodson in 1981 who proposed the concept of “preferred language” which accentuates the importance of the given language across time admitting the individual’s life experience and subjective self-assessment of language competence (Dodson, as cited in Baetens Beardsmore 1982: 35).

Looking though these numerous definitions it appears evident that each term has some inaccuracy and that it is impossible to choose one term which could be applied to all domains of bilingual studies. It would be rational to specify in cases where it is needed which term will be employed and whether the author intends to emphasize with its use the language first spoken, the highest competence in the language or personal preference to the language, etc.

The difficulty of the terminology partially comes from the fact that bilinguals are often compared with monolinguals. Some of the applied terms like “mother tongue”, “dominant language” or even “first language” reflect the idea that there is one language learnt from infancy on and which remains dominant the whole lifetime. This situation could be accurate for monolingual speakers; however, bilinguals could grow up with two mother tongues or master the second language later in life and due to its daily usage would have two equally dominant languages. The comparison of bilinguals with monoglots also persists in other fields of bilingual studies. For example, all the schemes of bilingual speech production are based on those of monoglots; bilinguals’ language competence is often compared with competence of monolingual speakers of each language.

As an illustration, numerous studies of aphasia in polyglots raised two perplexing issues: 1. Are the brains of bilinguals different from those of monolinguals? 2. Do bilinguals use their right hemisphere more than monolinguals? Neuroanatomical and neurophysiological approaches show no valid experimental evidence of different language lateralization in bilinguals and monolinguals. The differences are of a quantitative nature rather than of qualitative, that is to the extent bilingual and monolingual speakers use different cerebral mechanisms such as metalinguistic knowledge, linguistic and pragmatic competence (Paradis 2004, Gómez-Ruiz 2010). Another distinction between bilinguals and monolinguals is in their language choice. Both monolingual and bilingual speakers have a range of language varieties which they apply in different social situations. There are different ways of expressing the same idea. For example, you might say, "Could you please close the window?", "Do you mind closing the window?", "Close the window", "It's a bit cold in here", etc. "Not only can bilingual speakers, like their monolingual counterparts, choose among different varieties of a language but, when speaking to other bilinguals, they can also choose between two languages. Whereas a

monolingual can only switch from one language, a bilingual may change varieties in one language, change languages, or do both" (Grosjean 1982: 128). François Grosjean describes four factors which may influence a language choice (participants, situation, content of discourse and function of interaction may influence a language choice) and different combinations of which may explain why a speaker could say "Close the window" or "Could you please close the window" in this or that situation.

The research of a "perfect bilingual" used to chase after ideal mastery of both languages so that a bilingual speaker could be indistinguishable from monoglots of each language. Consequently, there should be not any language alternation in the speech of a "perfect" or "ideal" bilingual. Uriel Weinreich says:

The ideal bilingual switches from one language to the other according to appropriate changes in the speech situation (interlocutors, topics, etc.), but not in an unchanged speech situation, and certainly not within a single sentence. If he does include expressions from another language, he may mark them off explicitly as "quotations" by quotation marks in writing and by special voice modifications (slight pause, change in tempo, and the like) in speech. There is reason to suspect that considerable individual differences exist between those who have control of their switching, holding it close to this ideal pattern, and those who have difficulty in maintaining or switching codes as required (Weinreich 1953/1970: 73).

Weinreich's remark on perfect bilinguals has been criticized since then. "The bilingual is not two monolinguals in one person" says Grosjean (1989) and Pohl (1965) states that it is absurd to discuss a perfect bilingual just as a perfect monolingual.

The researchers' attention, thus, turned to bilinguals as such. The typologies, illustrated in the previous section, showed that a bilingual speaker is now regarded in a continuum of bilingualism and/or bilinguality—as one bilingual among other bilingual or even monolingual speakers; there is no more need in research of "perfectness".

Another factor in the gradual rejection of a vision of a bilingual as a person who is in effect two monolingual speakers in one has been the focus on the phenomenon of code-switching, whereby a single speaker transitions from one language to the other, and back again, very swiftly. It is a phenomenon which, while presumably as old as bilingualism itself, has only come under the

scrutiny of researchers in recent times. The next chapter will examine its appearance in studies of bilingualism, up until the more recent 21st century perspectives, and will set out how this author will use the concept in the context of the research project described here.

Chapter 2.

Language alternation

Globalization and the development of information technologies have increased contact among people all over the world and thus, of course, levels of language contact. Communication in all its forms brings with it different forms of language alternation in everyday life. Matras defines “[t]he alternation of languages within a conversation” as a code-switching (Matras 2009: 101). Many other terms are in use to designate language alternation according to the discipline which studies it and the environment where it occurs, such as code-switching, interference, etc. for language alternation in oral communication and language-switching in the field of L2 written production. The present chapter will look both at spoken code-switching and language alternation in writing, exploring these phenomena and giving an overview of existing research on these topics.

The section 2.1.1.1. will open this chapter with a brief historical overview of the development of the research on oral code-switching to show diachronically the different stages of how this phenomenon was approached. The next section 2.1.1.2. will define code-switching and make the distinction between code-switching, borrowing, interference and code-mixing; it will bring into light the origin of the term code-switching as well. The section 2.1.2. will view code-switching from the grammatical, psycholinguistic and social perspectives providing the major theories elaborated in these three respective fields; the models briefly mentioned in the historical overview (2.1.1.1) will be also studied in detail in the section 2.1.2.

The next big sub-section 2.2. will focus on language alternation in writing. It will illustrate the psycholinguistic and educationalist perspective (section 2.2.1.) exploring such fields of research as literacy of bilingual children (section 2.2.1.1) and second-language composition (section 2.2.1.2). Then, section 2.2.2. will present a sociolinguistic perspective as well, discussing such topics as code-switching in literature (section 2.2.2.1) and code-switching in electronic writing (section 2.2.2.2.). The closing section of this chapter 2.2.2.3. will introduce a multimodal frame of analysis elaborated by Mark Sebba in 2012. As things currently stand, this is the only model conceived for the multilingual texts.

2.1. Spoken code-switching

2.1.1. Foundational studies

2.1.1.1. Historical overview of code-switching

The phenomenon of code-switching has been known since the early 20th century. In 1913 a French linguist Jules Ronjat published his observations of his bilingual son - *Le développement du langage observé chez un enfant bilingue*. This was the first widely acknowledged recorded data on bilingualism. However, code-switching is a rather new phenomenon in linguistics. At the beginning of the 20th century, the greatest interest in bilingualism was shown by psychologists, who were trying to describe the connection between bilingualism, and therefore the code-switching which accompanies such bilingualism, and the mental skills and defects of the speakers. An important place in code-switching research is occupied by Uriel Weinreich's *Languages in Contact* (1953). He was interested in the effects of language contact on languages. He describes three types of bilingualism: coordinate, compound and subordinate. "Weinreich's description of switching codes suggested that bilingual individuals possess two separate linguistic varieties, which (ideally) they employ on separate occasions" (Nilep 2006: 5). The frequent use of code-switching was regarded as defective and was opposed to that of an "ideal bilingual". There was a notion of imperfect language learning. In the early literature code-switching was characterized as random, but now many linguists have come to the conclusion that it is a rule-governed phenomenon. However, there are still some discrepancies on the nature of the constraints.

In the second half of the 20th century the perspectives on code-switching began to change somewhat. One of the influential works on code-switching appeared in 1972 by Blom and Gumperz. They described the use of Bokmål and Ranamål in a small village in northern Norway called Hennesberget, showing the importance of social factors on the example of the use of these two Norwegian dialects. Blom and Gumperz introduced two types of code-switching, namely metaphorical, when the speakers change the language to produce some communicative effects on the listener, and situational code-switching, when there is a change of a language due to the change of a social setting. In 1982, Gumperz developed another concept of "conversational code-switching" which includes several conversational functions such as quotation, addressee specification,

interjection, reiteration, message qualification, and personification vs. objectification. His research inspired many subsequent linguists to explore and develop his ideas. Another important contribution to the study of code-switching in terms of social motivations was made by Myers-Scotton who built on Gumperz's situational and metaphorical code-switching her own model. In 1983 she published *The negotiation of identities in conversation: a theory of markedness and code choice*, where she proposes the Markedness Model. According to Myers-Scotton speakers choose languages to index their social roles that she calls a set of Rights and Obligations. Since 1970s there have been many studies producing various approaches, models and constraints on code-switching. In section 2.1.2. I will give a fuller presentation of the main theories of spoken code-switching.

2.1.1.2. Terminology and definitions

Originally, the term “code-switching” had nothing in common with linguistics. It was taken from communication technology and “refers [...] to a mechanism for the unambiguous transduction of signals between systems” (Gardner-Chloros 2009: 11). In 1950s R. M. Fano used the notion of “switching codes” in his article *The information theory point of view in speech communication* referring to the mental mechanism of decoding different accents (Alvarez-Cáccamo 2002: 31). Later in 1960s Hans Vogt employed the term “code-switching” designating bilingual behaviour. And since 1970s it has been used in a sense of “linguistic action (language alternation in speech, including grammar and lexicon)” (Alvarez-Cáccamo 2002: 33).

Once having started to study code-switching it is easy to notice that there are many confusions beginning with the spelling and terminology and ending with discrepancies in general constraints and “matrix language”. For example, in contemporary literature several possible writings can be found: code-switching, codeswitching and code switching; this inconsistency of the spelling emphasizes the lack of the standardized terminology, where by standardized I mean the terminology which is approved and used uniformly by scholars. One of the first general definitions of code-switching was given by Heller: “the use of more than one language in the course of a single communicative episode” (Heller 1988: 1). Many linguists give a definition of code-switching similar to Heller's. For example: “the use of two varieties in the same conversation” (Myers-Scotton 2006: 239); or “code-switching is the alternation of two languages within a single discourse, sentence or constituent” (Poplack 1979/1980/2000: 224). However, others explain code-switching in a descriptive way. For instance, Brian Hok-Shing Chan writes in his

book *Aspects of the Syntax, the Pragmatics, and the Production of Code-switching* that it “is a communicative resource available to bilinguals which is used to perform, presumably more effectively, various common purposes in everyday conversation. If pure languages are subject to a set of pragmatic principles, code-switching may well be governed by the same set of principles too” (Chan 2003: 259).

Though, to understand what code-switching is one should firstly ask oneself a question: what does “code” mean? It can be interpreted in two ways. The first one is that a code represents a language. It is a neutral term referring to a dialect, a variety of language or a language which does not have any pejorative or political connotation. For example, studies on code-switching constraints and grammar imply pre-existence of at least two linguistic varieties which the bilingual speakers alternate in communication (Pfaff 1979; Clyne 1980). Another approach is the one proposed by Jakobson and examined by Alvarez-Cáccamo (1998) in his quest to define “code” and code-switching: a language is not a code, but it is composed of codes. Speakers use codes to transmit a message and listeners use other codes to interpret it. A speaker can change a code to signal a change in a conversation, for example emphasizing that another participant has joined the discussion; or trying to change the meaning of the conversational output; or showing the change of their intention. “In other words, switching codes is a means by which language users may contextualize communication” (Nilep 2006: 17).

Another moot point in code-switching studies is terminology and in particular the way that the term code-switching enters into opposition, or not, with other notions such as code-mixing and borrowing. This is no easy task. As Eastman wrote: “Efforts to distinguish code-switching, code-mixing and borrowing are doomed” (Eastman, as cited in Gardner-Chloros 2009: 10).

The first distinction to be made is between code-switching and borrowing. Li Wei and Melissa G. Moyer say that borrowings “are the outcome of a historical process, and are used by monolinguals as well as bilinguals” (Li Wei and Moyer 2008: 60). This means that the speaker’s bilingualism can be hardly judged by the presence of borrowings in a discourse. Almost all the scholars are agreed that some borrowings or loans (these terms are often interchangeable) begin as code-switchings; but with time they become generalized and assimilated in the borrowing language; as they become a part of the lexicon they are used by monolingual speakers as well as by bilingual ones. There is a traditional definition: “a borrowing is a word or short expression that is adapted phonologically and morphologically to the language being spoken” (Grosjean 1982: 308). On the one hand, borrowings were studied from the point of view of degree of phonological and morphological

integration (Bloomfield 1933, Haugen 1950, Weinreich 1953, Poplack and Sankoff 1984, and others). Haugen (1950) provides the following classification of loans. On the phonemic level one can find unassimilated loans, hybrid loanwords/loanblends (partially assimilated), and wholly assimilated loans. For example, the English word "hamburger" /'hæm,bɜ:gə/ is totally assimilated in French /'ãburgœR/ and in Russian /'gamburger/. As for the English word "chewing-gum", it is only partially assimilated in French—the first phoneme /t/ of the word /'tʃu:ŋ.gʌm/ is omitted in French adapting the pronunciation of the letters "ch" to the French rules of pronunciation /ʃ/; however, the end of the word "-gum" is pronounced as in the original English word, rather than according to the French rules. And finally the word "restaurant" is totally phonemically unassimilated in German. On the morphemic level there can be total, partial, or no distinction. For example, the English verb "to check" was adapted to French adding to the English morpheme "check" the French ending of the infinitive "-er": "checker". Whereas the English word "shopping" is employed in French without any modification or the French word "déjà-vu" in English. As for the total morphemic assimilation, an example of the English verb "mock" can be given. It goes back to the Middle French "mocquer". And the third type of this classification is loanshift, also known as loan translation or calque. For instance, the English word "skyscraper", French "gratte-ciel" and Russian "небоскрѣб".

On the other hand, there are numerous attempts to contest the distinction between code-switching and borrowing which is based on the degree of assimilation (Mackey 1970; Poplack and Sankoff 1984). Myers-Scotton summarizes and gives three contradictions: first of all, degree of integration is not a concept, but a gradient that can give only a continuum for further evaluation; secondly, there are many borrowings which show very little phonological assimilation; and thirdly, not all the borrowings are assimilated both phonologically and morphologically; some can show either phonological or morphological integration (Myers-Scotton 1988/2000: 143–144). She proposes to distinguish borrowing and code-switching by a social content. Those elements which are socially significant in a conversation represent code-switchings, those which are not are borrowings. Nowadays many scholars incline to the conclusion that it is almost impossible to distinguish borrowing from code-switching by its linguistic form. "It is the nature of the sociolinguistic contact which prevails at the time when an element is switched or borrowed which determines in what manner it is adapted or altered" (Li Wei and Moyer 2008: 60).

The second distinction is between code-switching and interference. The later term seems to be very general. Many definitions which are rather vague can be

found. For example, according to Haugen, interference is the overlapping of two languages (Poplack and Sankoff 1984: 103), or Hugo Baetens Beardsmore says that it is “the use of formal elements of one code within the context of another, i.e. any phonological, morphological, lexical or syntactic elements in a given language, that could be explained by the effect of contact with another language” (Baetens Beardsmore 1986: 45). These definitions do not show any details, but they only state that interference is a phenomenon of language contact. The following questions can be brought up: When does interference become integration? What is the difference between interference and code-switching? To answer the first question, it is necessary to define integration. “Integration [is] the regular use of material from one language in another in the form of an established loan [...]” (Baetens Beardsmore 1986: 50). I.e. integration is equated with borrowing. When an element of one language incorporated into another ceases to be interference it becomes integration; and in its turn, integration differs from interference by the phonological and/or morphological shape of the element and its social integration. Consequently, the difference of both interference and code-switching from borrowing is recognized by the degree of integration. However, the distinction between interference and code-switching is not clear. One of the prominent ideas to distinguish these two terms was proposed by François Grosjean. He suggests that interference is a subconscious phenomenon and that a bilingual is not aware of producing something that a monolingual does not understand. Whereas code-switching occurs voluntarily, when it is meaningful to the participants of the communication. It can be observed in a conversation when a bilingual talking to a monolingual tries to avoid code-switching, but cannot control the use of unconscious interference which can lead to the misunderstanding (Grosjean 1982: 299). Another criterion of distinction between code-switching and interference is in the use of morpho-syntactic elements alone by the latter, whereas they are employed together with lexical elements by the former. Pieter Muysken explains:

in the case of *code switching* or *code mixing*, there is lexical material from two languages present in the clause, in addition to morpho-syntactic structure from both languages. Interference – in Weinreich’s (1953) sense of mutual influence between a bilingual’s two languages – involves morpho-syntactic structure from two languages, but lexical material from only one of them. (Muysken 2004/2006: 147).

The third distinction that can be made is between code-switching and code-mixing. This is the area of most confusion. Some scholars use code-mixing for what others refer to as switching and vice versa. Sometimes mixing is used to denote

intra-sentential alternations while switching is for inter-sentential shifts (Bokamba 1988; Matras 2009: 101; Niemiec 2010). Meisel (1994) refers to code-mixing as a fusion of language grammars, whereas he defines code-switching as a pragmatic skill with which the speaker can select the language according to the interlocutor, topic of the conversation, context of the situation, etc. For Muysken, code-mixing is a generic term that covers

- insertion of material (lexical items or entire constituents) from one language into a structure from the other language,
- alternation (that he calls code-switching) between structures from languages,
- congruent lexicalization of material from different lexical inventories into a shared grammatical structure. (Muysken 2000: 3)

In a more recent work Yakpo employs the term code-mixing for the use of non-native elements on a societal level with a strong presence of congruent lexicalization: "[...] the entrenched and systematic pattern of code-switching that I refer to as code-mixing in this article, and language mixing, i.e. the development and stabilization of a mixed language." (Yakpo 2015: 263)

Quite often knowing the confusion of terminology, linguists explain in their publications the terminology that they will use further. For instance, Kay McCormick in the book *Language in Cape Town's District Six* writes:

I refer to language varieties and to switching between them as 'codes' and 'code-switching' only where the varieties and the switching between them seems to have some significance for members of the speech community. Where switching is not significant, I refer to it as 'language switching'. [...]. I use 'switching' to refer to the alternation of phrases or longer chunks in one code or language with those in another, whereas 'mixing' is used to refer to the incorporation of single lexical items from one language into phrases in the other [...]. (McCormick 2002: 89).

However, what is "language switch" for McCormick (2002) would be a mixed "medium" for Gafaranga & Torras (2002), and what the latter entitle 'mixing' would be called 'borrowing' or 'nonce loan' by Poplack (1980). As for Pfaff (1979) and Romaine (1986), they employ 'mixing' to cover both 'borrowing' and 'code-switching'. There is no end to the attempts to define code-mixing and to distinguish it from the other terms of language contact. However, the overlap of referents and concepts results in the greater confusion.

That is the reason why some authors adhere either to the term of code-switching using it in a generalized meaning:

[...] many of the examples set forth might be labeled as codemixes, but we will keep the term code-switching for practical reasons. Bhatia & Ritchie also admit the controversy of this distinction [code-mixing vs code-switching] and the difficulty to draw a clear line between the two terms, and others (Evelyn Hatch 1976, Rodolfo Jacobson 1990) doubt the usefulness of this distinction. We will use the term code-switching to refer to the alternating use of two or more languages in the same discourse at the word, phrase or sentence level (Montes-Alcala 2000: 12).

or to both terms – code-switching and code-mixing – considering them synonymous:

In the absence of a general consensus, I shall be using the two terms – ‘codeswitching’ and ‘codemixing’ – interchangeably (Matras 2009: 101).

Having tried to convey the main points of a vast debate on terminology, I hope that this section has been sufficiently explicit on the difficulties of the terminological issue and the ways that researchers found out to be transparent in their works and understandable for their readers. This debate will be finished here allowing another equally broad issue to be opened, namely theoretical approaches to code-switching.

2.1.2. Theoretical framework

The phenomenon of code-switching can be viewed from within different perspectives. For example, “Cognitive approaches generally focus on code-switching as an index of a bilingual’s ability to selectively inhibit or activate a language” (Isurin, Winford and Kees de Bot 2009: 189). Psycholinguists try to understand the mechanism of the switch. Thus Walters explains code-switching taking into account both sociopragmatic and psycholinguistic factors:

Deliberate, L1-to-L2 codeswitching is accomplished by taking relevant social identity and contextual (setting, topic, and participant) information and applying it to a speaker’s intentions. That information is largely copied. When the speaker also intends to codeswitch (for ethnic solidarity, pragmatic focus or other sociopragmatic reasons), the copying mechanism gives way to variation (usually in the speech act component), permitting access to the language choice module and entry of the L2 word or expression” (Walters 2005: 204).

Linguistic approaches are interested in the knowledge of the language which is necessary to systematically make a switch between two languages. In particular, grammatical approaches focus on the language system, in our case on the structural aspects of code-switching. And sociolinguistics aims to explore why and under what conditions bilinguals code-switch. This section will try to show the phenomenon of code-switching from the three perspectives: grammatical, psycholinguistic and sociological.

2.1.2.1. Grammatical aspects of code-switching.

The grammatical approach has been one of the most prolific fields of study in code-switching. Some scholars (Weinreich 1953/1970, Green 1998, etc.) considered bilinguals as individuals operating with two independent linguistic systems each with its monolingual grammar and lexicon; code-switching results from the interaction of these systems. There have been many attempts to create universal grammars and rules predicting code-switching. Different constraints have been elaborated defining a grammatical and ungrammatical way of switching; many theories have been proposed to explain code-switching; several types of code-switching have been distinguished, etc.

A few grammatical assumptions which form the core of grammatical approaches to code-switching will be shown in this section.

Firstly, from grammatical perspectives code-switching can be viewed as

- single-word and multi-word CS
- intra-sentential (“a switch of codes within a simple utterance” (Wardhaugh 2006: 108)) and inter-sentential code-switching (“it can occur between sentences” (Ibid.: 101)).

Secondly, there are several grammatical constraints on code-switching; four of which will be presented below. Up until now, scholars have not come to an agreement concerning the grammatical constraints on code-switching. One can find some counter examples to all the rules in a different language pair. The issue of constraints on code-switching is still open.

The free morpheme constraint and the equivalence constraint

In the end of the 1970s Shana Poplack developed the two constraints, namely *the free morpheme constraint* and *the equivalence constraint*. Her study was based

on a large corpus collected in the New York Puerto-Rican community. It was one of the most influential works.

The free morpheme constraint was first formulated as the following: “Codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme” (Poplack 1979/1980/2000: 227)

In 1981 this statement was elaborated to a more precise definition:

A switch may not occur between a bound morpheme and a lexical item unless the latter has been phonologically integrated into the language of the bound morpheme (Sankoff and Poplack, as cited in Chan 2003: 10).

That means that ‘flicheando’ (flip-ING) is a possible English/Spanish switch. The stem ‘flip’ is pronounced in the same manner in English and in Spanish. Both the stem and the suffix are pronounced according to Spanish phonological rules. However, the following example is unacceptable: ‘eatiendo’ (eat-ING). According to the free morpheme constraint the lexical item ‘eat’ should be phonologically integrated into the language of the suffix ‘-iendo’; ‘eat-iendo’ should be pronounced with regard to the Spanish phonological rules: /eatjendo/ or reduced to /jatjendo/ but not /itjendo/. Besides, according to Poplack, idiomatic expressions like ‘Cross my heart and hope to die’ also behave as bound morphemes and are used monolingually.

In 1979 the equivalence constraint was stated as:

Code-switches will tend to occur at points in discourse where juxtaposition of L1 and L2 elements does not violate a syntactic rule of either language, i.e. at points around which the surface structure of the two languages map onto each other (Poplack 1979/1980/2000: 228).

Later it was reformulated in Sankoff and Poplack (1981) in a clearer way:

The order of the sentence constituents immediately adjacent to and on both sides of the switch point must be grammatical with respect to both languages involved simultaneously. This requires some specification: the local co-grammaticality or equivalence of the two languages in the vicinity of the switch holds as long as the order of any two sentence elements, one before the switch point and one after the switch point, is not excluded in either language (Sankoff and Poplack, as cited in Chan 2003: 12).

The equivalence constraint enables a switch at points in a sentence where the word order in the two languages is the same. A classic example is “I told him that pa’que la trajera ligero” (I told him that so that he would bring it fast/ (Yo) le deje eso pa’que (él) la trajera ligero) (Ibid.: 12). The surface structure around the switch point, which is in between ‘that’ and ‘pa’que’, is alike in both languages. However,

there are a number of counter examples often taken from such language pairs as: English/Japanese, Swahili/English, French/Moroccan Arabic, etc. For example, if there is a different order of adjective and noun in a NP in the involved languages there can be no switch. The equivalence constraint may lead to a conclusion that the more alike the structures of the two languages are the more switching points they may allow. The more dissimilar they are, the fewer possibilities of the switching there are.

Suzanne Romaine expounded several ideas explaining incompatibilities of the equivalence constraint:

One may be that individual language pairs in contact impose specific additional rules on switching. Or it may be that the equivalence constraint has not been formulated precisely enough in the first place. Notion of what is grammatical switching will need careful evaluation. Although most researchers stress the grammaticality of the majority of bilingual utterances, they assume that the grammatical norms of the two languages in isolation provide the basis for determining what is grammatical. But varieties of two languages in contact for a long time may become quite different from varieties of those languages spoken elsewhere. This has happened to the Austronesian languages of Papua New Guinea, as we saw in Chapter 1. Another factor which may affect switching behavior is the attitude of the participants concerned. In some Spanish/English bilingual communities, favorable attitudes to switching have been observed, which has led to an increase in the use of switching as a marker of ethnic identity. (Romaine 1994: 60)

The clitic constraint

Another constraint was proposed by Timm (1975) and Pfaff (1979), i.e. *the clitic constraint*. Pfaff says that the clitic subject or object pronouns should be “realized in the same language as that of the verb to which they are cliticized, and in the position specified by the syntactic rules of that language” (Pfaff, as cited in Atawneh 1992: 232). However, Meisel (1994) and Gardner-Chloros (2009) reported cases of code-switching between the verb and clitic pronouns. The violation of this constraint is shown in the following French–Alsatian example given by Gardner-Chloros:

Il koch güet

He cooks well (Gardner-Chloros 2009: 95)

The government constraint

In 1986 Di Sciullo and her colleagues (Muysken and Rajendra) developed a constraint which explains code-switching from the point of view of the government relations. *The government constraint* states that the switch can occur only between the elements which are not related to government, i.e. no switch can happen between a governor and the governed element. However, it cannot explain the common switch between verb and adverb, or subject noun phrase and the main verb phrase: “La plupart des canadiens scribono ‘c’”/ The majority of (the) Canadians write ‘c’ (the French–Italian example taken from Clyne 1987/2000: 264–265).

The Matrix Language Frame model

Another theoretical framework, which should not be left unmentioned, particularly in light of the fact that it will be employed in the analytical work to which this PhD will later turn, is the *Matrix Language Frame* (MLF) model. It was elaborated by Myers-Scotton in 1993 and differs from all the previous constraints by its complicated hierarchical framework. This model is based on the idea of insertion. There is a dominant language that provides a frame or matrix in which elements of the other language can be inserted. The participating languages are called the matrix language (ML) and the embedded language (EL) respectively. Myers-Scotton states the two principles of the model as the following:

The Morpheme-Order Principle: In ML + EL constituents [mixed constituents] consisting of singly-occurring EL lexemes and any number of ML morphemes, surface morpheme order (reflecting surface syntactic relations) will be that of the ML.

The System Morpheme Principle: In ML + EL constituents [mixed constituents], all system morphemes which have grammatical relations external to their head constituent (i.e., which participate in the sentence's thematic role grid) will come from the ML. (Myers-Scotton 1993, as cited in Myers-Scotton 2001: 35)

The MLF model is built on two oppositions. The first one is that of ML versus EL. Myers-Scotton says that ML can be identified only in opposition to EL due to the different roles that they bear and which are stipulated in the two principles of the model. She also describes ML as a more unmarked choice being the most appropriate in the interaction type where code-switching appears. The second opposition is between content versus system morphemes with later division of system morphemes into three subcategories as we shall see below.

Since the elaboration of the MLF model, Myers-Scotton with her colleague Jake have made a number of publications attempting to resolve controversies and defend the MLF model's viability. One of the subjects to modification has been the criteria for determining the ML. The ML was originally defined by the quantitative criterion. The language which supplies the greater number of morphemes than the other one is the ML. Gardner-Chloros and Edwards referring to Bentahila and Davis suggest that according to this criterion bilingual communications will change ML several times (Gardner-Chloros and Edwards 2004: 11). However, which segment of discourse can be considered sufficient within which to count morphemes? It is said that the sample should be more than one sentence, but the more exact size is unclear: "How large is 'large enough' is an unresolved issue" (Myers-Scotton, as cited in Gardner-Chloros 2009: 101). Later this criterion was abandoned.

Recently, the ML has been determined by the morpheme-type criterion. The ML provides the majority of system morphemes, whereas the EL supplies a number of content morphemes. However, the differentiation between content and system items was rather problematic. After successive amendments Myers-Scotton and Jake (2000b) proposed the 4-M model; preserving the distinction between content versus system morphemes, the 4-M model subdivides system morpheme into three categories: early system morphemes and late system morphemes, which in their turn are separated into bridge late system morphemes and outsider late system morphemes. The system morphemes are related to the process of language production and reflect different activation stages; they can be activated either at the lemma level in the mental lexicon or formulator. "However, this model is not so much a classification of morphemes but rather constitutes a hypothesis on how morphemes are conceived in linguistic competence and accessed in production" (Myers-Scotton 2001: 41).

Regarding the MLF model and all the amendments provided by Myers-Scotton and Jake, Gardner-Chloros says: "The MLF is an evolving, comprehensive, ad hoc attempt to describe the grammatical facts of code-switching, but it has proved difficult to maintain its basic tenet, that there is a single underlying grammar to all code-switching productions" (Gardner-Chloros 2008: 67).

This section showed several grammatical approaches to code-switching; these were the attempts to find the grammatical constraints of universal grammar of code-switching. However, all the above-mentioned examples could not be universally applied facing counterexamples in works/corpora of different authors. In the article concerning assumptions behind grammatical approaches to code-switching Gardner-Chloros and Edwards come to the following conclusion:

To sum up, although syntax plays an important role in CS, it cannot be assumed a priori that the constructs of syntacticians are the best means for characterising the processes of performance data such as CS. The possibility of throwing light on this question depends partly on whether or not it is right to assume that all bilinguals alternate in some meaningful way between two clearly distinguishable sets of rules – and this is a question which manifestly cannot be decided by grammatical analysis alone (Gardner-Chloros and Edwards 2004: 16).

They suggest that researchers would probably do better to turn to the variability of bilingual grammars than to focus on the research of a universal one.

With this Gardner-Chloros and Edwards's conclusion on grammatical approaches I would like to shift your attention to another model which regards code-switching from the psycholinguistic point of view.

2.1.2.2. Psycholinguistic model of triggering.

The first of a series of publications on triggering appeared in 1967 announcing one of the most famous psycholinguistic models of code-switching still used today. Michael Clyne proposed a psycholinguistic principle according to which words which are part of two languages facilitate non-intentional code-switching. Triggering is a linguistic phenomenon that promotes the activation of one language and the deactivation of the other. “‘Triggering’ [is] where an item of ambiguous affiliation (that is, one belonging to the speaker's two systems) triggers off a switch from one language to another” (Clyne 1987/2000: 262). Trigger-words are words in between the two language systems that influence the speaker to continue the sentence in the other language; they can belong to the two languages or they can be compromise forms between them. Clyne gives a classification of trigger-words: proper nouns (names of people and places), lexical transfers (elements of one language that have also become part of the other language), bilingual homophones (words of two languages that sound [almost] the same), and discourse markers (elements that help to organize the conversation). However, sometimes the nature of the triggering is unclear, in particular when the languages are close.

There are four types of triggering according to Clyne. The first one is consequential triggering that follows a trigger-word. In Example 1 below the prepositional phrase "kop of" could be either a phonologically integrated form [kɔp ɔf] of lexical transfer or a semantic transfer from English, which triggers the switch into English following this prepositional phase.

Example 1:

Ik hebt een kop of tea, tea or something / I had a cup of tea, tea or something
(Clyne 1987/2000: 273).

The second type is anticipational triggering that precedes a trigger-word. Example 2 shows that the noun "bridge" triggers the switch into English in the preposition and the article, which precede it.

Example 2:

Hij staat *on the* bridge / He is standing on the bridge (Clyne 1987/2000: 273).

The third is "sandwich words": a combination of the consequential and the anticipational triggering, where a switch is between the trigger-words. Example 3 illustrates a "sandwich" switch which occurs between the two English proper names.

Example 3:

Ich werde mir "Grange Hill" *and* "The Changes" ankucken. / I'll watch "Grange Hill" and "The Changes" (Saunders 1988: 88).

And the final type is contextual triggering when a switch occurs due to the context of the situation and not because of trigger-words.

Example 4:

Frank: "Mum, can you speak lots of German?"

Mother: "A fair bit. I speak a lot to Ilda's mother."

Frank: "Why a lot?"

Mother: "She doesn't understand English. You've heard me speaking German to Ilda's mother, haven't you?"

Frank: "Hm. And we play with Ilda."

Mother: "Hm. Do I speak good German then?"

Frank: "*Ja.*" (*Yes.*)

(Saunders 1988: 89).

Recently Broersma and De Bot have developed an adjusted triggering theory (Broersma and De Bot 2006). Keeping the general idea proposed by Clyne that trigger words may facilitate code-switching, they scrutinized the level where triggering may take place. Broersma and De Bot suggest that triggering occurs on the lemma level rather than in a surface structure. Empirical tests of both theories

proved that the presence of trigger words significantly enhanced the chance of code-switching as well as confirmed the authors' hypothesis. However, both versions of triggering theory explained in their turn occurrences of code-switching which the other version could not account for. Thus, Broersma and De Bot stated in their conclusion that "the adjusted triggering theory accounts only for some instances of codeswitching. Many forms of codeswitching, with different degrees of intentionality, can be envisioned. They may reflect different processes, at different levels in the speech production process. Triggering at the lemma level is just one of these processes" (2006: 12).

Broersma and De Bot's work showed the viability and still vivid interest in the triggering model proposed almost 50 years ago. They also left the area open for further research on code-switching in psycholinguistics.

The next section will give an overview of a much more prolific field of research on code-switching. The sociolinguistic perspective provides a wide range of studies focusing on the social and pragmatic functions of code-switching which will be discussed further.

2.1.2.3. Social aspects of code-switching.

Auer says "Grammatical restrictions on code-switching are but necessary conditions; code-switching is not merely a matter of linguistic well-formedness – it also has communicative content left unexplained by the analysis of the syntactic surface constraints" (Auer 1984: 2). A grammatical approach is not enough to answer the everlasting question: why does code-switching occur? Psycholinguistics does not treat the communicative content of code-switching either. It is the field of sociolinguistics which attempts to do it.

Code-switching is approached from two perspectives in sociolinguistics: macro-linguistic and micro-linguistic. From the macro-linguistic perspective code-switching has been examined on the community level where "High" and "Low" language varieties were introduced by Ferguson in a situation of "diglossia" (Namba 2005: 2). Gumperz and Hernandez noted that "code-switching occurs whenever minority language groups come into close contact with majority language groups under conditions of rapid social change" (Gardner-Chloros 2009: 54).

However, the majority of scholars focus on the micro-linguistic approach studying code-switching on the level of interaction. Benjamin Bailey writes that "language is our primary semiotic tool for representing and negotiating social

reality, including social identity categories. Through talk we position ourselves and others” (Bailey 2007: 341).

We-code vs they-code

Gumperz (1982) asserted that code-switching is an outcome of an alternation between a *we-code* and a *they-code*. A *we-code* is attributed to a minority language and associated with the informal style; while a *they-code* is regarded to be the majority language, often being the official language of the country, associated with a more formal style. There is a symbolic difference between “we” and “they”. It can be seen through the language choice: “esi engrezi sikhi e te why can't they learn? (‘We learn English, so why can’t they learn [Asian languages]?’)” where Punjabi is a *we-code* and English is a *they-code* (Romaine 1994: 61). “We” and “they” are opposed to each other. However, the relationship between code and identity is more complex.

Sebba and Wootton (1998/2013) show an example with an unusual configuration of *we/they* codes. They studied British-born Caribbeans who speak London English and London Jamaican. They found that the both languages serve as *we-codes* since they characterize the speakers as “Black British”. Gardner-Chloros (2009: 57–58) also provides two other examples of a less common distribution of *we-/they-codes* quoting Sing (1983) and Meeuwis & Blommaert (1998). Singh shows the reversed *we-* and *they-codes*. He says that in India some speakers from a high social class may use English as a *we-code* and Hindi as a *they-code* to show that they belong to a privileged minority, and that they are not like other groups. While Meeuwis and Blommaert describe another situation where a code-switched variety itself can take on the value of a *we-code*. In the Zairian community in Belgium the two most widespread national languages (Lingala and Swahili) are spoken as code-switching being mixed with French. Lingala–French and Swahili–French can represent two independent varieties with all the functions ascribed to a language.

Situational vs metaphorical/conversational code-switching

As it has already been mentioned another influential dichotomy of code-switching was proposed by Gumperz and his colleague Jan-Petter Blom (1972/2000), who suggested a distinction of *situational* and *metaphorical code-switching*.

Situational switching occurs when there is a change in the social settings due to some extra-linguistic reasons like a change in interlocutor, context, topic, mode of

interaction, etc. There is a direct relationship between a code and a social situation. They also clarify that “language choice is never completely fixed”; it is a subject of alternation throughout the interaction. In addition, they state that “sociolinguistic variables must be investigated empirically. Furthermore, situations differ in the amount of freedom of choice allowed to speakers” (Blom and Gumperz 1972/2000: 126).

In contrast to situational switching, *metaphorical* language choice is hard to foretell and it is left to the speaker’s decision. Metaphorical code-switching relates to subject matters and personal intentions of the speaker. It refers to different communicative effects that the speaker intends to produce on the listener. Suzanne Romaine (1994/2000) shows an example when a lecture in a college is performed in the official language, but then the local dialect is used to encourage the discussion among the students; even if the speakers, the topic and the listeners have remained the same, the code-switching has changed the mood of the conversation.

To understand metaphorical code-switching, it is not enough to simply observe the language use; it is also important to have some background knowledge of the local culture.

The semantic effect of metaphorical switching depends on the existence of regular relationship between variables and social situations of the type just discussed. The context in which one of a set of alternates is regularly used becomes part of its meaning, so that when this form is then employed in a context where it is not normal, it brings in some of the flavor of this original setting (Blom and Gumperz 1972/2000: 127).

However, Blom and Gumperz were much criticized. They were said to show “idealized entities” (Nilep 2006: 8). It is rather difficult to identify a language choice as strictly situational or metaphorical. The features may overlap. For example, personal decisions and intentions that characterize metaphorical switching may influence the situation. In addition, situational code-switching is rarely found in practice, it is considered as an idealized form. In 1982 Gumperz introduced the notion of *conversational code-switching*, perhaps recognizing the imperfection of the division between situational and metaphorical switching. However, the description of conversational code-switching resembles that of metaphorical. Gumperz suggests six conversational functions of code-switching, noting that this list is not exhaustive. He proposes the following functions: quotation, addressee specification, interjection, reiteration, message qualification, and personification vs. objectification.

Discourse-related vs participant-related language alternation

Gumperz's works on code-switching inspired many subsequent scholars. Peter Auer (1988) proposed two other types of code-switching, namely *discourse-related* and *participant-related* language alternations. Discourse-related code-switching includes a topic change and a change in addressee constellation, while participant-related switching relates to the participant's language competence and participant's preferences.

If we compare participant and discourse-related language alternation we note that the main difference is the object of the signalling process. Whereas in the case of participant-related alternation, co-participants display or ascribe certain predicates to each other (competence, preference), they signal a change of conversational context in the case of discourse-related switching. (Auer 1988: 198–199)

Markedness model

In 1983 an influential work on code-switching in terms of social motivations was developed by Carol Myers-Scotton who proposed the *Markedness model*. She states that social factors influence the choice of language in code-switching context. A participant in the conversation chooses a language according to the set of rights and obligations (RO) showing his/her place in the existing social situation. Therefore, a speaker's choice can be interpreted with regard to their motivations which are linked to the speaker's vision of the appropriate set of RO in the situation.

The Markedness model is based on one principle and a set of maxims. The negotiation principle formulates the main idea of the model: "Choose the form of your conversational contribution such that it indexes the set of rights and obligations which you wish to be in force between the speaker and addressee for the current exchange" (Myers-Scotton 1988/2000: 137).

The unmarked choice maxim conducts the speaker to choose an expected language variety for the present situation. It implies that the speaker tries to establish the unmarked RO often negotiating the status quo. According to the Markedness model situational code-switching proposed by Blom and Gumperz refers to a movement from one unmarked choice to another.

Code-switching can itself be an unmarked choice where the two applied varieties are normally indexed in the current community. Myers-Scotton gives an example of educated bilingual speakers in Kenya and Zimbabwe who alternate between their first language and English in conversations with peers from the same ethnic group. "The two varieties are both indexical of positively valued identities,

but from different arenas, such as ethnic group membership and being part of an educated and/or urban elite." (Myers-Scotton 1988/2000: 146).

The second is a marked choice maxim. When the language choice is socially unexpected it demonstrates a marked code-switching. A marked choice redefines or violates the actual RO. Gardner-Chloros (2009: 69) gives the following example: switching to a vernacular variety while talking about home or family represents an unmarked choice; however, the same switch in a public speech will be a marked choice. The marked choice can be positive and negative. The former narrows the social distance "if it is indexical of a relationship of solidarity"; and the latter can increase the social distance "because it encodes anger or the desire to make a power differential salient (when it would not be salient ordinarily)" (Myers-Scotton 1988/2000: 149).

When a speaker does not know which choice is expected in a given situation, they can make an exploratory switching, which represents the third exploratory choice maxim. It can help to find out which language choice will suit the RO that they desire.

In the Markedness model speakers are represented as "rational actors". In any situation they can make a language choice which may either fit the current set of RO or disrupt it.

To say that rationality drives choices is simply to say that speakers seek to optimize their rewards and minimize their costs. In deciding to follow norms or not (i.e., to make the unmarked choice) speakers recognize that all choices impose costs and supply benefits to both groups and individuals. True, speakers who are behaving rationally are also behaving subjectively, and they are not necessarily successful in achieving their goals. However, what they are doing is recognizing a feasible set of choices and selecting the most feasible one; this, for them, is making the rational choice (Myers-Scotton 1999: 1270).

The Markedness model will close the list of sociolinguistic theories on code-switching presented here. The several approaches discussed in this section draw a general overview of socio-pragmatic explanations of language alternation including some of the most prominent works. Most of the above-mentioned models have been applied not only to oral but also to the written code-switching. The next chapter will deal with the issue of language alternation in writing which embraces interests of different disciplines.

2.2. Language alternation in writing

Since the 1970s, many linguists working in the field of bilingualism have been interested in language alternations, namely code-switching and related phenomena, such as code-mixing, interference, borrowing, etc. However, the majority of the studies have been focused on code-switching in oral speech, especially in its spontaneous usage. Many works describe methods and procedures for code-switching analysis, such as ways of conducting observations, interviews, spontaneous and semi-spontaneous conversations; means of recording audio and video and transcribing the data; not to mention all the theories we have just been investigating: Markedness Model, triggering, they/we-code, or such types of code-switching as situational and metaphorical switching, etc., all of which were elaborated on and for spoken code-switching. Most researchers who have shown an interest in written code-switching data have done so by trying to apply methods originally developed for spoken code-switching. Authors compare whether the tendencies of the written code-switching reflect those of the oral. Dawn Marley, in her analysis of code-switching on websites for the Moroccan diaspora says:

This small sample of forum discussions shows clearly that in this context code-switching seems to be used in very similar ways to those used in face to face (or telephone-mediated) conversations, often replicating patterns observed among immigrants and their descendants in Europe (Marley 2011: 51).

However, while some writing practices such as correspondence, literature, online communications, etc. could be easily studied from the perspective of oral code-switching, other writing types, where the interlocutors are not clearly identified such as advertisements, posters, signage, etc., require another approach highlighting the value of these literary genres. The recent Sebba's multimodal frame of analysis is the only model worked out specifically for multilingual texts, taking into account their diversity.

Talking about code-switching in writing, we tend to forget about the multidisciplinary nature of this phenomenon. It interests not only sociolinguistics but also education studies and psychology. This chapter will throw light on different perspectives of language alternation in writing, focusing on the four following topics: reading/writing of bilingual children, L2 writing, language alternation in literature and electronic writing; as well as it will explain the only approach to multilingual text analysis.

2.2.1. Psychological and educational perspective.

2.2.1.1. Literacy of bilingual children.

There are families where parents speak to their children in a language different from that of the society they live in. These children find themselves in a different environment when they go to school; they have the same conditions as the monolinguals, but they do not have the same cultural and linguistic background. Do bilingual children need special programs of education or individual language courses? Many studies focus on the education of bilingual children. Krashen (1993/2004) states that reading is a powerful tool which provides a foundation to reach a higher levels of proficiency both of L1 and L2. Datta (2000) provides five approaches to literacy, which were earlier proposed by Baker (1993) according to the role of reading in the multilingual education that will assist language-learning and develop cultural identity. However, not only reading, but also listening, speaking and interpreting are of great importance for bilinguals to enhance their knowledge of literacy and literature. In the book *Bilinguality and Literacy: Principles and Practice*, Manjula Datta shows interesting examples of bilingual creative writing. She notes that bilinguals are anxious about their second language level; they are afraid of writing in L2, they feel more at ease speaking. Datta considers bilingualism as an asset and that teachers and education should “aspire toward multicompetence... enabling learners to mediate between their languages rather than become ersatz native speakers” (Cook, as cited in Datta 2000: 11). During her teaching career she worked with multilingual classes. She proposed to children at a stage of writing development to create stories and to write them down, of course with the help of a teacher. We can find drawings with some explanations; captions are usually written in one language. Another group of children was encouraged to write stories according to given topics. What can be noticed is that children choose the language conforming to their ideas and memories with which the stories are linked. Sometimes we can find the beginning in one language and the end in another one. As in the following example:

Oleg Ozaliny _____ 22nd October, 1998

Как I spent my weekend
in Kazakstan.

На некоторые выходные мы выезжали на дачу. Дача у нас
была в горах, потому мы купались в речке рядом с дачой.
И зимой мы находили там одну или две ночи но не больше
потому что дача там была.

Я расскажу вам как я проводил лето на даче с моей дво-
юродной сестрой и бабушкой.

Мы приезжали на дачу на один месяц мы купались, играли в карты
и т.д.

Oleg Ozaliny in Year 6 has lived in England for eighteen months. He speaks and writes confidently in Russian. In his article he tells of visits to the family villa in the mountains of Kazakstan. In summer they would stay for nearly a month, walking and swimming in the river. In winter they would stay for only a day or two.

Example taken from Datta 2000: 113.

What is interesting is that Oleg's code-switching starts in the word "Kazakstan". The correct spelling of this word in English is "Kazakhstan", whereas its Russian equivalent is "Казахстан". In the text the first part of this word is in English as everything that precedes it and the ending of the word is in Russian, employing Cyrillic alphabet, as is everything that follows it: "Kazakstan". It can be considered as an example of triggering in a written text.

Datta says that "[i]n a typical intercultural literate community children engage fluidly in all aspects of language, using all modes of learning—speaking, listening, reading, writing—feeling free to switch between both their languages to make meaning" (Data 2000: 114).

Another study of written code-switching in a multilingual community was presented at the American Educational Research Association Annual Meeting in 2007 by Kathy Escamilla and Susan Hopewell. In their paper *The Role of Code-Switching in the Written Expression of Early Elementary Simultaneous Bilinguals* they examined English-Spanish and Spanish-English code-switching in 327 writing samples of first, second and third grade children. They analysed code-switching

from the point of view of phonics, syntax, rhetorical structures, concepts, and also using traditional approaches to lexical code-switches, borrowing, and inter and intra-sentential code-switches. “[F]indings from this study indicate that the most interesting findings relate to cross-language use of phonics rather than to the inter and intra sentential code-switching which is frequently employed when looking at code-switching in oral language” (Escamilla and Hopewell 2007: 17). Cases of code-switching related to the transfer of phonics were much more numerous than those of inter- or intra-sentential code-switching. For example, Spanish phonics were used to write in English: "may" for "my", "chi" for "she", "laiket" for "like it", etc.; and English phonics to write in Spanish: "kavalelero" for "caballero", "agwa" for "agua", "welan" for "huelan", etc. The authors also paid attention to the learning environment and to the way children are taught to write in Spanish and in English. And “[f]inally, a major finding of this study was that children do not simply transfer skills, and strategies across languages, they transfer themselves and their experiences living and growing up in simultaneous worlds” (Ibid.: 19). Bilingual children may perceive linguistic and cultural environments of their worlds as simultaneous entities.

The use of two or more languages in writing has been studied not only among bilingual children but also among second language learners, who are of interest for both educationalists and cognitive psychologists. The outline of the research on L1 use in second language writing will be drawn in the next section.

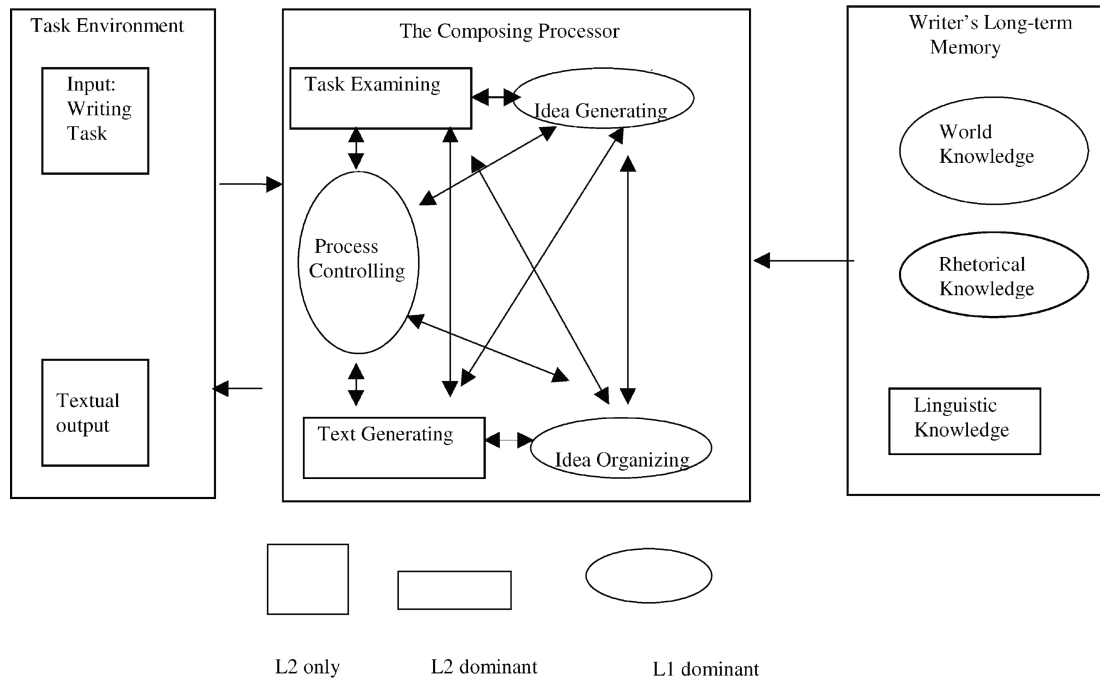
2.2.1.2. The use of L1 in second-language composing

Research on second language writing is a vast field of multidisciplinary studies (education, linguistics, psychology and their derivatives) which has come to the fore since the 1980s. Much focus has been on the comparison of L1 and L2 writing processes. Researchers have come to a consensus that one salient difference between L1 and L2 writing is that L2 writers may use more than one language to deal with problems that they face while composing in L2; they often switch back and forth between their languages and may use both L1 and L2 for cognitive operations.

Wang and Wen (2002) worked out a model of the L2 writing process based on the model of L1 composing proposed by Hayes and Flower in 1981. The models are similar in that they consist of three parts: the task environment, the composing processor, and the writer’s long-term memory. Using different figures (a square, a

rectangle and an ellipse) Wang and Wen indicated the language a writer may think in during L2 composing.

Figure 1: A descriptive model of L2 composing process by Wang and Wen (2002)



The task environment contains the input for the writing task (a writing prompt) and the output (a produced text). Usually for the L2 writing activity both input and output are supposed to be in L2.

The composing processor represents the writer's cognitive activities including five categories: task-examining, idea-generating, process-controlling, text-generating and idea-organizing. The main difference from the Hayes & Flower model lies in the recursive nature of the writing process. All the composing activities are connected with bidirectional arrows, erasing the former rigidity of a sequential order.

The third part of the model is the writer's long-term memory. This component is adopted from the Hayes & Flower model. It shows that during the entire process of writing, both the writer's language system and the content networks of the long-term memory are activated and provide world knowledge (information related to the topic), rhetorical knowledge (information about organization of ideas), and linguistic knowledge (information required for sentence construction) (Wang and Wen 2002).

Wang and Wen developed this model after their study of 16 Chinese students enrolled in the English department of Nanjing University. The results proved that L2 writing is a “bilingual event” (Wang and Wen 2002: 239): both L1 and L2 were used in L2 composing tasks. The participants’ L1 use represents 30% of the think-aloud data (the notion of think-aloud protocols will be explained below). Other researchers (Kobayashi and Rinnert 1992 and Manchon et al. 2000, as cited in Wang and Wen 2002: 239; Wang 2003) also measured the L1 use. However, the results varied showing a great deal of inconsistency from one research to another. There are at least two reasons that may explain the disparity in the percentage of the L1 use.

First, the process of measurement itself was performed differently. According to Wang and Wen (2002), Kobayashi and Rinnert in 1992 used the participants’ self-declared data about the amount of L1 in their L2 composing; while Manchon et al. in their study of 2000 calculated the amount of L1 by deduction of the number of the written words from the total of words in the think-aloud protocols. Regarding Wang and Wen (2002), they evaluated the amount of L1 by dividing the number of L1 words by the total number of the words in the protocol. This difference in methodologies is one of the reasons for the discrepancy in the results.

Second, there is one more inconsistency, namely differences in data coding. Researchers define and thus distinguish language alternation in the L2 writing process differently. For example, Woodall (2002) defines language-switching as “any use of the L1 while engaged in the L2 writing process. A switch started with an utterance in the L1 and ended with the next utterance in the L2” (Woodall 2002: 14). Another definition was proposed by Wang (2003), who analyzed what preceded and followed the L1 sequence; as a consequence, Wang distinguished language-switching from the notion of direct translation from L1 to L2 as studied in Uzawa (1996). Wang defined language-switching as:

[...] mental operations that went from L2 to L1 as cognitive processes of problem-solving and decision-making while writers were engaging in their L2 writing. On this basis, the participants’ utterances in their L1 that they directly translated into English were not counted as a switch. Thus, a switch was identified as an idea unit in the L1 prompted by an idea unit in the L2 (Wang 2003: 356).

Discordance between authors concerning the way to define language-switching results in different data coding and thus in different number of L1 sequences.

It is important to note that the term code-switching, as we have seen above, is not employed in L2 writing process research; instead, the term *language-switching* is used. Alister Cumming (1990) explains that from linguistic perspective language-

switching behaviour in L2 writing resembles code-switching in conversation, however, they have different social purposes. According to Odlin (as mentioned in Cumming 1990: 502), code-switching performs an augmentative function when, for example, an interlocutor indicates their membership or marks a special status for a phrase in a conversation. On the contrary, language-switching fulfils compensatory function dealing with problem solution in L2 writing.

Faerch and Kasper (1987: 128–129) labeled this process strategic transfer, one kind of communication strategy that appears when learners perceive a problem in their language production, particularly for lexical choices, and then use relevant knowledge resources in their mother tongue to analyze and solve the problem (Cumming 1990: 502).

Going further, Woodall (2002: 8) suggests that language-switching in L2 writing can be described as *non-instructed* use vs *instructed* use of L1 in translating. He explains that translating a text from L1 into L2 implies the use of L1 during L2 production. In this case L1 is an essential part of the task prompt and that is why its use is instructed. As opposed to L1 use in translating, language-switching in L2 writing represents spontaneous and non-instructed use of L1. It is a kind of private switching or, in other words, “talking to oneself” (Woodall 2002: 8).

Woodall’s expression “talking to oneself” can be viewed both metaphorically and literally. On the one hand, it is a mental operation used to deal with problems faced while L2 composing; on the other hand, an essential part of the L2 writing research data is composed of think-aloud protocols; that means that participants are asked to verbalize all their thoughts that come to their mind during composing “without evaluating one’s processing” (Ransdell and Barbier 2002: 3). Firstly, participants usually practice to say aloud their thoughts in a training session; only then do they proceed to the writing tasks. Afterwards, their think-aloud protocols are transcribed and analysed. Hence, the use of L1 in L2 writing represents verbalized the writer’s thoughts.

Of course, such variables as L2 proficiency, task difficulty and language groups (cognate/non-cognate) may influence the L2 writing processes. Still much research is needed to clarify their inter-relationships. Without wishing to go into detail of this vast issue, I would like to draw attention to some of the results produced by Wang and Wen (2002) and Wang (2003), which would be interesting to observe.

Wang and Wen (2002) confirmed by their study the involvement of L1 in composing activities, more precisely “L1 is more likely to occur in process-

controlling, idea-generating, and idea-organizing activities than in text-generating activities” (p. 239). These results were corroborated by Wang (2003).

Lurong Wang classified all the language-switching sequences into six categories: discourse, idea generation, language use, translation, lexical search, and metacomments. The description of the categories is given in the table below.

Categories for coding language-switching sequences (Wang 2003: 358)	
Category	Description
Discourse	Sequences that focus on the organization of written discourse, its structure beyond the level of the clause. Reference is made to organizational or rhetorical units, their interrelationships, or discourse types, such as paragraphs, introductions, conclusions, points, examples, essay, summary, or letter.
Idea generation	Sequences focus on the substantive content of the writing — the writer’s thoughts or ideas. Statements may indicate participants are formulating, considering, reconsidering or searching for content in their writing.
Language use	Sequences refer to grammatical, punctuation, or orthographic conventions.
Lexical searching	Sequences focus on finding a lexical item and phrasal-level lexical expression.
Translation	Sequences are directly translated into L1 from L2 in order to generate content, or monitor written production.
Metacomments	Sequences involve self-evaluation and metaconcerns about the appropriateness or qualities of one single element or text production.

The results showed that a significant number of L1 were employed for the three following composing activities: idea generation, lexical searching, and metacomments. These were the three common purposes for language-switching in both groups, with high and low proficiency, revealing qualitative differences in language-switching behaviours of high and low proficient participants (for more detail see Wang 2003).

These results will end the outline of research on second-language composing and will close the psychological perspective on language alternation in writing. The next section will view code-switching from a more customary perspective, that of sociolinguistics.

2.2.2. Sociolinguistic perspective.

2.2.2.1. Code-switching in literature.

Literature is a rich source of written code-switching. For example, language alternation can be found in such literary works as *Henry V* by Shakespeare, *Jane Eyre* by Charlotte Brontë, *War and Peace* by Tolstoy, *Popular Music* by Mikael Niemi, *La Vide es un Special* by Roberto Fernández, *Floricanto en Aztlán* by Alurista and many others. Code-switching can be employed for various purposes in literature.

The use of several languages or varieties within the same text, or code-switching, has a multitude of possible functions within multilingual literature. It includes different languages being used for different characters or voices; to mark out different parts of the text; to represent a mixed speech mode which characterizes the community; or to bring in different registers or sets of allusions. (Gardner-Chloros and Weston 2015: 186)

Some literary scholars have applied the models which were elaborated for code-switching in speech and which were presented in Chapter 2 to multilingual literary texts. For instance, Lenora Timm (1978) gave a full range of constraints on intra-sentential code-switching, focusing on Russian–French switchings in Tolstoy's *War and Peace*. She says that switchings between subject and predicate, within verb phrases or between verb and adverbial phrases are rather common in *War and Peace*.

Callahan (2004) applied Myers-Scotton's Matrix Language Frame model to a selected corpus of Hispanic American literature. She comes to the conclusion that the majority of the code-switches found in the studied corpus accounts for the MLF model, namely in the narratives. However, the difficulties were compounded when dealing with dialogues, where the application of MLF model was rather problematic.

Literary texts have also been examined from a social perspective. A number of researches have focused on the socio-pragmatic functions of literary code-switching. Studies on Hispanic literature account for most of these.

For example, Montes-Alcalá (2000; 2012) conducted research on a corpus of novels by writers of three Spanish-speaking communities in the United States: Cuban–American, Chicano and Nuyorican. She categorized the Spanish–English switches found in the corpus according to the socio-pragmatic functions of code-switching in speech, basing her research on the classifications proposed by Valdés-Fallis (1976), Poplack (1979/1980/2000), Zentella (1997), McClure (1981) and Gumperz (1982). The classification and the functions found by Montes-Alcalá (2000) will be applied in the research of the present PhD and will be further discussed in detail in Part III, section 4.3. In her work, Montes-Alcala comes to the conclusion that “code-switching is a valid option for the bilingual writer. Be it an aesthetic device in poetry, or a characterization technique in drama and fiction, code-switching serves a wide array of socio-pragmatic and stylistic purposes and it constitutes a substantial part of contemporary Latino literatures” (2000: 237).

Two other studies on socio-pragmatic functions of literary code-switching were undertaken by Carla Jonsson. The author analysed Chicano theatre play scripts by Moraga (2005; 2010) and in another work she studied multilingual texts in Sweden dealing with code-switching between Swedish and two national minority languages: Sami and Meänkiele (2012). In her studies, Jonsson shows that code-switching in bilingual literature performs both *local* and *global* functions. The *local function* can be seen on the textual level and corresponds to some functions in the typologies developed for oral code-switching; for example, quotation, interjection, reiteration, etc. *Global functions* are of a macro level and are “meaningful [...] in a more global sense” (Auer, as cited in Jonsson 2012: 223). Global functions relate to the social, political, historical and economic context of the literary works. Carla Jonsson emphasizes two of them: power relations and identity construction.

When speaking of code-switching in literature, one issue that cannot be omitted is “*authentic*” versus “*artificial*” switchings. Written code-switching is often compared to the oral case, inasmuch as it is expected to be identical to the spoken code-switching, “*authentic*” so to speak. This type can be found in dialogs. They claim to reproduce oral conversations that can happen in a multilingual community. However, code-switching can also appear in monologues representing the author’s voice. It is said to be “*artificial*” as it does not represent real bilingual speech. “Artificial” code-switching is often criticized for “decrease[ing] the artistic merit of the work” (Callahan 2004, 99). However, this view is countered by a group of

scholars who advocate that literary language is different from oral communication. They say that readers assume that they will not find language heard in everyday speech in a literary work.

Acknowledgement of the separate reality of all literary texts carries with it a crucial complication for understanding bilingual literature. The bilingual language of literary texts is not the same as the language of a given bilingual community. To presume that bilingual literature directly corresponds to usage in a given bilingual community entails a total misapprehension about the relationship between literary language and communal language (Keller, as cited in Callahan 2004: 100).

Even if the written text imitates the oral communication, there are still some differences between them. Laura Callahan (2004) describes three main distinctions. The first one is “due to the difference in channel”. Written language has to be more expressive because here the nonverbal communication is replaced by the lexical representation. Information that is usually transmitted by the physical context, facial expression, or tone of voice should be expressed by written words. However, at the same time writing allows more freedom from phonetic constraints. The writer does not know how his work will be read, aloud or to oneself, whether there will be a foreign accent or not, etc. Phonetics is an issue in spoken code-switching, but not in written.

Callahan also mentions visual effects of code-switching on readers. “The most basic of these is the visible contrast between the two codes” (Callahan 2004: 102). The effect can be increased if the languages used in the text have different ways of writing; for instance, alphabet and hieroglyphs, or Latin and Cyrillic alphabets. A different typeface can also accentuate the contrast between the languages. The standard type is usually used for the Matrix Language and italics for the Embedded Language. It can also add some thematic emphasis if the reader is used to italics which serve as a focus on the thematic content of the word/phrase in monolingual literature.

The second difference of written code-switching from oral, according to Laura Callahan, is caused by effects of written code-switching on the reader. Code-switching in speech is usually employed if both of the speakers are bilingual; meanwhile readers of bilingual literature can be either bilingual, or monolingual, or they can even be learners of L2. Writers often translate words and phrases that can cause some difficulties for the readers with an insufficient level of the Embedded language. Callahan refers to Rudin (1996), who described three different types of translation that can be used: literal, nonliteral, and contextual. Both literal and nonliteral translation can follow or precede the Embedded language. In the literal

translation, parentheses and different forms of punctuation can be used; the nonliteral one represents a paraphrase; as for contextual translation the sense of words/phrases in the Embedded language can be deduced from the context. Moreover, books often comprise a glossary with the translation of foreign elements.

It is said that bilingual writing is harder for the reader. It requires more effort, but at the same time it provides satisfaction; “Translation for non-fluent in some ways resemble puzzle-solving. [...] The native speaker may be pleased in the belief he has understood a poem, but the translator has a much greater sense of accomplishment” (Graham, as cited in Callahan 2004: 110). Texts where two different languages are switched demand not only to decode one more language which may be unfamiliar for the reader in a written form, but also words in the Embedded language may carry cultural or political connotation which may be unknown for the reader.

Is the effect of such texts the same on the monolingual and bilingual reader? Translation methods are used to facilitate the understanding of the text for a monolingual reader, to make it more accessible for someone who has not sufficient knowledge of the Embedded language. However, translation can have a negative effect on a bilingual reader—they may find it excessive.

The third distinction between oral and written code-switching is associated with register and thematic content. Laura Callahan conducted a survey on oral and written code-switching habits. The majority of her informants, who were approached in the academic settings, admitted that they sometimes code-switched while speaking, but half of them asserted that they would never switch Spanish and English in writing. Code-switching is often considered as unprofessional and “colloquial”. Informants often associate speaking with informal style while writing with a formal one. “Writing is virtually synonymous in the public mind with more formal register⁵ than those found in speech” (Ferguson, as cited in Callahan 2004: 113). However, oral and written texts do not correspond to informal and formal registers respectively. Focus on the content can be more important than the interpersonal involvement that is usually considered as one of the main markers of formality and informality of the text. Montes-Alcalá (2000) shows that code-switching is frequently used in hand-written letters and notes exchanged between Spanish–English bilingual speakers:

It has been revealed that those individuals who normally code-switch when speaking will do so when writing. Moreover, even those subjects

⁵ “Broadly conceived, a register is a language variety viewed with respect to its context of use” (Biber and Finegan 1994: 4).

who normally do not engage in oral code-switching will shift languages when writing to another bilingual speaker. Once more we speculate that code-switching is more acceptable in written forms among bilinguals, serving as an in-group code at the same time (Montes-Alcalá 2000: 71–72).

2.2.2.2. Code-switching in electronic writing.

Another rather new field where written code-switching has been examined is electronic writing. “[...] [C]ode-switching in electronic writing appears to be a natural phenomenon showing a trend in our modern and even more globalized society. Linguistic contacts do not involve only those languages used in the interactions, but also the writers, societies and cultures of which they are an exceptional medium” (Laroussi 2011: 13). Electronic writing has become not only a way of exchanging information, but a predominant social tool for communication. With the development of technologies, it is known that people, especially students write much more than ever but chiefly for non academic purposes. One of the features of the electronic communication is that it is multilingual. There are websites in different languages; it allows people to interact in any language they want. There are also cases of language contacts, where code-switching occurs.

Even though there are many publications on the code-switching in electronic writing, it still remains under-researched, as do all the studies on written code-switching. Issues on electronic code-switching cannot be regarded independently; “[a] generally accepted methodology that takes the specifics of CMC [Computer-mediated communication] into account has not yet been developed. Researchers draw on frameworks originally developed for the analysis of spoken discourse, despite criticisms of the adequacy of such frameworks” (Hinrichs, as cited in Androutsopoulos 2013: 668). Electronic code-switching has been mainly regarded from the point of view of pragmatics and sociolinguistics discussing the issues of identity, ideology, social change, strategy of communication, etc. (Crystal 2006; Laroussi 2011; Halim and Maros 2014).

Such web resources as websites, blogs, chats, emails and online diaries have attracted scholars’ attention. Dora Carpenter (2011) examines <http://harissa.com> which is a website for Jews of Tunisian descent. After the independence in 1956 and Arab–Israeli conflicts, many Jews, who represented a considerable religious community of the country at that time, left Tunisia for France or Israel. Carpenter analyses French and Judeo–Arabic code-switching. She says that code-switching is

very frequent on this website. However these switches are often translated or explained to facilitate its comprehension for other web users. The most recurrent language alternations are lexical switches, i.e. Judeo–Arabic words and phrases are inserted in the French Matrix language. Also intra and inter-sentential code-switching can be found. One of the main issues studied by the author is transmission of identity through language. According to Carpenter it is clearly seen that the users of this website were born outside of Tunisia. She justifies this statement by the frequent use of switchings and the limitations in language skills in various semantic fields. Dora Carpenter concludes that:

The virtual space of Harissa.com is definitely a *lieu de mémoire* where Tunisian-Jews revisit their past and where they can share a sense of community. [...]. As a virtual Tunisia where the multiple identities of the Tunisian Jews are on display, it also allows new generations to adapt their own response to their issues of identity building. Code-switching – the alternation between languages – and the variations on the religious, cultural, linguistic paradigms on Harissa.com become a symbolic representation of the flexible identities in constant re-invention among the Tunes of today (Carpenter 2011: 31).

Régine Delamotte and Cécile Desoutter (Delamotte and Desoutter 2011) study business emails and work or course based messages posted on a forum. Even if French is the main language in their corpus, code-switching is still an alternative for those who are in difficulty using French. They state that business emails are more monolingual in comparison with business telephone communication. As for students' emails, the frequency of code-switching depends on the status of the speakers: “the more equal the relationship with the intended addressee (student, tutor, administrator, teacher), the greater will be the likelihood of other language appearing” (Ibid.: 71). It is believed that in writing you have time to choose the language before you start, you have all the time to think it over. However, the authors note that the final decision is made during the conversation, depending on the choice of the interlocutor. If there is a dominant language, this does not mean that the correspondence will be monolingual. It starts in French, but if someone employs another language, the interlocutor is glad to have a choice to switch to continue the conversation.

Another interesting study on code-switching in electronic writing was made by Fabien Liénard and Marie-Claude Penloup (2011) who examined the role of code-switching in personal writing. The authors collected different types of data: they conducted interviews with bloggers and also analysed the private blogs:

skyblog, overblog and canablog. Their findings are the following: code-switching may represent

- a marker for closeness in a network of bilingual bloggers
- a strong identity claim (both young identity marker and the double linguistic and cultural identity)
- [a] painful relation to language (“a wrench, an insecurity”) (Lienard and Penloup 2011: 86).

The authors state that there is an opposition between spoken and written code-switching. While speakers are at ease to code-switch orally; they hesitate in writing. Code-switching occurs more in informal writing which is closer to spontaneous oral conversation. It seems that the writer shows closeness to his interlocutor and also takes into consideration his addressee’s bilinguality. Code-switching often appears in intimate writing such as online diaries. The authors admit that it would be an interesting case to study.

2.2.2.3. Theoretical framework for multilingual texts

Multilingual writing includes a range of written multilingual texts. They may be described as a continuum from more to less conversational genres. The majority of the research on code-switching in writing has tested models and theories elaborated for oral code-switching, applying them to written multilingual texts. This has proved relatively successful if the text under study is close to conversational genres, for example in the case of rap lyrics (Sarkar and Winer 2006), emails, blogs, social networks (Crystal 2006; Laroussi 2011; Halim and Maros 2014), etc. However, for some kinds of texts (e.g. posters, newspaper articles, advertisements, etc.) where the interlocutor is a quite distant reader, it could be difficult to apply such models of spoken code-switching as Gumperz’s conversational functions or Auer’s participant-/discourse-related code-switching, etc.

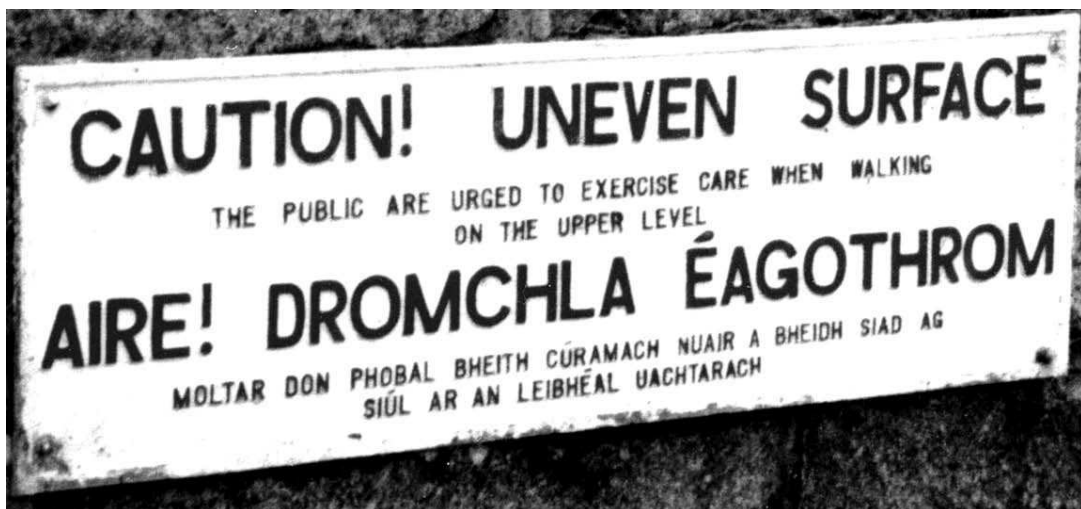
The first model for multilingual text analysis appeared quite recently. Mark Sebba (2012) developed a *multimodal frame of analysis* which could be applicable to a wide range of texts.

He explains that while spoken code-switching has mainly one dimension—namely “the juxtaposition of spoken linguistic units from two languages within a single interactional event”—language alternation in a written text is “potentially multidimensional” including juxtaposition on both linguistic and visual levels

(Sebba 2012: 106). The proposed multimodal framework takes into consideration both linguistic and visual perspectives and consists of four components:

1. *Units of analysis*, which in their turn include grammatical units (sentences, morphemes), genre-specific units (paragraphs, headings, etc.), and visual/spatial units (column, box, frame, etc.).
2. *Language-spatial relationships* which may describe texts as being symmetrically placed in a given space, asymmetrically or a mixture of both.
3. *Language-content relationships* which distinguish equivalent (texts have a similar content), disjoint (texts have different content) and overlapping texts (some of the content is repeated in the other language while some is not).
4. *Linguistic mixing types* containing monolingual units, mixed (elements of two or more languages) and language-neutral (elements that cannot be assigned to one particular language, but belong to both (all) languages used in the text) units.

With this framework, Sebba describes two categories of multilingual texts. The first one is parallel texts which have symmetrical language–spatial relationships, equivalent language–content relationships and an exclusively monolingual linguistic mixing type. “In other words, the ‘parallel’ type of multilingual text consists of matched units, symmetrically arranged and containing identical content in each language, without any language mixing” (Sebba 2012: 109). The author gives the following bilingual warning sign as an example of a parallel text:



Bilingual warning sign (Sebba 2012: 107).

The second category is complementary texts where language-spatial relationships are asymmetrical and language-content relationships are disjoint. Concerning linguistic mixing types, it varies; some texts may contain only monolingual units, while other texts may have a combination of several or all the mixing types (monolingual, mixed and neutral units). The following example of the advertisement for “Cien Alegrías Tire and Wheel service”, which Sebba provides for complementary texts, clearly shows that the languages are arranged asymmetrically, as well as their mixed ordering (English units constitute the bands 4, 5 and 6, and are followed by the Spanish units 7 and 8, and then more English appears in the bands 9 and 10). In terms of the content relationships, what is written in one language is not repeated in another one. Linguistic mixing types are monolingual except for the band 2 which is of a mixed type containing the name of the business in Spanish and clarifying its nature in English.

CIEN ALEGRÍAS Tire and Wheel						1
LOGO LOGO LOGO LOGO LOGO LOGO LOGO LOGO						2
BRAND 1 - BRAND 2 - BRAND 3 - BRAND 4 - BRAND 5 - BRAND 6 -						3
17 inch Starting from	18 inch Starting from	20 inch Starting from	17 inch Starting from	22 inch Starting from	24 inch Starting from	4
\$\$\$.99	\$\$\$.99	\$\$\$.99	\$\$\$.99	\$\$\$.99	\$\$\$.99	5
INCLUDES MOUNTING AND BALANCING, LUGS, LOCKS AND FREE						6
¿Necesitas Reparar?						7
Reparación de rines de aluminio Doblados - Quebrados				COMPRAMOS RINES DE ALUMINIO QUEBRADOS		8
OPEN 7 DAYS A WEEK! MON-SAT 8AM - 8PM - SUN 9AM - 5PM - 7 LOCATIONS						9
9999 Southern Drive - Hometown - 555-444-						10

Advertisement for “Cien Alegrías Tire and Wheel service” (Sebba 2012: 104).

Of course these two text categories do not exclude other intermediate types. In this case the framework may help to describe and analyse their distinctive features.

To conclude, Sebba's multimodal frame of analysis is a significant and long-awaited step in the research on language alternation. It is the first approach specifically designed for written language which thus forms a path to various genres of multilingual texts. The framework outlined here was conceived to facilitate the description and analysis of multilingual texts "provid[ing] the basis for a productive analysis of the wealth of multilingual texts that now exist" (Seba 2012: 113–114).

As this chapter has shown, there are various forms of multilingual texts ranging from children's/students' compositions to webpages, literary works and even signage. This list is not exhaustive: for example, it did not mention language alternation in such types of personal writing as diaries or notes where the author is talking to oneself. One of the few examples of research in this field is the PhD dissertation of Cecilia Montes-Alcala (2000). She analysed socio-pragmatic functions of code-switching in Spanish–English multilingual texts. One of the text types under analysis was a personal diary; due to the difficulty of such data collection Montes-Alcala examined only one diary that of her own.

In analyzing the entries of a bilingual individual's journal, we can conclude that even when code-switching is not the daily pattern of oral communication, language switching does occur in writing more often than not. When provided with two codes, bilingual individuals have a "superset" of resources to better explain their ideas, and these are all put to use in writing (Montes-Alcala 2000: 43).

Personal journals as well as notes are not intended for an external reader. The notes are often taken for personal purposes. Thus, such kind of writing is very different from all the types of texts discussed above. The next chapter will study the notion of notes itself showing its various forms and fields of application.

Chapter 3. Notes

“[T]here’s no reason why anyone bright enough to be accepted as a medical student should not learn to write his own language tolerably well” (Osmond 1955: 417).

The third notion to explore is that of “notes”. In particular, we need to address two questions, both of which revolve around the cognitive aspects involved, and in many ways, the two questions are in fact two different ways of looking at the same issue. The first takes its inspiration from the observation that we “take” notes, rather than “write” notes. The same is true, it would seem, of other languages: *prendre des notes*, *tomar notas*, *notizen machen* (here, in German, you “do” notes – the same as in a Russian example), *делать записи*. This would seem to suggest that there is a fundamental difference between what is happening during the process of creative writing and that of note-taking. Second, we need to see what differences exist between a “note” and a “text”, in other words, what formal criteria we can use to define the “note”.

3.1. Note-taking

The issue of note-taking has been of interest to scholars for some considerable time, now. In 1955 Humphry Osmond devoted much of his article “Notes on note-taking”, from which the above epigraph is taken, to the tools physicians used while noting their patients’ case-histories such as a writing board, paper and ink. Even if contemporary scientists no longer search for the longevity of paper and ink, and even though medical practitioners are still well-known for their illegible handwriting, such issues raised by Osmond as methods and shorthand are still topical issues both in medical and educational fields. Up until the mid-20th century, the main focus was writing materials themselves (paper, ink, etc.), whereas later works showed more interest in the cognitive aspects of note-taking. For example, Guy Gimenez and Sophie Barthelemy elaborated a note-taking method for psychiatrists dealing with group therapy (Gimenez and Barthélémy 2011); while Faraco (1997), Barbier et al. (2003), Piolat (2006), Teng (2011), and others examine students’ note-taking manners and habits.

Note-taking occupies a lot of time of students’ life and studies. That is why it is of such great interest for educationalists. Each country has its own particular

features of the educational system; taking into consideration these differences and even comparing them in various environments, researchers with different scholarly backgrounds and from different educational systems show their interest in note-taking. The process of note-taking can be defined among other factors as a cognitive activity, which implies recording information in a written form from an oral or written source, the information that note-takers reckon important (Piolat 2006: 7). The word “cognitive” plays an important role in this definition. The cognitive effort employed in note-taking should not be underestimated. This activity does not consist only in writing down what is said or written. “Note-taking is a complex activity that requires comprehension and selection of information and written production processes” (Piolat et al. 2005: 291). Note-takers have to select information while listening, as they do not have time to lose before and while writing the information down. The average speed of student’s writing while taking notes is from 5 to 10 times slower than of lecturer’s speaking. Françoise Boch and Annie Piolat (2005: 102) provide the following figures: “the average writing speed of a student is around 0.3 to 0.4 words/second, whereas a lecturer speaks at a rate of around 2 to 3 words/second”.

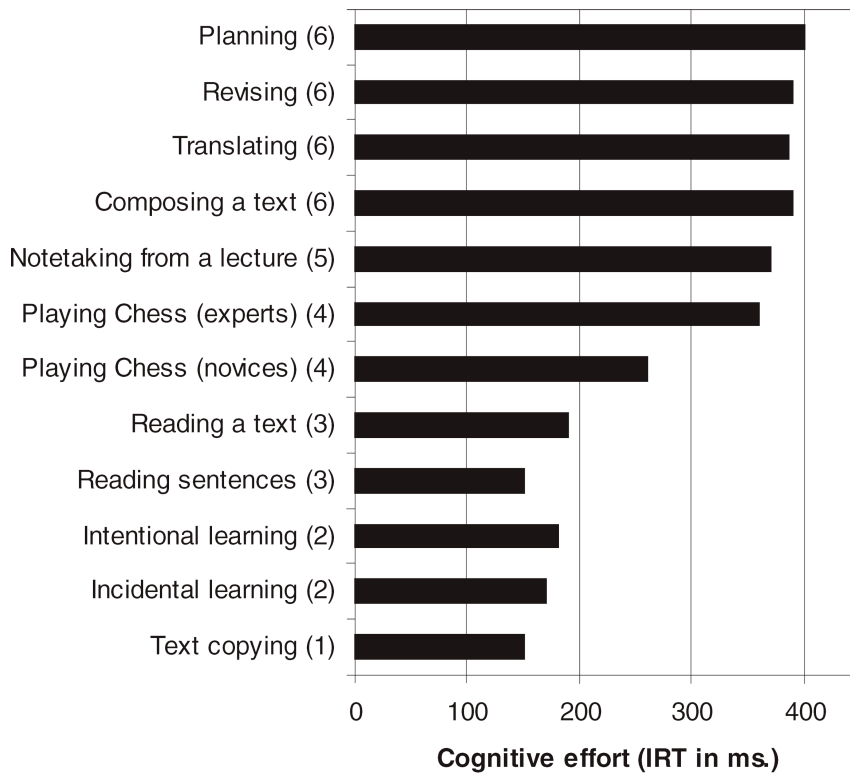
Compared to the estimation of speaking and writing speeds, the measurement of cognitive effort may seem more sophisticated. “Cognitive effort refers to the fraction of limited attentional resources that are momentarily allocated to a process” (Piolat et al. 2005: 298). Depending on the research objective/s different techniques for measuring cognitive effort could be applied. The two of them, which I would like to mention, are a *dual-task paradigm* and a *triple-task paradigm*.

A *dual-task paradigm* is often used to assess and/or compare the cognitive effort expended on higher-order cognitive activities such as reading, text composition, etc. “Dual” makes reference to the two simultaneous tasks performed by participants; they should carry out a primary activity such as reading, writing or note-taking, etc., and a secondary stimulus which consists in reacting to sound tones. For example, while taking notes from a written document, participants have to react as quickly as possible to the tones that they hear either by pressing a computer/mouse key or by saying the word “stop” into a microphone; the tones are randomly placed with time intervals of 15–45 seconds. Reaction time (RT) then is compared to control time when participants react only to the stimulus without conducting any primary activity. Inter-response/reaction time (IRT) shows the measure of the cognitive effort devoted to the primary activity. Comparing two different activities, for example note-taking and text composition, the equal IRT will mean the equal cognitive effort employed in these activities; however, if one of the activities is more cognitively demanding the IRT will show it by a bigger score.

While a dual-task paradigm provides a thorough comparison of a holistic cognitive effort expended on different activities, it gives no information on the cognitive effort of specific cognitive processes of each activity. This role could be fulfilled by a *triple-task paradigm*. In addition to a primary activity and a secondary stimulus, participants have a third verbalization task to perform (immediately after the stimulus) where they have to name the process that was intruded by a sound tone. Of course, participants are firstly trained to recognize the activated cognitive processes by remembering their conscious thoughts at the interrupted moment. In the majority of texts composing studies the three following processes are analyzed and consequently are a focus of identification by participants: planning of ideas, transformation of ideas into sentences and reviewing of texts/ideas (Piolat et al. 2005). The selection of the aimed cognitive processes is guided by the research question leaving it open to choose any processes of interest for the research. A triple-task paradigm thus provides information on processing time of cognitive processes as well as a cognitive effort assessment of each chosen process.

Depending on the aims of the study, researchers can use a dual-task or a triple task paradigm or even both if they are intended to make a double comparison of the cognitive effort used in different activities as well as the cognitive effort of particular cognitive processes in all the compared activities.

Using a dual-task and triple-task paradigm Annie Piolat shows in her works a range of quite interesting results of measuring cognitive effort (Barbier and Piolat 2005; Piolat et al. 2005; Piolat 2006). Here is one of the figures she provides in Piolat et al. 2005 (p. 299), showing the cognitive effort of undergraduates while performing various tasks:



Interestingly this bar chart shows that while “Text copying” matches “Reading sentences” and not even “Reading a text”, the cognitive effort expended on “Note-taking” from an oral lecture is situated in between that spent by an expert “Playing Chess” and the cognitive effort needed for “Composing a text”. All three latter activities comprise such processes as prior knowledge retrieval, planning and tackling problematic situations, which demand more cognitive effort than those activities (such as copying or reading) that do not need such operations.

Piolat also explains that note-taking from a lecture requires more cognitive effort than note-taking from a written document due to a bigger number of concurrent operations while listening; that is why there is a bigger load on working memory (Piolat et al. 2005: 303). A distinction should be made between long-term memory, which keeps prior knowledge, and working memory, which emerges from the concept of short-term memory and “which involves the temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities” (Baddeley 2003: 189). Both memory systems are linked. Wetzels et al. (2011) worked on the retrieval-directed function of note-taking in prior knowledge activation.

Prior knowledge activation involves the transfer of available knowledge from long-term memory to working memory. If new, potentially meaningful information is related to the assimilative context of existing knowledge held in working memory, this information can be integrated in the existing knowledge base. [...] So, the availability of prior knowledge is not sufficient to achieve

higher learning outcomes. This knowledge should be retrieved and activated in order to establish relationships between existing knowledge and new information (Wetzels et al. 2011: 275).

They suggest that note-taking may have a favourable effect on learning. The findings show that note-taking is beneficial for students with high prior knowledge as it facilitates activation of their prior knowledge by decreasing the load on working memory.

While talking about the load it is necessary to mention that working memory has a limited capacity. An easy way to illustrate a memory span is to ask a person to remember and immediately repeat a sequence of digits, like a telephone number. An average person can hold in memory about 7 digits (Baddeley 1998: 169). There is another interesting study that discusses working memory limits (Guichon 2006). Nicolas Guichon conducted a study where second-year students, enrolled in the Bachelor of Economics courses in France, were asked to make a summary in L2 (English) of a video taken from BBC. The majority of the subjects had an overload of working memory after 58 seconds and again after the second minute of the video, which lasted a total of 2.40 minutes. In both cases, the overload was observed empirically by an obviously apparent, brief and temporary difficulty which the subjects manifested in the note-taking process. He states that working memory operates by cycles, which during an exercise of L2 listening has a span of about 1 minute. His study deals with note-taking in L2, which demands more cognitive resources than taking notes in L1. Besides a bigger cognitive load during L2 note-taking, Marie-Laure Barbier and Annie Piolat (2005) show in their study that note-takers are less fluent and use fewer abbreviations in L2. However, that does not mean that the difference in L1 and L2 note-taking is due to the linguistic level of note-takers. After 3 years of studying a foreign language in the university students write in L2 with almost equal cognitive effort and fluency as in L1. Performance in note-taking differs due to the lack or insufficient amount of automatic procedures in L2 note-taking (Barbier and Piolat 2005). The idea that institutions should help students to be prepared for the process of note-taking including work on automatic procedures and elaboration of their own system of note-taking has persisted since Humphry Osmond's article (1955) where he wrote: "[...] schools should insist that their students develop a legible handwriting which will deteriorate as little as possible when used at high speed" (Osmond 1955: 417). Nowadays another aspect of this issue has appeared where second-language learners are the centre of attention: "teaching a second language at university should prepare the students to develop linguistic skills, but not only in writing tasks: also in specific activities such as note-taking" (Barbier and Piolat 2005: 5).

However, while note-taking may be thought of as a general cognitive activity (taking in account physical aspects such as text production, writing speed, etc.), evidence shows that personal and cultural factors also have a role to play. For example, students may have developed personal skills or have elaborated a personal system of note-taking during their education and other experiences. Furthermore, the teaching tradition of each country influences the way notes are taken. Barbier et al. (2006) give the examples of French and American students who have different strategies for note-taking. The French write down as much information as possible to use their notes as an external memory afterwards; as for the Americans, they do not write a lot, noting only the key elements. This difference could be due to the different educational systems. In France the information given by a lecturer could be sufficient to pass an exam, while in America a lecturer adds comments to what is written in a recommended to students book/books and students' notes depend on the information in the book and not on what the lecturer has said (Omer 2003). Having my own experience in the Russian academic system, I could say that it has some features of both American and French educational approaches. My sense is that Russian students try to write down everything that the lecturer says using their notes as an external memory; however, they also expand their notes or check if anything was not clear from the lecturer with the information from the recommended books. In 1997 Martine Faraco conducted a research, which confirms the necessity to take into consideration personal and cultural factors while analysing and working with foreign students' note-taking. She studied notes taken by foreign students (3 English, 5 German and 4 Spanish natives) who were attending courses in a French university. The degree of their proficiency in French can be estimated to be comparable including from 3 to 7 years of French classes in a high school and at least 1 year of French language courses at their universities. For all the participants it was the first academic experience in a French university. The results show that after 3 months of classes foreign note-takers can still be distinguished according to their linguistic groups. However, all of them, because they lacked a specifically taught system of French abbreviations, instead developed their own systems of note-taking.

A particular feature that distinguishes native and non-native note-taking is code-switching. While in the notes of the Spanish students there were almost no traces of the L1, the English students used it massively to provide an efficient substitution for their French deficiency and it served as a meta-discursive tool for the German students. To visualize it, here are the examples (translation into English

is given in square brackets) that Martine Faraco provides in her research (Faraco 1997):

1. The use of German as L1 for structuring. The student takes the distance from the lecture tagging its different parts with L1:

BESPRECHUNG [= discussion]
ou encore faire référence à des documents annexes : [= or refer to the annexed documents]
figure 23 (Zettel) [= figure 23 (piece of paper)]

2. The use of English for inter-lingual translation rather than for intra-lingual one. When the lecturer's explanations are quite fast or too long the student translates them directly into L1:

fin XIXe : laïque et matérialiste [= end of the 19th: secular and materialistic]
---> Marx Le Capital
opposite to Smith
---> Durkheim Le Suicide

Durkheim says that there exist economic situations which can affect your life and how you feel

Lien social : people living together [= social ties]

When the students were asked about the difference of their note-taking in L1 and L2 one of the English interviewees, who speaks both French and German as foreign languages, said:

In French [...] [I] often use all the 3 languages to be able to write better and faster. I have developed a system and I think that I'm managing it well – we'll see after the exams⁶. (Faraco 1997: 4).

Development of a personal system of note-taking characterized by particular abbreviation procedures, space management or the use of several languages, etc., is a necessary tool to succeed in one's studies because, as it was mentioned before, note-taking is an activity which follows students throughout their academic career. However, it is not their only "studies' companion". According to Annie Piolat and Marie-Laure Barbier (2007) students continuously drift between documentary note-taking (from a lecture or a text) and editing (using rough copies).

⁶ Translation by the author of the present work.

3.2. Genetic criticism

After having looked at documentary note-taking at the beginning of this chapter, I would like to focus attention on another type of personal notes and the second students' "companion" which is represented in rough copies. Different terminology is used for both tangible support and preliminary writing as a whole such as: rough copies, drafts, manuscripts, etc. An interesting proposition was made in 1972 by Jean Bellemin-Noël, who coined the term "*avant-texte*". The verbatim translation of this term would be "pre-text", however in the Anglophone literature, authors have tended to maintain the original term. Jean Bellemin-Noël introduced *avant-texte* as:

l'ensemble constitué par les brouillons, les manuscrits, les épreuves, les « variantes », vu sous l'angle de ce qui précède matériellement un ouvrage quand celui-ci est traité comme un texte, et qui peut faire système avec lui⁷ (Bellemin-Noël, as cited in Grésillon 1994: 108).

The approach studying such sets of papers is called *genetic criticism*. It has nothing in common with biology and genetics, on the contrary, it deals with genesis of literary works. Specialists or so-called geneticists are often of a literary or linguistic background. Such fields as theatre play scripts or musical sketchbooks are also covered by genetic criticism. Its history is rooted in the French structuralism movement, which gave birth to a new branch in the 1970s. The object of genetic criticism is not exactly the same as that of structuralism: the latter analyses a text as a "finished form", while the former considers a text as being in "movement" (Grésillon 2008: 23). Moreover, a published text provides linearity of reading and owes its social aspect to a potential reader for whom it was written and published; while an *avant-texte* is normally written for oneself with interlinear and marginal notes and corrections, intermittent regresses and different graphic symbols which break the linearity of reading. Documents that constitute a genetic dossier represent both written evidence such as plans, characters' synopses, worksheets, accessory notes, etc., and at the same time something in evolution, something that is preparing to come out as well as something "improper" (considering the number of crossing-outs and revising) and unfinished. Through all these tangible materials, which form

⁷ "a set which consists of drafts, manuscripts, trails, and « variants » which physically precede a work that is concerned as a text, with which they could form a system"— translation by the author of the present work.

a “prehistory” of a text, geneticists try to discover a creative mechanism of text genesis (Grésillon 1994). “[T]here is nothing mystical in the activities of genetic criticism, which pursues an immaterial object (a process) through the concrete analysis of the material traces left by that process” (Deppman et al. 2004: 11).

Jean-Louis Lebrave (1983) explained the duality in the positions that readers can adopt with regard to *avant-textes*. For the majority of people, to deal with a rough copy means to read it. However, when we start performing this habitual activity we understand that it is quite different from reading a text; it is during this pseudo-reading that people generate genetic information quasi-spontaneously that leads them to hypotheses of text production. At the same time, we are all producers of rough copies; this position as a writer makes us feel an enigmatic relatedness with genetic processes. Thus, Lebrave proposed a model of text production analysis based on these two aspects of being both a reader and a writer. Fundamentally, reading is an attempt to organize an *avant-texte* in a double perspective. Firstly, we read a text containing author’s corrections and substitutions which we intend to set in a chronological order; this process may be complicated by a great number of corrections, as for example *avant-textes* of such authors as Heine, Proust or Flaubert are composed almost exclusively of crossing-outs. Secondly, we try to set an evaluative order in an *avant-texte*; going beyond the cultural conventionality of text perfection, an evaluative and at the same time interpretative system could be an attempt to explain to ourselves why there are so many corrections in *avant-textes*. Having proceeded from a reader’s point of view to a writer’s perspective, Lebrave highlighted the difference between oral and written production; namely the fact that it is always possible to come back and correct, delete, enrich, etc. already written text, while in an oral discourse one may only insert a meta-linguistic comment (“no”, “I apologize, I misspoke”, “I will correct myself”, etc.). The fact which determines this difference and writing in general, is that “l’écrit produit quelque chose qui se trouve, non pas dans le temps, mais dans l’espace, et qui est un objet matériel manipulable à deux dimensions”⁸ (Lebrave 1983: 18). The two dimensions, that Lebrave referred to, are a textual and a graphic level of analysis. Lebrave explains that the textual level provides some kind of material for reading in its general meaning; this level corresponds to Grésillon’s description of *avant-textes* analyses which may include language (paraphrase, time representation, syntax, discourse, etc.), narration, themes, psychoanalysis (unconscious of the text) and author’s input vs historical epoch input (Grésillon 1994). Concerning the graphic level, Lebrave states that it brings *avant-textes* closer to visual arts with the

⁸ “writing produces something which is situated not in time, but in space; it is a tangible object which can be manipulated in two dimensions” – translation by the author of the present work.

orientation on its author's writing activity and not on the readers' judgment of a rough copy. Writing is considered to be a visual activity because it is guided with one's eyes, and also a spatial one because a writer organizes a text, in other words graphic outline or signs, in a space (Olive et al. 2010: 35).

The main focus of genetic criticism is on literary works, less on music and theatre scripts; works that allow building a genetic dossier with authentic papers written by the author and also external evidence such as correspondence between author and close friends, testimony of family members, etc. With the arrival of new information technologies, a static representation of a text, which is immobilized on printed pages, has been gradually changing into a dynamic one. Already in 1991 Lebrave depicted a software Hypercard, which was new at that time, and which allowed a dynamic visualization of genetic dossiers. Lebrave showed an example of the beginning of Flaubert's *Hérodias*. The uploaded genetic dossier could be characterized by its comprehensiveness and variety, it included facsimiles of the text; some visual documents and maps, that Flaubert used working on *Hérodias*, such as photographs of the site of Machaerous, maps of Palestine, sketches that Flaubert made on the margins of his manuscript; publications that Flaubert read, for example Tristram's works about Palestine; avant-textes organized in such groups as plans, scenarios, rewritings, etc.; editor's notes, comments and information. Hypercard provided a possibility to browse through all these documents, to search for a particular string of words, to show corrections made in a studied document and even perform a dynamic simulation of entire rewritings of a phase under concern.

Being a software of the end of the 20th century, Hypercard had many limitations and namely could not perform any functions except containing all the documents of a genetic dossier and exhibit corrections that had already been detected and marked by geneticists; moreover, Hypercard required at least minimum knowledge and competency in computer science and programming, which according to a survey was off-putting for potential users (Lebrave 1991). However, we should not underestimate this, as it was one of the first word processing programs which made the first step towards hypertexts. This first step was followed by many others with the advantages brought by a further progressive development of information technologies. The 2004 saw the description of a word processing software MEDITE which was elaborated "to facilitate the textual genetic studies" (Ganascia, Fenoglio, and Lebrave 2004: 91). This program allows comparison of two texts or more precisely two versions of a text by detecting textual transformations. The advantages of MEDITE in comparison to other word processing software such as, for example, Word are quite numerous. Firstly, MEDITE enables automatic detection of 4 transformations – namely insertion, removal, replacement and shifting; while Word

distinguishes only insertion and removal. Secondly, MEDITE has a higher level of reliability and accuracy. Thirdly, the program's interface is more convenient, marking different transformations with respective colours and allowing its users to take notes. And finally, MEDITE makes quantitative analysis possible, providing information on transformations for statistics. This software opens a new perspective on linguistics of writing.

Still, we should not leave unmentioned digital philology which aims at the creation of digital publishing collections and developing its scientific use. For example, "Philectre", which is a blend of "Philologie électronique", is a multidisciplinary project on digital philology. It was founded by the National Center for Scientific Research in France and is now composed of a medieval corpus of troubadours and a modern corpus which contains Flaubert's *The Legend of Saint Julian the Hospitalier*.

The technical progress, which has led to the accessibility of personal computers and consequently different word processing software, has at the same time dramatically changed the very notion of a text. Henceforth each user of a personal computer is their own editor who may write, correct and print their work, a first in the long history of writing and printing. Moreover, in the past all texts had traces, notes on the margins, rewritings, versions, etc.; these traces together allowed a text's history to be reproduced. Nowadays, writers may alter their document in a such way that nothing or nobody is able recognize and detect these modifications. It is almost impossible to reconstruct the genesis of a digital document. A text has become

[...] indéfiniment malléable et néanmoins toujours parfait dans sa réalisation matérielle, texte qui réunit donc en lui des couples d'attributs contradictoires, à la fois parfait et imparfait, achevé et inachevable, définitivement provisoire, ou provisoirement définitif.⁹ (Lebrave 2006b).

3.3. A study with a dual corpus

An original genetic dossier was collected and examined by Annie Piolat and Marie-Laure Barbier (2007). They managed to group together two types of avant-texte: both documentary note-taking and rough copies, which precede the writing

⁹ “[...] infinitely malleable and nevertheless always perfect in its tangible realisation, a text which unites contradictory pairs – at once perfect and imperfect, finished and unfinished, permanently temporary and temporary permanent” – translation by the author of the present work.

process. Psychology undergraduates, firstly, had to take notes of a 13-minute lecture on creativity, writing and madness; note-taking is considered as an *avant-texte* because students prepared notes with the purpose of their subsequent use in an essay writing exercise. These notes are said to be “provoked” *avant-textes* because this activity was obligatory for the participants of the experiment. Secondly the undergraduates produced two types of rough copies: one “provoked” and another “spontaneous”. A. Piolat and M.-L. Barbier adopted Almuth Grésillon’s definition of “*brouillon*” [rough copy] which refers to the writing/editing phase succeeding preliminary works such as notes, plans, scenarios, rough outlines and sketches (Grésillon 1994: 74). For 6 minutes, without looking in their notes, the students had to outline their ideas on the following topic: Do you think that literary and scientific creativity are of the same kind? Afterwards the participants wrote an essay using both their notes and provoked rough copies. Two months later, students sat an exam (the examination essay’s theme was “Children’s acquisition of calligraphy and reading”). 80% of students gave their examination’s rough copies for the aim of the research; these examination’s drafts are considered as spontaneous rough copies.

The research showed interesting results. The Genetic dossier of the psychology undergraduates represents elliptical writing. While discursive condensation¹⁰ lessens from note-taking to spontaneous rough copies, lexical condensation remains very dense in all *avant-textes*. Students are guided by the function of their writing to give it a particular form for the easy access to their notes, ideas, plans, etc. in future. For example, they can write linearly, tabularly, in a form of a list, or mixing these different forms. Rough copies often serve as a tool for conceptualisation that does not require text formatting; in other words, students use their drafts to “think” what to say rather than to look for a way to say it (Piolat and Barbier 2007).

With this study Annie Piolat and Marie-Laure Barbier pointed out a new field of personal writing that is rich in diversity of materials and which is still understudied. Their third type of *avant-textes* – namely, spontaneous rough copies – will compose the corpus of the present PhD and will be described further on. The direction opened by these authors could be highly prolific for other sciences such as education, cognitive psychology, linguistics, psycholinguistics, etc. *Avant-textes*, which do not have a potential reader except for their own authors, show language

¹⁰ Discursive condensation is a way of structuring and summarising information using pictograms (⇒, >, <, etc.), through list structures using dashes or stars (-, *, →, etc.), or even through syntax. (Piolat and Barbier 2007: 95)

and language production from a point of view which surprisingly has not yet been taken into consideration by researchers, or at least not enough.

There exist at least two variables that are applied to the analysis of both note-taking and rough copies: quantitative and qualitative (Barbier et al. 2006). The first one involves the total number of words, abbreviations, crossing-outs, foreign words, etc. Quantitative analyses show that foreign students use much fewer abbreviations and symbols as well as fewer condensation techniques while taking notes in L2. Numbers are also of great interest for geneticists who count the total of abbreviations, regressions or even crossing-outs to compare different stages of the author's writing process.

The qualitative variable concerns the content of avant-textes (topics, new words, words in L1 and L2, etc.), organization of ideas and grapho-spatial structuring (graphic symbols (arrows, multiple exclamation marks, etc.), graphic techniques (underlining, the use of different colours, font sizes, etc.), and space structuring (columns, text separation by a line or different text organisation on a sheet of paper, etc.)). In the analyses of both note-taking and rough copies researchers pay much attention to such an element as abbreviation (whether the author prefers truncation of the ending, or concentrates on the consonants of words, whether they use already established abbreviations or invent their own techniques, etc.); the use of symbols and other graphic elements, as well as the way the author structures the writing and manages the space are indispensable in avant-texte analyses being particular features that distinguish notes and published texts.

3.4. Interpreters' note-taking

Techniques of note-taking are also of great importance in interpreting. This discipline being self-sufficient has its own long scholarly history. A debate lasted over half a century over the necessity of teaching note-taking with a conclusion that:

[...] teaching note-taking was not only possible, but essential, and that – provided the students' notes were concise and structured – this should provide them with a solid basis for understanding their individual note-taking needs and tailoring their technique to suiting them (Kalina, as cited in Dingfelder Stone 2015: 149).

Note-taking is a substantial part of interpreting; clear and essential notes relieve interpreter's working memory, let them keep eye contact, be faithful to a source text and be efficient. Much literature is devoted to interpreters' training in note-taking, explaining different models and theories, highlighting such aspects as

communication, quality, fidelity, comprehension, knowledge acquisition, etc. (Gile 2009); one may even find interpreters' handbooks of note-taking such as, for example, Gillies's *Note-taking for Consecutive Interpreting* published in 2005. Throughout its long history, the Interpreting Studies community has finally agreed over the fact that systematic and structured notes are beneficial; however, subtlety of some points is still under discussion. Dingfelder Stone (2015) outlined the eight basic principles of note-taking in her work as following:

1. Dig deeper: noting the sense behind the words
2. Listen first, note later: noting only the minimum
3. Mix and match: noting a mixture of symbols, abbreviations, and full words
4. Keep it sweet and simple: noting unambiguously and efficiently
5. Structure determines content: noting vertically, diagonally or circular
6. Pick a language, any language: noting in whatever language is most efficient
7. Unburden your memory: noting names, figures and lists
8. Chronology and décalage: noting the right item at the right time

These eight instructions reflect textual, visuo-spatial and temporal aspects. However, very few studies have explored interpreters' notes empirically. Publications on interpreters' notes mainly concern note-taking systems, particular aspects of note-taking (such as comprehension, symbols, etc.), and interpreters' training (Szabó 2006: 130). A large number of works represent guides for teaching interpreting and contain suggestions, advice, exercises and even whole courses that a teacher may employ in class. The few empirical data-based studies performed on this subject usually concern either interpreters' expertise or language choice.

Csilla Szabó (2006) conducted an interesting study on interpreters' language choice in their note-taking. She examined the notes of second year student trainees enrolled in European Masters in Conference Interpreting with a Hungarian–English language combination. The results showed that 50% of the participants were guided by the Source/Target language strategy, while 50% had a preference for English (their second or B language) regardless of whether English was a source or target language. Having compared her study with previous works on the subject, Szabó agreed with Dam's statement that “interpreters are likely to take notes in whichever

language is easier and therefore faster” (Dam, as cited in Szabó 2006: 140). However, the characteristics of being “easier” and “faster” should not be restricted to one particular language, it should also take into consideration the nature of the language pair. For example, in Szabó’s study, Hungarian is an agglutinative language which produces longer words in comparison to the English language; “[...] the average word length is longer in Hungarian by approximately 35%” (Szabó 2006: 141; Szablya Consultants, Inc. [online]). The economical quality of English probably played a role in the subjects’ preference for the English language in their notes. In conclusion, Szabó suggested that there is a potential interest in the analogous comparative studies of Indo-European vs non-Indo-European language pairs as well as Indo-European in-group language comparison.

Assuming the significance of interpreters’ note-taking in a larger field of notes and note-taking, I will bring this very brief description of this issue to a close here, without going into further detail. This is because interpreters’ notes are not entirely consistent with the corpus and objectives of the present work. Firstly, interpreting deals with at least two languages at a time, meaning that while working, interpreters have two active languages, which facilitate a language alternation; moreover, the use of several languages in note-taking is widely applied in interpreting and sometimes it is even a necessary technique. The corpus of the present PhD was collected with a condition to have as little language influence on the subjects as possible; the instructed interpreters’ use of L1 is opposed to the non-instructed, spontaneous use of L1 by the participants of the present research. And the second inconsistency arises from the fact that interpreters’ notes represent a taught and learnt system which is lately adapted to one’s own personal use; that is the opposite situation to the personal systems of note-taking developed by students under study here.

3.5. Conclusion

This chapter went through the notion of notes, its forms and contexts. Various disciplines show interest in its different aspects such as language use by students, cognitive efforts while note-taking, genesis of writing through *avant-textes*, note-taking as a system for interpreters, etc. Each discipline has its own skew to notes, examining and hypothesising only about a particular side of this phenomenon, which is always studied in confines of a research objectives and not as such.

At the risk of repeating myself, I would like to emphasize that notes represent a rich source of data for further research which can be approached both on the textual and the visuo-spatial levels. Besides, notes reveal a temporal perspective in certain conditions. Both quantitative and qualitative variables are common in analyses of avant-textes of all types. Could it be possible to direct these variables on linguistic studies of avant-textes? Could avant-textes tell us more than they have already said? The rest of this study is an attempt to answer these, and other, questions about avant-textes.

Part II

Data description and results

The first part of this work outlined a number of studies, each contributing to a field or several fields of social sciences. The strength of the latter is in its multiple diachronic components. They appear in different moments in history and bear their own particular values. Even the least significant component is indispensable for the science, serving as a footing for its construction and development.

Moving on from the state of the art, which was depicted above, Part II will focus on the study, which was conducted within the scope of the current PhD thesis and, which, I hope, will contribute to various areas of social research in its turn. It will examine bilingual patterns, namely code-switching, in personal writing. As Part I shows, the phenomenon of code-switching has been well studied since the beginning of the 20th century. However, the intersection of bilingual writing and genetic criticism has never been the subject of the interest it deserves. Is it simply an area of research, which is waiting for its time to come? I believe so. The multiple reasons for choosing to study this particular topic, which will be presented in Chapter 4, will advance arguments in favour of its interest and importance.

The first chapter of Part II will expose how the study was undertaken, from the very early stages of data collection and methodological procedures, to its results. Chapters 5, 6 and 7 will discuss the three sets of data and propose a model of bilingual patterns in personal writing.

Chapter 4.

Methodological matters

4.1. The choice of the subject

As often happens the origin of the choice is rooted in one's personal interest and experiences. The interaction between the needs and wants of individuals and the way that these feed into the way a community functions have always fascinated me. Abiding by inductive reasoning, I believe that comprehension of the individual allows us to understand the functioning of the society as a whole. Language is one of society's characteristics. Their mutual influence—the influence of society on language of an individual and vice versa—results in language systems, which form the object of linguistic studies. Thus, research both on the societal and individual levels are of considerable importance, as what we learn from the one is valuable to the other, and vice versa.

For the past century research into the phenomenon of bi-/multilingualism has gained in importance. Countries where two or more languages have long co-existed started to invest in research projects aiming at the enhancement of a devaluated language and at an equal accessibility to languages in such social institutions as primary, secondary and high school (Baker 1993 for Europe, US, Canada, etc.; Escamilla and Hopewell 2007 for US; Lewis and Andrews 2014 for Wales, etc.). At the same time globalization and increases in migration have favoured the rise of interest in bi-/multilingualism of the countries, which were not concerned before (Wang and Wen 2002 for China; Georgieva 2011 for Bulgaria; Teng 2011 for Taiwan). Recently a new term has been coined: "new speakers" (Walsh, O'Rourke, and Rowland 2015). This term allows researchers to make a distinction between native speakers (both monolingual and bilingual) and people who acquire a new language/languages in different periods of life, for different purposes and through different means. The creation of a new term shows that research follows social changes and studies new directions of the development of bi-/multilingualism in society.

The number of language pairs and fields of research in bi-/multilingualism has expanded over recent years. Some of the prominent areas are bilingual discourse

(both oral and written), bilingual education and the bilingual brain. The ability to use two or more languages by a single person in a unitary speech act, i.e. code-switching or as others may call it code-mixing, was questioned in all the above-mentioned areas of research. Scientists tried to explain socio-pragmatic functions, linguistic constraints and mental schemes behind this phenomenon.

Some bilinguals do not accept code-switching in their everyday life, separating spheres of their daily activity according to the language they use. This could be the result of political, social or cultural constraints, or it could be their personal attitude to the choice and use of their languages. Yet, code-switching is a common element of life for many bilinguals. One may come across it in oral conversations, online discussions, emails, songs, newspapers, literature, etc. This enumeration is far from exhaustive. When talking to a bilingual person about their bilingual practices, they often easily accept the fact that they use code-switching in oral conversations. On a question about their writing they react more reluctantly about code-switching in written correspondence, however, in informal conversations which I have had, they freely affirm using two or more languages in their personal writing giving such examples as to-do lists or diaries. The majority of the research has been focused on code-switching in both oral and written communications which implies at least two participants: a speaker/writer and a listener/reader. Very few works have examined so-called personal writing (one exception is Montes-Alcalá 2000). The specificity of this linguistic form is that there is only one person involved, because the writer plays the role of the reader; personal writing resembles a monologue with oneself.

It is a potentially rich ground for research. This type of writing is inherent to every literate individual. It constitutes a kind of a personal space where writers feel free to express themselves without any instructions or judgement, whether moral or formal. This space is of great interest for linguists. Besides, personal writing may play the role of an intermediary between the work of the brain and the final output, it is reflected in such examples as self-corrections that are present and acceptable in personal writing and all types of drafts that one makes as a preliminary sketch of a final or bigger work. Thus the personal writing of bilingual individuals can bring a valuable contribution to psycholinguistics, providing some tangible and natural data for the analysis of bilingual production. Furthermore, individuals form communities and societies. At the same time, these social structures exert an influence on individuals, an influence which can come to be embodied in social and moral norms, culture, politics, etc. Therefore, the individual reflects the community and/or the society, which they belong to, in the same way as personal writing mirrors the individual. As a result, personal writing raises a number of sociolinguistic concerns:

How much is personal writing representative of the community/society? Which social variables affect personal writing? Does code-switching in bilingual personal writing vary with different language pairs? As paradoxical as it may seem at first, personal writing, as a subtype of writing in general, can be analysed from the sociolinguistic perspective. It is a paradox that has been picked up on and exploited by other authors: "[...] [W]riting represents a special form of linguistic performance that is unavoidably shaped by sociolinguistic conditions of its production, so texts of whatever kind also constitute a type of linguistic source that may be of interest in its own right from a sociolinguistic perspective" (Schneider 2013: 170).

Bilingual personal writing represents quite an innovative source for bilingual studies. We first need to define personal writing to allow it to be applied further to research in bilingualism. The present work will refer, from now on, to personal writing as written elements or sequences of written elements produced by an author and which initially were not intended for external readers.

The latter characteristic of personal writing, that it is not meant to be read by other people except its author, makes the collection of such data difficult. According to Montes-Alcalá, "we are dealing with highly confidential material" (2000: 35). This confidentiality may restrain researchers to work with a limited or even insufficient quantity of data due to its authors' reluctance to share their writing for scientific purposes, a problem that I myself faced, as will become apparent. The next section will elucidate the solutions, which were found to deal with this constraint.

4.2. Corpus

The novelty of the work on code-switching in personal writing has, as one inevitable consequence, the fact that there are no pre-existing corpora. First, it was necessary to define all possible types of personal writing, which could be then collected. According to the definition given in the previous section, personal writing includes diaries, personal notes, shopping and to-do lists, and all sorts of drafts. It is important to take into account that all these writing practices are not intended for an external reader. It is also important not to confuse them with such sources of writing as notes, which are used to exchange information and communicate with others; shopping lists, which can be shared with a person who one is going to make some purchases with; or drafts, which can be a part of a team work project where each member should have access to them.

The difficulty of collecting diaries and personal notes lies in the confidentiality of the information. Not everyone is willing to share their intimate

thoughts and feelings. Regarding shopping and to-do lists, these documents usually consist of isolated words or expressions. It would be a shame to limit the study to a corpus which contains only single words and strings of words, as this would obviate the possibility of analyzing code-switching in complete sentences. The only remaining alternative is drafts. They combine the advantages of the lists and diaries referred to above, because they could contain not only isolated words, but also expressions and sentences, which enlarges the variety of analyzable material in comparison with lists. Besides, this type of data is more easily accessible in a greater quantity especially in the academic environment. Students very often use rough copies to prepare their examination papers. These drafts are written for the personal use of the student and generally speaking they are read neither by other students, nor by an examiner. Thus, students' examination drafts represent the best possible source of personal writing data accessible for the purposes of this work.

Examination drafts were collected in three respective English departments of University Lumière Lyon 2 (France), Volgograd State University (Russia) and Sofia University St. Kliment Ohridski (Bulgaria). All the participants were second-language learners, because they were enrolled in the university courses of English as a first foreign language. The French students were either in the 3^d or in the 4th year of their studies. The Russian participants were either 4th or 5th year students. As for the Bulgarian data, it differs in two aspects: first, only 30% of participants were in the 4th year of studies, while 70% were in their 1st year; secondly, only 44,5% of the data constitutes rough copies, while the remaining 55,5% are fair copies. Despite the fact that the disparity of the Bulgarian data will complicate the comparison between the three groups, this data merits to be studied on its own, as will be demonstrated below.

The difference of examination rules and requirements in academic institutions sometimes worked in my favour, allowing me to collect drafts a long time after the corresponding examinations. However, in some cases they were against me, preventing any accessibility to the data. The academic policy in Russian and Bulgarian universities requires the collection of all papers including question sheets and students' drafts after some examinations, especially such as entrance and final tests. Regulations stipulate that these papers have to be kept in faculty archives for a certain period of time. This university commitment allowed me to access some drafts months or even years after the examinations had taken place. In French universities only fair copies are collected, which is why drafts had to be gathered right after examinations with the students' consent. Students were free to decide whether to hand over their drafts for my research purposes or to keep them. I was surprised that some students preferred to put their drafts into a rubbish bin in front of

me, rather than let other people access their personal notes. These drafts are not re-used afterwards by the students because they only contain the support information for the examination answer. Besides, this information is not as private and intimate as in personal diaries. Thus, I did not expect that some of the students would not like to share their examination rough copies for research purposes.

There were also attempts to gather drafts from second-language learners in other countries. However, they were unsuccessful in the three other contexts where I had the necessary contacts, namely Austria, the United Kingdom and Ireland. The Austrian academic system provides assessment through essays and reports that students prepare at home. Word processing programs, which are used nowadays to edit text documents, do not allow the genesis of the texts to be retraced. Drafts no longer exist in the context of a computer-aided editing. Thus it was impossible to collect any rough copies from Austrian students. Similarly, a contact was established with two universities in United Kingdom and one university in Ireland. This time the hold-up was in the complex and protracted administrative procedures required by the universities' respective Ethics Committees. Being external to the university, I needed to have the permission of the Ethics Committee to conduct a research project on the materials collected in the corresponding universities. Only then could I seek the students' consent to collect their examination rough copies. The procedures required to obtain the Ethic Committee's permission were too lengthy and protracted to enable me to undertake the process, in light of the time constraints for this PhD thesis.

In the end, the data collection resulted in 404 French, 93 Russian and 75 Bulgarian examination drafts, as well as 60 Bulgarian examination fair copies. It was an impossible task to have an exact balance. First of all, it was logistically difficult: I live in France and I can obtain French scripts easily, whereas for the Russian and Bulgarian scripts, I had to rely entirely on other people, despite being able to go there. Given the fact that examinations take place more or less at the same periods of the academic year in the three universities under consideration, the priority was given to stay in France during the examination periods due to the impossibility, already mentioned, of accessing drafts afterwards. Russian examination drafts, which are stored in the English department archive, have to be consulted directly on site, because it is prohibited to take them out of the university, whilst Bulgarian scripts can be taken away a certain period of time after the examination. Secondly, the English departments of the three universities varied in their size. In the 2015–2016 academic year there were 673 students enrolled for both Bachelor and Master of Arts degree in the English department of University Lumière Lyon 2 (France), 275 students in the English department of Volgograd

State University (Russia) and 134 in the English department of Sofia University St. Kliment Ohridski (Bulgaria). For example, during two consecutive summers I was provided with access to the scripts of Russian students who had taken their final examination in the specialty Linguistics or Theoretical and Applied Linguistics (TAL). These are 4 and 5 year programmes, respectively, proposed by the English department of Volgograd State University. The total number of students in the 5th year in TAL programme for five consecutive years (from 2011 to 2015) comprised 41. And the number of students in the 4th year of Linguistics for three consecutive years (from 2012 to 2014) equalled 52. Consequently, the number of drafts it was possible to collect was limited. And finally, a human factor plays a very important role. Both the access to archives, which are not public, and the possibility to attend examinations to gather rough copies immediately afterwards depended on the professors and persons in charge. Their willingness to assist was indispensable for data collection. Due to these three reasons, the balance of collected data was difficult to achieve. Even though this imbalance jeopardizes a thorough statistical analysis, the sheer quantity of texts collected is sufficient for meaningful trends and tendencies to be observed, and is also not a problem if the focus is more on a qualitative than a quantitative approach, which is what will be attempted here.

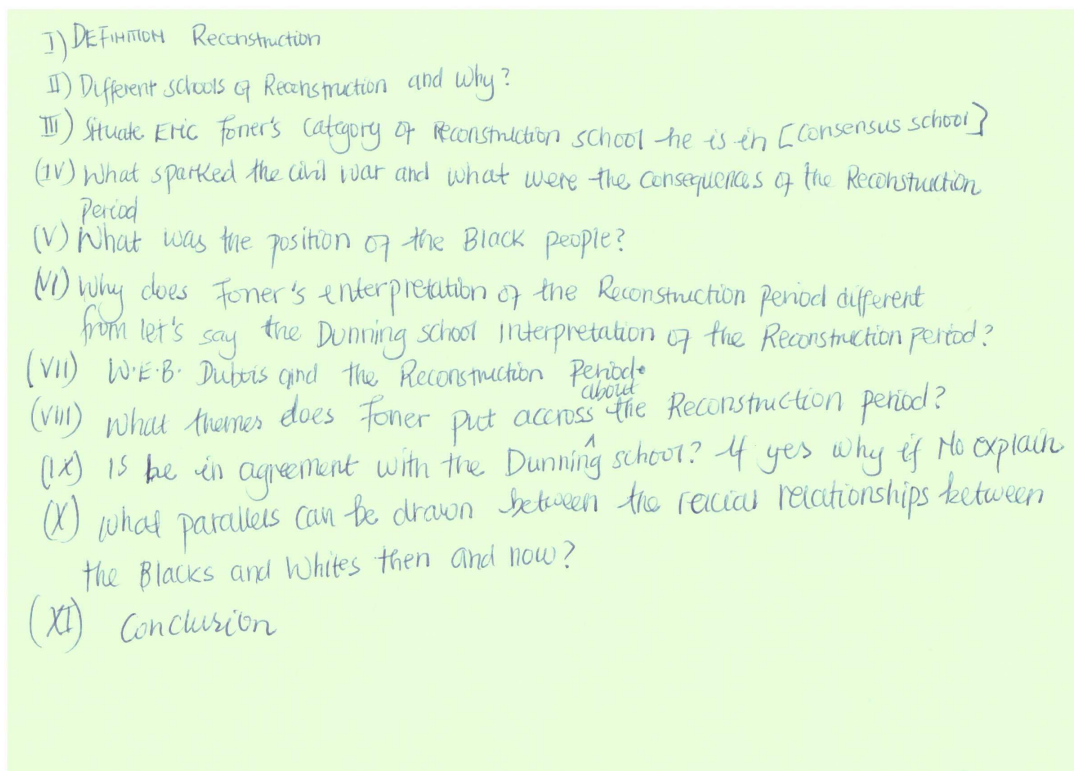
Being an adherent to the point of view that bilingualism is a relative concept, which represents a continuum (Mackey 1962/2000; Baetens Beardsmore 1982; Hamers and Blanc 1989)—a notion which was examined in Part I—I consider second-language learners in the 3^d, 4th, and 5th year of studies in the English department of a respective University to be bilingual for the purposes of this study. During the 3 to 5 years of education in the English departments of the respective universities, the students increase their English language proficiency in the four language skills. Besides, they receive a certain amount of cultural and historical information about the English speaking countries. Hence, the 3^d, 4th, and 5th year students move up in the continuum of bilingualism, moving away from the monolingual group and coming nearer to the bilingual end of the spectrum.

All the scripts were collected after the examinations, which were based on a language stimulus. By this is meant, first, that the language of the examination answer was imposed: the French and the Bulgarian students had to answer in English, while the Russian students had some tasks to be written in English and some in Russian. Second, examination question sheets and source texts were in the language in which the answer was required. (More examination details will be given in the data description for each language group). The total amount of collected data is 632 scripts (572 drafts and 60 fair copies). However, only papers with the presence of two or more languages will be analysed, these are 167 drafts and 12 fair

copies (107, 28 and 32 French, Russian and Bulgarian rough copies, respectively, and the 12 Bulgarian fair copies).

Each piece of data is a unique personal work, with its own specific handwriting features, its abbreviations and shortenings, its spatial layout, etc. For example, the length of a draft is quite personal. A rough copy could consist of a few words or it could be made of up to five pages. Besides, the handwriting of its author plays a key role. Sometimes the reading is fluent because both handwriting and the draft's structure are lucid (Image 1).

Image 1. Draft of a French student.



Even when the draft is much longer and its structure is more complicated, for example with additional lines being added as an afterthought, provided that the handwriting remains legible, it is easy to work with (Image 2).

Image 2. Draft of a French student.

Transparence and obstacle:

- 'elaborated from sight' l.10.
- 'the curtain of falling snow' l.11.
- 'a sheet of ice covering my mouth' l.18-19.
- 'the hand is thickly gloved' l.19-20.
- 'hooded figure' l.9; 'hooded face' l.33.
- 'So this what it is to see!' l.38.

- The body:

- hard to walk: l.11-12; 'trudging' l.48. 'my features are frozen' l.17-18, can't see.
- can't speak. 'No sound comes from my mouth, in which my tongue lies frozen like a fish' l.30-32.
- 'I want to speak to her through my clumsy frozen words' l.39-40., 'mumbling' l.41.
- 'my scalp feels numb with cold' l.44.
- girl 'huddled in a ball' l.45: enclosure.

Paradox: ~~the girl is~~ the blind one is the girl, but in the dream things are different: the Magistrate has to overcome countless obstacles in order to experience vision at least: 'So this is what it is to see!', while the eyes of the girl are full of light and glancing.

→ dealing with surfaces, screens, etc., yearning for ~~something~~ ^{the thing} behind the curtain.

Precisely, the girl remains impenetrable throughout the novel, an ~~opaque~~ ^{opaque} surface.

Magistrate's endeavours outside the dream too: 'I ask the girl' l.51, 'I also ask' l.56.

↳ frustrated in all his attempts to get to the heart of the matter, but also to understand hi-self: the passage begins and ends with introspection, ~~always~~ but never quite satisfactory (Irony of the 'inspection paradox' l.67).

II - Surfaces

Intro: Recurrent dream → issues central to the novel.

I. Dealing with surfaces

II. the body (dysfunctional)

III. I - possibility to write

the whole town is white and mute and empty' l.28-29.
 → cf. blank page at the end of the book?
 unwritable story of the frontier.
 'the waste of the snowswept square' l.48; 'the town she is building is empty of life' l.49-50.

how does this passage deal with explore the Magistrate's way of relating to the world, but also to hi-self - since introspection is a represent

However, the writing is not always as clear as this, far from it. After all, as we have already mentioned, there is in principle no other reader than the writer. Clarity is therefore not an over-riding concern. Nonetheless, it is very important to read the whole draft to detect the use of two or more languages, which is difficult to notice without reading. The difficulty of deciphering the handwriting can double or

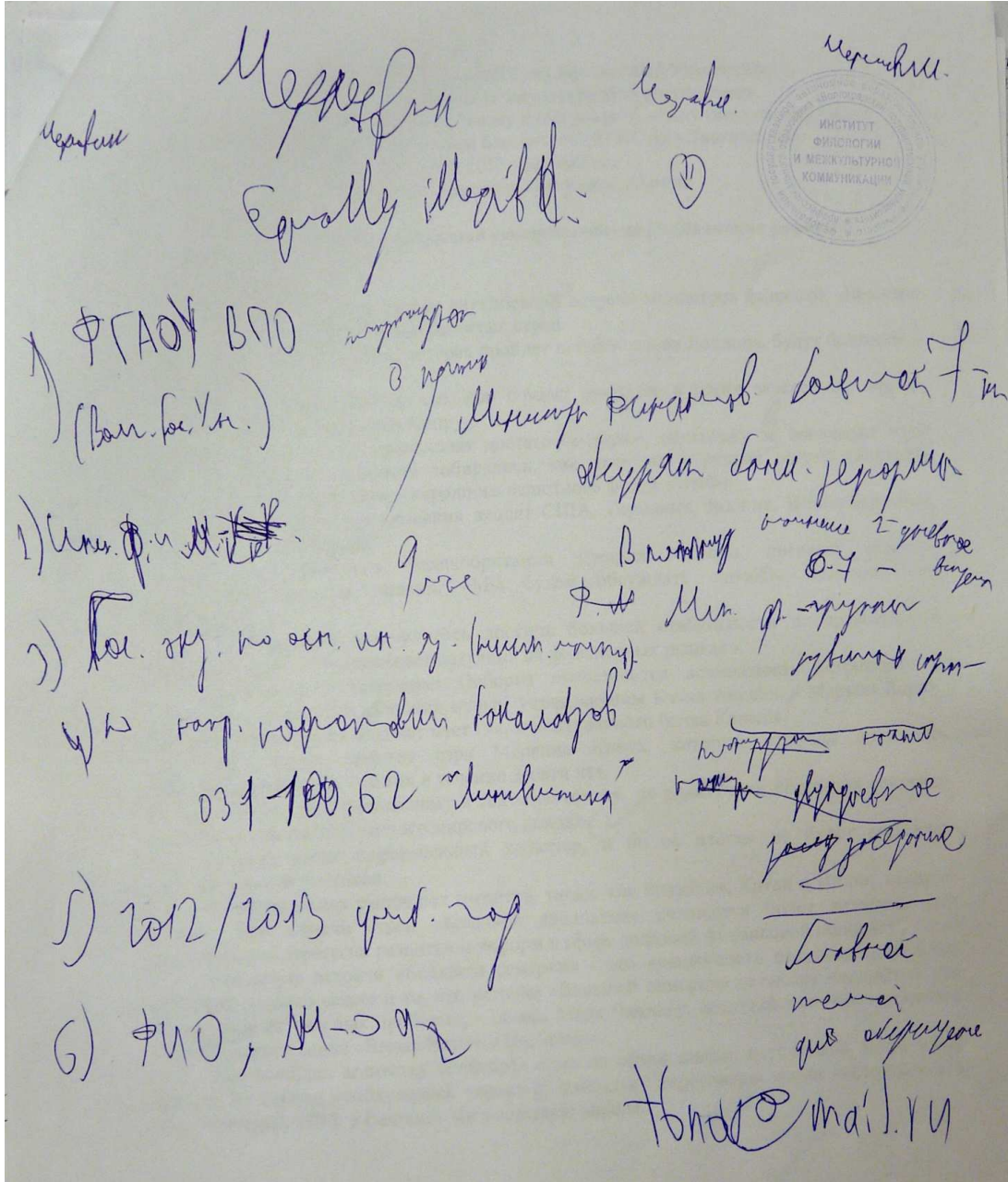
even triple the time necessary for a researcher to work over the script. Image 3 shows one page out of the total of three pages, which compose the draft in question.

Image 3. Draft of a French student.

the CFAA massacre of its freedom by white supremacist organizations
the day when "freedom died" as one journalist described that sad
event - (From goes further, and explains that modern conservatism
By the 1920s, the intellectual debate in the America was steeply
rooted in progressivism which came to establish more social justice and
efficient democracy to wash away the several of ethical and economic scandals
that surrounded the presidential cabinet for several decades since the
Grant Administration. ~~But the 1920s~~ Intellectuals, either government or
or social activists wanted to question the essence of the federal gov.
~~History as well~~ Similarly historians wanted to question the reliability
of the political history of the 19th century, in doing so, historians
like Charles Beard wrote a revisionist history paper on "the
interpretation of the declaration of independence" where he expressed the
economic factors behind the "revolutionary" event of the founding fathers
regarding the Reconstruction History, we already mentioned one
of its spokesmen historians against the so-called school of William
Douglass, W. E. B. DuBois, as the father of the civil rights movement
and African American Intellectual; ~~and~~ ~~strongly~~ ~~appeared~~ in his
historical study "Black reconstruction", He pioneered the American historians
in one major aspect that reflects the influence of progressive ideas and paradigms
in fact, He ~~tries~~ is relate the exploitation of the "dark proletariat"
to the emergence of the American industrial and capitalist power,
Here, DuBois, ~~states~~ directs his study within a Marxist
economic framework when he ~~puts~~ ~~the~~ ~~black~~ ~~slaves~~ ~~emphasizes~~
on the centrality of the black slaves within a global struggle for
emancipation - in his previous work of 1903 "The souls of the
black folks", He argued to his ^{racist} ~~opponents~~ of the Doug school
that blacks are not required superstitions, but like "Sambes".
~~DuBois~~ refers The reference of DuBois to Marxist dialectic
is one of the early forms of ~~diverging~~ ~~national~~ ~~issues~~ ~~from~~ ~~an~~ ~~im~~
other ~~perspective~~, an aspect of his writing which deserves

When a scribbled text is accompanied by an inexplicit structure and randomly placed words, the reading of a text comes even more complicated (Image 4).

Image 4. Draft of a Russian student.



All the data was collected either in the handwritten format or was photographed if the original paper could not be taken away. All the samples were carefully read. However, those that were entirely monolingual—English, French or Russian—were discarded, and only those in which more than one language was present were transcribed. Transcription is an analytical tool, which allows a transfer

of a handwritten material into a digital work scape, i.e. onto a screen and a word processing programme. Besides, it provides a clearer text perception. For example, if comparing a draft of a few pages with a hardly legible handwriting on a purple sheet of paper, with a digital transcription with no ground colour and with an easily readable typed text, one's attention is less perplexed by external factors. It favours the work with the text and sometimes even reveals details which were left unseen in the original sample. The transcription system adopted in this research attempted to reproduce, as closely as possible, the original scripts preserving authors' grapho-spatial elements, crossing-outs, grammar and spelling mistakes, etc. If language alternation appeared frequently throughout the whole paper, the entire script was transcribed. If there were a few occurrences of language alternation, then only the language alternation in its context was transcribed. The passages of language alternation are italicized in the transcription. It should be noted that the embedded language varies from script to script, thus both L1 and L2 could be found in italics depending on the matrix language of the sample. The matrix language is defined both on a quantitative principle and the order of L1 vs L2. The less represented language as well as the language, which appears second in the draft is defined as the embedded language.

For example, here is a draft from a French student (Example 1) where French is quantitatively more present than English and appears first in the rough copy; thus it is seen as the matrix language and English is embedded.

Example 1:

Civilisation Britannique
Intro - article de journal de ___ dans
le ___
A la fin des mandats de Blair
il éretiq critiq le bilan de B denexqant
qui est le titre de l'article
il lurent sur les débuts, → 1 ^e joue bcp de'epocrs
→ <i>Northern</i> sur les accomplies
ce la the la paix en Ireland Nord
ou l'économie

~~L'éc~~ ~~1~~ ~~1~~ Enfin il denoue toutes les reformes
~~fais~~ ~~1~~ qui ont eu 1 bilan très contesté.
→ *problematic*
→ *plan*

Sometimes the two principles can contradict one another. In such cases, the draft needs to be reread in order to review the language choice in its structure. For instance, the draft can start with the language A, but the whole content will be in the language B. In such situation, the quantitative principle is favoured (Example 2).

Example 2:

Brouillon [1]
[...]
Brouillon [2]
[...]

Translation:

Brouillon = draft

The next example (Example 3) illustrates a sample where the transition from English into French can be observed on the documental level. Its first part (which is also written on a separate page) is composed mainly in English and it is the first language in use; this part also contains some French code-switchings. The second part/page is dominated by neutral units. However, as English appears first in this part, it is considered to be its matrix language. And the last part/page is written almost entirely in French and it is also the French language which is the first in use in this part. In such situations, each paragraph has its own matrix language; and it is not appropriate to assign one matrix language to the whole document. Such cases of matrix language distribution are rare, only five drafts show such characteristics.

Example 3:

Bubble.
Authority ⇒ dictionary → *pas* right or wrong

→ collecting all words.

→ *pas authority superieur.*

→ check spelling & *contexte*?

liberty in speech

• evolution of language.

• English = all around the world.

- *français.*

- canadian

- british

- american (+ -

- different

→ corruption? or evolution?

- *anglais*)

→

Academie d'anglais?

~~surprisingly~~

surprisely

surpriseling

I was astonished. We're being followed. Stunned into silence.

→ preterit

→ 1 = Preterit + preterit.

2 = be + ing *à la forme passive* -ed.

3 = preterit.

1) Identification.

1) IdentificaO

2) liste des possibilités de traduction.	
3) contraintes de ses traductions.	
4) Choix de trad. justifiée.	
→ <u>la casquette</u> enfoncée...	⇒ <i>That on.</i>
→ avait <u>la peau</u> brune.	⇒ <i>be had a dark skin.</i>
→ haussant <u>les épaules</u>	⇒ <i>shrugging her shoulder.</i>
→ qu'il y a <u>les loups</u>	⇒ <i>since the wolves</i>
↳ déterminant + nom.	
2 nom singulier	2 nom masculins ^{pluriel.}
3 nom féminin	1 nom masculin

When the collection of the data was completed and the bilingual scripts were transcribed, then the transcribed data was described from three perspectives (visual, pragmatic and linguistic). The methodology of description is explained in the next section.

4.3. Methodology

Besides the non-existent corpus, the novelty of the present thesis exposes another difficulty, which is the non-existent methodology. A new corpus of examination drafts contains a lot of information of different levels such as the grammar and syntax of personal writing, language choice, organisation of ideas and grapho-spatial layout, note-taking system (abbreviations, shortenings, etc.) and so forth. It was necessary to select the information, which is significant to answer the main question of this PhD work, namely: "what is the reason for code-switching in examination drafts?".

Without wishing to go into a terminological discussion, and given that (as we saw in the first part) uniformity has not been reached yet in distinguishing the phenomenon of code-switching from code-mixing, as well as the fact that the current study does not depend on any such distinction being made, I conceive of code-switching in this work as the use of two or more languages in a unitary written act produced by a single author. The term code-switching, from now on, will be

used to designate all forms of language alternation (Montes-Alcalá 2000; Muysken 2000). Other terms such as (language) switch and language alternation can be found in the present work as interchangeable for code-switching.

To understand the reasons for language alternation in drafts, a multi-perspective approach is essential. The pragmatic perspective, to be defined below, highlights the pragmatic functions of code-switching; however, it is tightly linked with the linguistic perspective, which allows patterns of code-switching to be examined, which (in turn) often reflect pragmatic functions. In addition, the visual perspective provides some complementary information, which may either prove and clarify the use of pragmatic functions or indicate the direction, which has been left unseen.

In this work the visual perspective will be used to demonstrate whether code-switching is visually marked. By visual marking I understand any graphic and structural means used to visually emphasize language alternation. Such means include the use of capitalisation, bold handwriting, quotation marks, parentheses, underlining, framing, highlighting, the use of different colours and structural marking, a term I use to refer to when an element is marked out by its prominent position in the structure.

The methodology of the pragmatic perspective is based on the work by Montes-Alcalá (2000), where she explored socio-pragmatic functions in written Spanish–English code-switching. In her study of personal journal entries, electronic mails, letters and personal notes, bilingual narratives, magazines and literature, Montes-Alcalá proposes a list of 12 socio-pragmatic functions, inspired by the works of McClure (1981), Gumperz (1982), Poplack (1979/1980/2000), Valdés-Fallis (1976) and Zentella (1997). Adapting this list to the corpus of personal writing seven functions were selected. The remaining five functions were discarded because they were not found in the present corpus.

1. Direct quotation: another language is used to cite an original document with or without the employment of quotation marks.

2. Clarification and/or elaboration: another language is used to work out the given idea with more details or to make it more clear.

3. Parenthetical comment: another language is used to give qualifying or appositive information both with and without parenthesis; "code-switching constitutes a sort of 'aside'"(Montes-Alcalá 2000: 38).

4. Triggered switch: a switch is provoked either by another lexical item of the same switch or a neutral item, which precedes or follows it.

5. Stylistic switch: another language is used for a stylistic purpose; "[...] to convey the message more effectively than if it was a monolingual text" (Montes-Alcalá 2000: 41).

6. Lexical need: another language is used either because of a lack of an exact equivalent or to fill a memory lapse.

7. Free switching: another language is used without any explicit pragmatic function.

The seven above-mentioned functions do not cover all the pragmatic uses of language alternation in the collected corpus of personal writing. Seven more pragmatic functions were added to embrace the specific features of writing in examination rough copies. The functions reflect the structural and semantic features of draft composition. While the functions numbered 4 and 5 below are also observed in the oral bilingual communication (Gardner-Chloros 2009; Bernard and McLellan 2014) and in online discussions (Halim and Maros 2014), the remaining five functions (numbers 1, 2, 3, 6, 7) were elaborated specifically to explain the language alternation in the students' examination drafts.

1. Framing: another language is used to give administrative and functional information, which is not related to the content of the examination answer. Rough copies contain two types of information: the content of the examination answer and the external information such as the student's name, the title of the examination, the number of words, etc.

2. Structuring: another language is used to organize the ideas in the content of the examination answer. For example, structuring may be achieved through such words as "Introduction", "Part I", "Conclusion", etc.

3. Terminological switch: a switch in a term related to the course in which the student sits an examination. The terms in both languages are available for the student, but they give preference to the term in a particular language.

4. Language economy: another language is used for lexical items, which are shorter than their equivalents in another language and efficient in use.

5. Reiteration: another language is used to repeat the information that has already been written.

6. Vocabulary list: another language is used either in providing a lexical item or its translation in composing a sort of a dictionary entry.

7. Spelling check: a switch in lexical items, which serve to recall a correct spelling and which are written apart from the main text.

All occurrences of code-switching in the collected data will be analysed according to the 14 above-mentioned functions. The description of the data according to the pragmatic functions requires understanding of the text under analysis. To make this information accessible for the reader, the passages in French, Bulgarian and Russian will be provided with a translation into English made by the author of the present work.

The third, linguistic perspective will examine code-switching on a textual level. The switch may occur within the boundaries of a word, within the boundaries of a sentence and between neighbouring sentences. The notion of an intra-word switch and the notion of a sentence are both quite dubious as applied to the corpus of examination drafts.

Code-switching inside a word is called *word-internal* or *intra-word code-switching* or it is sometimes referred to as *blending*. The phenomenon of blending represents grammatical adaptation of a lexical item in another language to the language of the utterance—or a situation when part of one word in language A is joined to part of one word in language B (Hamers and Blanc 1989/2000). The term blending will be avoided in what follows, given its ambiguity. It is of course also used by lexicologists to refer to situations where a new word is created by the combination of elements from two existing words in the same language, such as *motel* (motor + hotel), *malware* (malicious + software) or *Spanglish* (Spanish + English). For this reason, only the term word-internal code-switching will be used. Given that drafts provide a space both for a preliminary outline of ideas and grammatical and spelling check, the distinction between word-internal code-switching and a spelling mistake or ignorance of a word can be difficult to make. Whereas in the Bulgarian and the Russian data the use of two alphabets facilitates the distinction between a switch and a spelling mistake, in the French rough copies the situation is more complex. On the one hand, the omission of diacritics in a French word may turn it into an English lexical item, on the other hand, the addition of the final "-e" or double "-ll-", "-nn-", etc. may transform an English word into a French item (e.g. "paradox" vs "paradoxe"). And if several such changes are performed in one word, this lexical item does not belong anymore to either of the used languages. For instance, one of the French samples contains the word "extremiste"; this word is neither the French "extrémiste", nor the English "extremist". Another example is the word "constitutionnal", which is either the English word "constitutional" with a double "-nn-" as a spelling mistake or a word-internal switch of the French word "constitutionnel" and the English equivalent. Taking into account that writing in examination drafts is close to note-taking and thus presents such features as shortenings, abbreviations, auto-correction, etc. the

limit between a conscious or even unconscious language switch and a spelling mistake for an external reader is tiny. For this reason, the cases of diacritic omission, letter doubling and omission or addition of the final "-e" will be considered as spelling mistakes rather than word-internal code-switching.

Another feature of drafts is unfinished and incomplete sentences. Quite often a new idea starts with a new line without any full stops at the end, or in a list format (Example 1).

Example 1:

A cet égard, problématique → in his letter from
Birmingham City Jail, what is MLK trying to
advocate & explain ?

Ebauche de plan

I. Figerting injustice : a matter of survival

A ~~Ye~~A time of violence & ~~strifes~~ and struggle

B Non-violent dvect act° suitable program

Translation:

A cet égard, problématique = *With regard to this, problematic*

Ebauche de plan = *Rough outline of the plan*

When the writing is structured this way, it creates a problem for applying the intra-/inter-sentential classification of code-switching, which is based on the notion of clause or utterance. In cases where the majority of the draft is composed of ideas and incomplete sentences, clauses are not always distinguishable; they are often connected or divided by graphic signs such as =, →, ⇒, &, +, etc. and a punctuation mark of colon. Without wishing to go into overinterpretation of such signs to decide whether the string of words should be considered as a coordinate or subordinate clause, I adhere to the notion of utterance in the same way as Wardhaugh (2006). However, as Wardhaugh does not provide a definition of utterance, we can turn to Harris, who distinguishes utterance from sentence explaining that an utterance may consist of a word or even unfinished sentences. He defines utterance as "any stretch of talk, by one person, before and after which there is silence on the part of that person" (Harris as cited in Lyons 1968: 172). The utterance can be characterized by the syntactic ambiguity unlike the sentence. The utterance should be regarded as a

statement¹¹. If the sentences are not separated by a full stop, then each free-standing line, unless they represent one semantic unit, or each point of a list will be regarded as a separate utterance. For instance, Example 2 illustrates an excerpt from a French rough copy, which contains eight utterances. Each utterance is marked with a red number (added by the author of the present work).

Example 2:

1 §4 – “its fear... W class” → ↑ WC, < opposit° (b) classes	
2 → <i>a-t-on tjrs des classes</i>	6 - Society = < citizens with identity
3 → <i>evolut° dans le temps</i>	
4 rise middle class	7 - <i>evolut° identity in time</i> → to in time and Social & Societal relationships
5 loosing the idea of class	
	8 - Hope for a loss of it?

Translation:

a-t-on tjrs des classes = are there still classes

evolut° dans le temps = *evolut°* (shortening from evolution) in time

Thus intra-sentential code-switching will henceforth refer to switching within utterances, whereas inter-sentential code-switching to switching between utterances. Code-switching between a main clause and a subordinate clause will be considered as an intra-sentential switch, because they are related both semantically and syntactically, in the cases where the sentences are grammatically correct (Treffers-Daller 1992: 144). No switches between coordinate clauses were detected in the present corpus. Yet, the dilemma lies in the use of an introductory colon. This punctuation mark often introduces a next section or part of the student's answer (e.g. "Introduction: ", "Conclusion: ", etc.). The notion of utterance is still more suitable than division into clauses. Hence, if an introductory colon is followed by a phrase or a sentence situated on a same line, it will be considered as a single utterance. If an introductory colon is followed by several sentences or a paragraph situated on different lines, it will be considered as two or more separate utterances. Here is an example of intra-/inter-sentential code-switching in the utterances with colon (Example 3 and 4).

¹¹ In *Merriam-Webster* [Online], retrieved 24/10/2016, from <http://www.merriam-webster.com/dictionary/utterance>

Example 3. Intra-sentential code-switching.

Pbtique: In which way these histo. try to define Am. identity

Translation:

Pbtique = *Problematic* (shortening from "Problématique")

Example 4. Inter-sentential code-switching.

Intro rédigée:

This passage is an extract from Act III scene 3. [...].

Translation:

Intro rédigée = *written introduction* (shortening from "introduction")

After having clarified the methodology of the data description according to the three perspectives, I would like to draw your attention to the fact that each sub-corpus (French, Russian and Bulgarian) comprises features which should be taken into account when studying code-switching. First, the number of neutral items varies due to the alphabet in use and the type of examination. The French data consist of a much greater amount of neutral items than the Russian and the Bulgarian data. Besides the French and English words, which have an identical spelling, omissions of diacritics and shortenings may transform a word into a lexical unit, which equally belongs to both languages (e.g. "plan", "opposit° (b) classes", "Intro", "societal", "defense", etc.). Inasmuch as an external reader will never know in which language the neutral item was written, the neutral items, which are detached from an utterance, will not be reckoned as a switch. However, the neutral items, which are situated within the boundaries of an utterance, will be considered as items belonging to the language of an utterance. Furthermore, the neutral items, which stand at the beginning and the end of a switch, will be conceived of as triggers (Clyne 1980). Second, the titles of laws, historical events, companies, books, etc., which may even have an official translation into French, Russian and Bulgarian, and which appear in the examination task, source texts and/or were used in the course during the semester, will not be counted as a code-switching; because these are proper names, which were delivered in English during the course and/or examination and which could be elements of priming (e.g. "Civil Right Act", "Lenovo", etc.); however, they may trigger a switch. Third, the Russian examination consisted of several parts; each part was entitled either in Russian or English. If the title of the task is repeated in

students' drafts, it will not be regarded as code-switching to avoid the effects of priming (e.g. "Аннотация", "Реферирование", etc.).

The description of the corpus following the visual, pragmatic and linguistic perspectives of code-switching will be presented in respective chapters devoted to the French, Russian and Bulgarian data. The opening of each chapter will provide the numbers of the collected samples and the details of the examination procedures.

Chapter 5.

French data

Four hundred and four French drafts were collected after written examinations, which were held in the English department of University Lumière Lyon 2 (France) from 2013 to 2016. The students who participated in this study were either in the last year of their Bachelor of Arts degree (3^d year of studies) or in the first year of their Master of Arts degree (4th year of studies) in Foreign Languages and Cultures, specialty English studies; 252 and 152 students, respectively.

The drafts were collected after written examinations in various courses, such as English and American history, literature of English speaking countries, sociolinguistics and English as a media language. All the courses were read in English throughout a semester. During the examinations students had to comment a text and/or to give explicit answers to some questions. Both the examination instructions and the given texts were in English. Students' answers were also required to be written in English. Nevertheless, 107 out of 404 students used both the English and the French languages in their examination drafts. Table 1 shows the distribution of language alternation according to the year of study:

Table 1. The ratio of French/English bilingual drafts to the year of university enrolment.

	All collected drafts		Bilingual drafts	
		% (of total)		% (of the respective year)
3 ^d year	252	62.4%	66	26.2%
4 th year	152	37.6%	41	27%
Total	404	100%	107	26.5%

Just over a quarter of students code-switched in their drafts, which intuitively constitutes quite a significant number. Moreover, the distribution of bilingual drafts in the 3^d and the 4th year is almost equal 26.2% and 27%, respectively.

5.1. Visual perspective

French drafts are rather homogeneous from the visual perspective. First, both the French and the English language use the Latin alphabet. Second, the majority of the French students do not highlight code-switching. Thus, the switch from one language to another is imperceptible and one needs to read the draft extremely closely to perceive the coexistence of two languages. Here is an example:

Example 1. French draft.

Extrait de "Letter from Birmingham City Jail

16 avril 1963 → en plein dans le mvt de déségrégation

Civil Right Act of 1964

Lettre de MLK, leader du SCLC, figure de la lutte afro-américaine.

→ à ses collègues pasteurs des autres communautés : chrétiennes, juives.

Au moment où il écrit sa lettre, il est en prison dans la ville de Birmingham, AL. ~~B~~ Il est en prison pour ses actions

Birmingham → ville la + ségrégationniste des USA.

Plus qu'une demande particulière, ~~le texte~~ la lettre critique les religieux et les modérés.

PB : ~~It~~ ^{to by the condemn} what extent does Martin Luther King, in his letter, justify ~~his actions~~ ^{the need for a direct action} in the name of freedom † and justice?

I/
A
-
-
-

B disappointed promises

by the religious, by the white moderate, leaders of the eco community

wait → tranquilizing thalidomide → emotions

because negotiation fail → need for a direct action

II / ^{The} Need for a direct action

A

4th basic step : direct action because nego failed →rappel to
recent history : freedom rides, Montgomery Boycott

B But ~~the~~ direct action : non violent

opposition to several violent groups : Black Panthers, Malcom X → spoke to
clergymen, non christian, justify by religion god created equal.

The language switch in the draft, represented in Example 1, occurs between different parts of the text as well as inside them. The first half of the draft is written mainly in French giving a general description of a historical situation. The second half represents a plan and the problematic of the text, both of which are written in English. However, a few cases of code-switching are found in both parts of the draft. Example 1.1 illustrates code-switching, which occurs at the beginning of the sample (henceforth code-switched elements are italicized):

Example 1.1.

16 avril 1963 → en plein dans le mvt de déségrégation

Civil Right Act *of* 1964

Lettre de MLK, leader du SCLC, figure de la lutte afro-américaine.

Translation:

16 April 1963 → in the middle of the movement (shortening from "mouvement") of desegregation

Civil Rights Act *of* 1964

Letter from MLK, leader of SCLC, figure of the Afro-American struggle.

It may seem that these three lines are written in French. However, there is the English preposition “of”, which is triggered by the title "Civil Right Act"¹² and which may ascribe the whole line to the English language. Of course, the latter is a hypothetical observation of a researcher without any possibility to check in which language the student was writing it.

¹² As was explained in Section 4.3. proper names are not taken into consideration in the analysis to avoid the effect of priming. Nevertheless, they may perform the function of a trigger for the neighboring elements.

The next example illustrates code-switching at the very end of the sample:

Example 1.2.

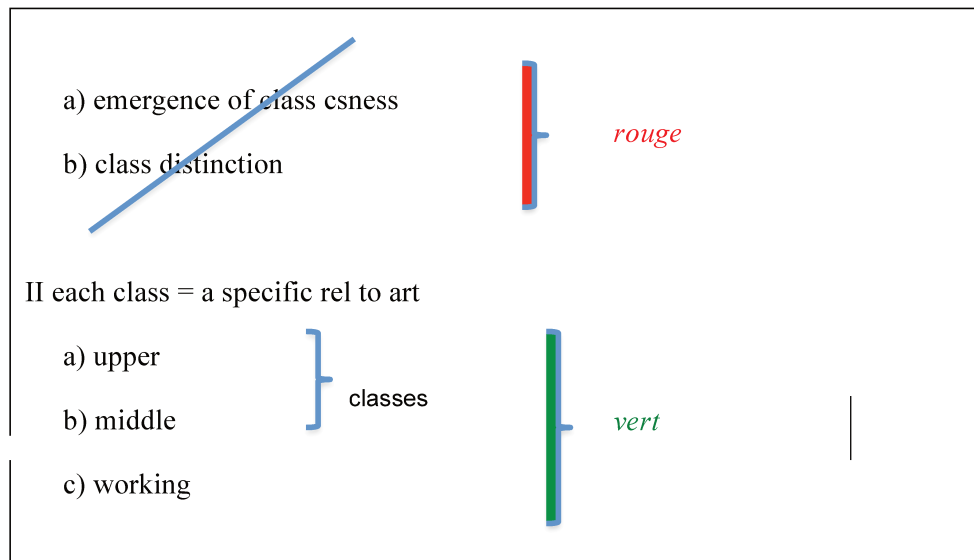
4th basic step : direct action because nego failed → *rappel* to

Besides shortening the word "steps" to "step" and the word "negotiation" to "nego", the student also switches into French using a French word "rappel" instead of its English equivalent "reminder".

Almost a half of French students visually mark code-switching, 45 out of 107 samples contain marked language alternation. By visual marking I mean the use of graphic and structural methods to make code-switching stand out of the text. The French drafts contain such types of marking as the use of capitalisation, quotation marks, parentheses, underlining, framing, highlighting, the use of different colours and structural marking. Such examples are demonstrated in Example 2 below.

Example 2. Different types of visually marked code-switching.

1. *Using colours and structural marking.*



2. *Underlining.*

Classes au service les unes des autres (domestic servants) & then art (provided by the M-class)

3. Structural marking.

accroche term coined in the 1830s, first used for the M-CI –CI fundamental to und pol, sec, cult evo in the UK & social catego.

[...]

pbq In that aspect I shall ~~argue that~~ explain how B s^{ty} had evolved for a csness so^{ty} to a clsless s^{ty} with cult identificat°

4. Highlighting with colours, parentheses and structural marking.

1780 – 1830 = I. Rev°

1830 anthro of classes

middle classes

(inné) (acquis)

fin commentaire en ouverture

5. Quotation marks.

VI Academy of the Anglo-American language.

No. This isn't a salvation. They would create a "standard English" which would be taught to everyone & to strangers too. // *Académie Française* & "*bon français*". → there is NOT one "standard" language bc we all speak differently, there are ≠ accents + dialects....

6. Parentheses

x Richard 3 (2 comédiens → contemporain), debut 2^e semestre (master).

7. Framing, capitalisation and parentheses.

SAUTER DES LIGNES !!!

_ 7 rhetorical devices (tous différents!) _

Trump = opportunist _

Nonetheless, students who apply visual marking to code-switching represent the minority. They constitute 42.1%. Table 2 illustrates the ratio of marked code-switching to the total amount of marking.

Table 2. Visually marked code-switching.

	All code-switching is visually marked	Some code-switching is visually marked
Monolingual content is also visually marked	17	26
Only code-switched content is marked	1	1
Total (N ^o of drafts)	18	27

Only 1 student out of 107 applied visual marking to all their code-switched passages without applying it to any other elements of the draft; while 43 out of 45 students who visually marked the language alternation applied visual marking both to code-switching passages and to monolingual content. Thus the numbers suggest that if students use visual marking in their drafts, they apply it to the content of the answer where code-switching may be one of its components. There is no direct connection between visual marking and code-switching. This is itself a potentially interesting finding. Given the apparent lack in the literature of similar work, taking the visual perspective into account, and in light also of the fact that this was not the central focus of this work, as initially conceived, it is a finding that will perhaps go unexplored here. This is a potential path for further research.

5.2. Pragmatic perspective

5.2.1. Pragmatic functions

107 bilingual French drafts represent 13 pragmatic functions, which vary in their frequency quite significantly. Each function will be illustrated with examples below. Afterwards, a summary table will provide the numbers of occurrences of each function.

1. Free switching.

Example 1:

women 1918 *droit de vote*

Translation:

women 1918 *suffrage*

Example 2:

Intro Conc plan *détaillé*. 64/69 woods toek.

the same year

a few years before dying. Kennedy.

Translation:

plan *détaillé* = *detailed plan*

Example 3:

But MK tried to adress the whole mat° to push them closer to Black cause (mo
agress°; he explains the motivat° of his people, w/ creating an agress° against
the white)

Si le mnde avance, prk sommes-nous en retard ?

désolidarisat°

I. The How to pressure the white about the Non-violent intentions of the Martin
K.'S followers: [...]

Translation:

Si le mnde avance, prk sommes-nous en retard ? = *If the world*
(shortening from "monde") *is moving forward, why* (shortening from
"pourquoi") *are we staying behind?*

désolidarisat° = *lack of solidarity* (shortening from *désolidarisation*)

Example 4:

Newlabour = New Right + old labour. → keeping much of NR policies.

§1: election of T.B.

Rhetorical question.

Ferme la boucle à la fin de l'article.

I. T.B's ^{election} & New Labour

[...]

Translation:

Ferme la boucle à la fin de l'article = Closes the loop at the end of the article.

Example 5:

I IMPORTANCE OF THE KNOWLEDGE OF THE PAST THAT
INFLUENCE THE PRESENCE AND THE FUTURE

1. Importance of the knowledge of the past

A. *Histoire Source de savoir*

B. Pre-colonial - culture

1. Pre-colonial culture of education

2. Importance according to the text

3. Influence on the ~~past~~ _{present} and the future

Translation:

Histoire Source de savoir = History - Source of knowledge

2. *Stylistic switches.*

In the first example the student switches into French in the second line. The word "constat" represents a stylistic choice of the student. The French word implies both meanings—"observation" and "report"—which makes it difficult to choose one English word instead.

Example 1:

I The injustice of segregation

A) *constat de l'injustice.*

B) Wait

Translation:

constat de l'injustice = acknowledgement of injustice.

The second and the third examples show a switch, which appears several times in the French corpus. The French word "bilan" incorporates the meaning of such English words as "appraisal", "annual accounts", "results", and "resumé".

Example 2:

II Positive *bilan* of British economy

Example 3:

Bilan

I -

Par. 3 = New Labour "This time it would be ≠".

Another French constructions such as "mise à/en ..." and "rapport de ..." remain untouched in the students' drafts (Examples 4, 5, 6).

Example 4:

mise au point sur knowledge

Translation:

focus on knowledge

Example 5:

1) Staging of the scene.

- public & decor.

- oral person. TMTTC bbey

- *mise en abyme* B avec livre ds livre.

Translation:

mise en abyme [B] *avec livre ds livre* = *mise en abyme* [B] *with a book inside (shortening from "dans") a book*

Example 6:

B) Cultures clash

→ W/Af → *idée de rapport de force / influ colo/.*

→ "Rehabilitate = undo cob = own idtity = dupt.

Translation:

idée de rapport de force = *idea of ratio of power*

Several stylistic switches are due to the English words of a particular register such as "storyteller", "gentleman", "clergymen", and others (Examples 7, 8 and 9).

Example 7:

- *perte substance, reproduction non fiable,*

→ *alors que l'homme y arrive, lui* → *irréel*

storyteller, représentations vivantes → *artiste créateur*

→ *voyageur*

Translation:

- *substance loss, not reliable reproduction,*

→ *so that a man arrives there, he* → *fictional*

storyteller, live performances → *artist creator*

→ *traveller*

Example 8:

recherche de réalisme

(pas d'apparition de l'auteur)

reliable English aristocrat gentleman

descriptions très précise / pas du narrateur,

prend le temps de décrire

Translation:

search for realism

(no signs of the author)

reliable English aristocrat gentleman

very precise descriptions / no narrator,

takes time to describe

Example 9:

1. Un discours envers *clergymen* pr se justifier

Translation:

1. A discourse towards clergymen to justify oneself

3. *Triggered switches.*

The triggering word is in bold typeface (added by the author of the present work for easier visual representation).

Example 1:

Aristocracy / Middle-class / W class



rassemblement

en 1832 = Great **Reform Act**

puis *séparation après ça* → W Class → Chartism in reaction to the Great R. Act (they were left out) → Failed (seen as rebels, threat to the established orders).

Translation:

rassemblement = gathering

en 1832 = in 1832

puis *séparation après ça* = then separation after that

Example 1 contains three cases of code-switching. The first one—the word "ressemblément"—represents a free switch and at the same time it triggers the beginning of the next phrase in French "en 1832". The third case of code-switching—"*puis séparation après ça*"—is triggered by the title of the law, which keeps its spelling in French "Le Reform Act".

In the next example the title of the law triggers a switch into English unlike the previous example where the English title triggered a switch into French.

Example 2:

16 avril 1963 → en plein dans le mvt de déségrégation

Civil Right Act of 1964

Lettre de MLK, leader du SCLC, figure de la lutte afro-américaine.

Translation:

16 April 1963 → in the middle of the movement (shortening from "mouvement") of desegregation

Civil Rights Act of 1964

Letter from MLK, leader of SCLC, figure of the Afro-American struggle.

Quite often a draft contains several occurrences of code-switching, which represent different pragmatic functions. Example 3 illustrates an excerpt of such draft.

Example 3:

Blairism, not an efficient gov. passivity.

I

II/ Internal Politic

economy, justice,

Maire de Londre un clampin

III/External politic

Iraq, follow *Ricain petit chien chien*

Scotland. **IRA**, *rien en 98 ca a pas marché. Cénariot*

Sant Andrews agreement. *S'obstine.*

Translation:

Blairism, not an efficient gov. passivity.

I

II/ Internal Politic

economy, justice,

Mayor of London an ordinary man

III/External politic

Iraq, follow *the poodle of the Yank*

Scotland. IRA, *nothing in 98 it didn't work. Scenario*

Sant Andrews agreement. *Persist.*

The line "*Maire de Londre un clampin*" shows a triggered switch "*Maire de Londre*", which precedes a stylistic switch "un clampin" by which it is triggered. The next occurrence of language alternation is a stylistic switch: "*Ricain petit chien chien*". The third code-switching is triggered by the name of the political party; the switch follows the triggering word: "*IRA, rien en 98 ca a pas marché. Cénariot*". And the final line contains a free switch "*S'obstine*".

4. Clarification and/or elaboration.

Example 1:

Transit^o : *entre les parties*

Translation:

Transit^o (shortening from "transition"): *between the sections*

Example 2:

2. inkhorn = *mot savant qui vient du latin bcp introduit ds la langue anglaise*
au 18^{ème} -19^{ème}
↳ *Shakespeare s'en moque.*

Translation:

2. inkhorn = *scholarly tem which comes from Latin much* (shortening from "beaucoup") *introduced in* (shortening from "dans") *the English language in the 18th - 19th*

↳ *Shakespeare makes fun of.*

Example 3:

→ Engl doesn't always demand agreement my little ponies as opposed to french

one could argue everyone represent many people : should be plural

No one → *privatif pas de raison qu'il soit sing ou plur ni masc ou femin.*

Translation:

privatif pas de raison qu'il soit sing ou plur ni masc ou femin = privative no reason to be either singular (shortening from "singulier") *or plural* (shortening from "pluriel"), *or masculine* (shortening from "masculin") *or feminine* (shortening from "féminin")

Example 4:

Logos = *Logique, clarté.*

Alliteration = *consonnes pareilles*

✓ ethos = *crédibilité*

Assonance = *voyelles pareilles.*

Pathos = *sentiment*

hyperbolis = *exagération → paroxysme*

1-3 sentences + quote for analysis. ✓ ~~Time~~ Allusion

✓ irony : fate, situationnal, dramatic. ✓ satire

✓ Repetition

✓ irony

Translation:

Logique, clarté = logic, clarity

crédibilité = credibility

sentiment = emotion

consonnes pareilles = the same consonants

voyelles pareilles = the same vowels

exagération → paroxysme = exaggeration → paroxysm

5. Structuring.

One of the most recurrent cases of structuring is the use of the French word "problématique" before explaining the problematic of the given text. This word is represented both in full and different shortened forms (Examples 1, 2, 3 and 4).

Example 1:

Pblmtq: To what extent did the ~~aristocrats~~^{upper class} have less and less influence on politics and social classes in Britain, which makes a great difference today.

Example 2:

P^{que} ~~He~~ To what extent the author of this article ~~makes~~ draws a critic of T.B's work during his term?

Example 3:

Problématique : ~~Non-violence was the method chosen to show discontent, was it viewed as non-violence? No — MLK writes to explain hotheir non-violence...~~
~~Plan~~ : Their actions where viewed as "unwise + untimely." Does MKL manage to justify ~~h~~that the opposite is true?

Example 4:

INTRO

accroche term cl coined in the 1830s, first used for the M-Cl –Cl fundamental to und pol, sec, cult evo in the UK & social catego.

doc deal with the evo of cl in B^{sty} = either at its stake^(19th – early 20th) or absent for s^{ty} is a ty written by MB in... , & the new place cult has taken.

explicat^o The author ment^o cl as being ~~abs~~ This fall in cl csness may be ~~unlied~~_{unleashed} ~~bt~~oy the rise of the cult , ~~ana~~as a new aspect of soc catego.

pbq In that aspect I shall ~~argue that~~ explain how B s^{ty} had evolved for a csness so^{ty} to a clsless s^{ty} with cult identificat^o

plan

While the shortened words "INTRO", "doc", "explicat°" and "plan" are neutral items, Example 4 contains two cases of French code-switching; these are "accroche" (hook line) and the shortening "pbq" (problematic).

Example 5:

Ebauche de plan

I. Fighting injustice : a matter of survival

A ~~Ye~~A time of violence & ~~strifes~~ and struggle

B Non-violent direct act° suitable program

II. ~~seeking Enhancing his posit° towards~~ Fighting injustice : enhancing & clarifying his posit°

A Disillus° & disappointment

B Internal | divisions
 | strifes

Plan l'intro

→ The 2nd WW restored America's grandeur along – [...]

Translation:

Ebauche de plan = Rough outline of the plan

Plan l'intro = Plan of the introduction (shortening from "introduction").

Example 6:

Représentation d'un personnage

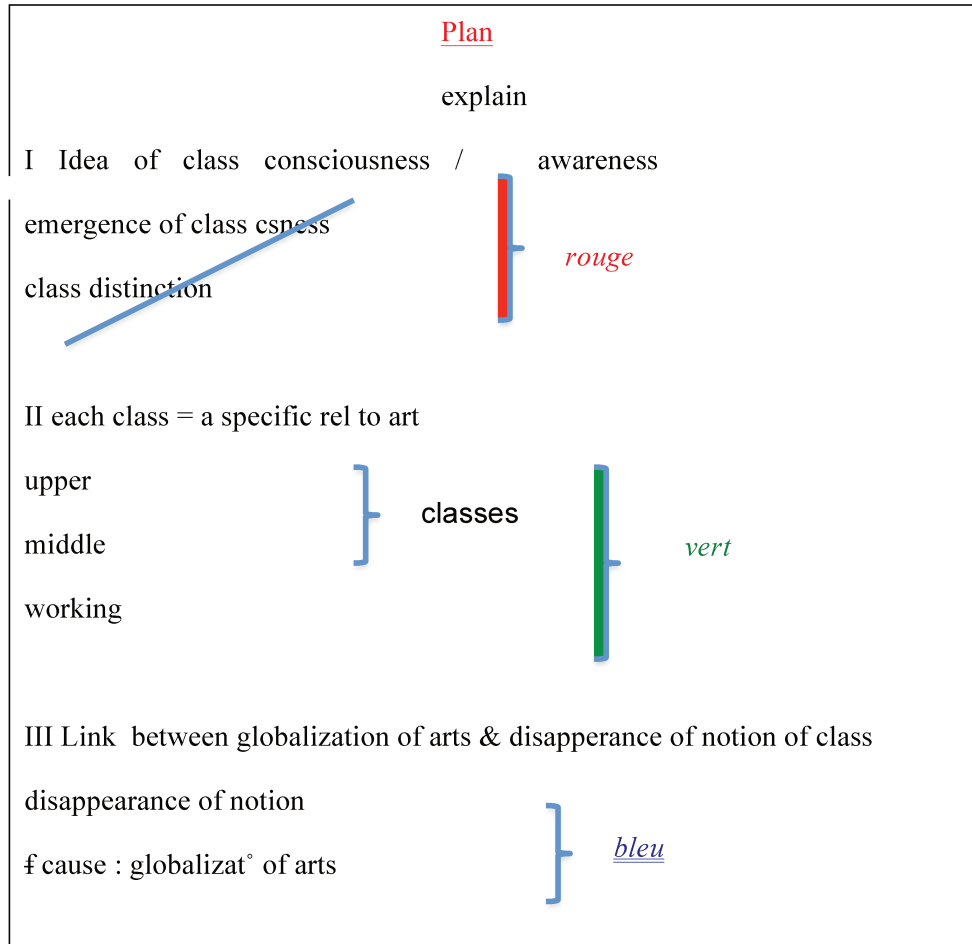
a new figure → vague - mysterious (silhouette)

Translation:

Représentation d'un personnage = Character's representation

6. Framing.

Example 1:



Translation:

rouge = red

vert = green

bleu = blue

Example 2:

- party dealignment also a factor to soc class it with culture nowadays.
+ *autre feuille*.

Translation:

autre feuille = another sheet of paper

Example 3:

BROUILLON

Translation:

DRAFT

Example 4:

Brouillon 1

[...]

Brouillon 2

[...]

Translation:

Brouillon = draft

7. Direct quotation.

All quotations in the drafts represent direct quotations.

Example 1 contains one direct quotation in the line three and another one in the line five. As for the first line, it depicts a stylistic switch and the second line a free switch.

Example 1:

-
- 1. Un discours envers *clergymen* pr se justifier
- *beginning of the text*
- “you”
- étapes
- “my Christian & Juish brothers”

In Example 2 the student surprisingly quotes the original document translating the original text into French.

Example 2:

Doc A
→ forgot. pharaoh. > books pyramids. <i>"Notre histoire"</i> River <i>"apport des civ Af"</i> Civ Eur Antiq.

Translation:

"Notre histoire" = "Our history"

apport des civ Af = contribution of African civilisation (shortening from "civilisation Africaine")

In the next example the student provides the original expressions (without using quotation marks) and translates them.

Example 3:

1) IdentificaO	
2) liste des possibilités de traduction.	
3) contraintes de ses traductions.	
4) Choix de trad. justifiée.	
→ <u>la casquette</u> enfoncée...	⇒ <i>That on.</i>
→ avait <u>la peau</u> brune.	⇒ <i>be had a dark skin.</i>
→ haussant <u>les épaules</u>	⇒ <i>shrugging her shoulder.</i>
→ qu'il y a <u>les loups</u>	⇒ <i>since the wolves</i>
↳ déterminant + nom.	
2 nom singulier	2 nom masculins ^{pluriel.}

3 nom féminin	1 nom masculin
---------------	----------------

Translation:

- 1) Identification (the original word is a shortening)
- 2) liste of translation possibilities.
- 3) constraints of these translations.
- 4) Justified choice of translation (the original word is a shortening).

→ a <u>baseball cap</u> pulled down...	⇒ <i>Ø</i> hat on.
→ had a dark <u>skin</u> .	⇒ be had a dark skin.
→ shrugging <u>the shoulders</u>	⇒ shrugging her shoulder.
→ that there are <u>the wolves</u>	⇒ since the wolves

↳ determiner + noun.

2 singular noun	2 noun masculine ^{plural} .
-----------------	---

3 feminine noun	1 masculine noun
-----------------	------------------

8. Lexical need.

Example 1:

4.He blames his (*interlocuteur*)
 their inactions.
 their accusations.

Translation:

interlocuteur = *interlocutor*

Example 2:

BiLAN=

Intro

This text is an extract from "gap between rhetoric and achievement defined this era" written by Philip Johnston in the Daily Telegraph. It has been published on

May 10th 2007, at the end of Tony Blair's last term, as a nutshell "*BiLAN*". The author : point is to show the ≠ between ~~what has been done~~ the *promesses* and what has been done ~~during~~ between 1994 and 2007 ~~on~~^{under} Blair's government. [...].

In Example 2 it is necessary to distinguish the stylistic function of the word "*BiLAN*" (*summary*) from the function of the lexical need. The first occurrence of "*BiLAN*" is used stylistically, while its second appearance emphasized by the quotation marks demonstrates the student's failure to find its French equivalent. The third case of code-switching in Example 2 is the French word "*promesses*". The lack of the quotation marks around it in the text raises a supposition that the student code-switched unconsciously.

Example 3:

The way we would accomodate our speech to certain specific situations such as : job interviews or police *arrestation*. How we would standardize our speech to make a good impression.

Translation:

arrestation = *arrest*

Example 4:

- I agree with you , "correct" Eng is **needed** *définissable* .

Translation:

définissable = *definable*

In the next Example 5 it seems that the author wrote down the idea in French trying then to find the best way to express it in English.

Example 2:

7) Why do you think we always refer to "THE" dictionary when there are in fact dozens of dictionaries.
→ **concept** of dictionary: giving a set of tools/words which are widely acknowledged as correct although some might differ from one another.
→ In Fr - *Académie Française*.

Translation:

Académie Française = *French Academy*

Example 3:

x → sublime
x → *arche de Noé* (manifest destiny)
x → enlightened through the darkness, desolated land.

Translation:

arche de Noé = *Noah's Ark*

The next example contains two cases of code-switching. "*Authorial Onomastics*" represents a terminological switch, while the word "voice" below is triggered by the above-mentioned terminological switch.

Example 4:

Authorial Onomastics : *Day* = *daisy*
voice *Day* = jour = lumière

Translation:

jour = day

lumière = light

10. *Parenthetical comment*.

The code-switching employed as a parenthetical comment can be inside parentheses (Examples 1 and 2).

Example 1:

-> plan (*annonce*).

Translation:

annonce = *announcement*

Example 2:

_ 7 rhetorical devices (*tous différents!*) _

Trump = opportunist _

Translation:

tous différents = *all different*

Some code-switching performing the parenthetical comment function could be used without parenthesis at all (Examples 3 and 4).

Example 3:

hanged criminals → *tout un rituel*... ⇒ supposed to be dissuasive

Translation:

tout un rituel = *a ritual*

Example 4:

Gives the impression of an harmonious life, even though times were harsh for farmers. (shoes → *dégués*). Everyday tasks → as if doing their part [etc.].

Translation:

dégués = *icky*

11. *Language economy.*

Example 1 shows a switch in writing the time period, which is shorter to express in English ("60's") than in French ("années 60"). Such a case appears three times in the French corpus.

Example 1:

Negociat° mais vers milieu des 60's → *radicalizat°* du mvt

Translation:

Negociat° (shortening from "négociation") but closer to the middle of the 60's → *radicalizat°* (shortening from "radicalization") of the movement (shortening from "mouvement")

The second example illustrates a switch in the last line where the student prefers to use French to say "*des oublis*" than to say "memory lapses" or "forgotten issues" in English. One French word "oubli" bears the semantic meaning, which is expressed in a more complex structure in English.

Example 2:

I - Critiq Labour

II - Critiq Cons.

III - *Des oublis.*

Translation:

Des oublis. = *memory lapses.*

The third example provides a case of code-switching in the phrase where the student switches into French in the expression "*au quotidien*". This word expression could be translated as "daily" or "every day". However, given the syntaxes of the phrase, such a translation may change the meaning of the whole line. The translation that could fit the present syntaxes is "*on a daily basis*" and it is longer than the original French expression.

Example 3:

→ *segregation au quotidien.*

Translation:

au quotidien = *on a daily basis*

12. Spelling check.

In Example 1 it seems that the student was looking for a correct spelling of the word "intrinsic"; first having written this word in French and then having started to write it in English, but never finishing it.

Example 1:

<i>intrinsèque</i> intri

Translation:

intrinsèque = *intrinsic*

Example 2 illustrates a few words, which are situated randomly on top of the student's draft. Probably, the phrase "*droite parallèles*" was used to find the right word to express this idea in English.

Example 2:

parallel	
	struggling
<i>droite parallèles</i>	

Translation:

*droite parallèles*¹³ = *parallel line*

Example 3:

desillusioned
desillusionned
occurrence
<i>fidele</i>

Translation:

fidèle = *faithful*

¹³ As mentioned earlier, all of the original drafts are transcribed with all of the original spelling and potential grammatical mistakes. Some of the original drafts are available in Appendix 1, and the entire corpus has been made available online at the address https://drive.google.com/drive/folders/0B_C6pnkGJ6jHQmVneThYa0JSMkk?usp=sharing

In the last example the spelling of the first two words is based on their French equivalent "désillusionner" and not the English word "disillusion". The last word lacks the diacritic, which should be over the letter "è" in the correct spelling of the French word "fidèle". Probably the author wanted to adapt the French word to the English spelling not recalling the English equivalent "faithful".

13. *Reiteration.*

The pragmatic function of reiteration is found only once in the French corpus.

Example 1:

attempts to feminize them *feminiser*

Probably the student was looking for an appropriate word in English and to be sure of the meaning, which the student wanted to express, they doubled the word in French.

5.2.2. Summary table

The 13 above-mentioned pragmatic functions are summarized in the table below according to the number of their occurrences in the French corpus. The first six functions are the most recurrent. The largest frequency belongs to the free switchings, which amount to 96. Such functions as triggered switches, stylistic switches, clarification and/or elaboration, framing and structuring fluctuate between 27 and 46. The remaining seven functions do not appear very frequently throughout the corpus; they vary from 1 to 11 occurrences.

Table 3. Pragmatic functions of code-switching in the French drafts.

Pragmatic functions	Code-switching (Nº of occurrences)
Free switching	96
Triggered switches	46
Stylistic switches	42
Clarification and/or elaboration	31

Framing	28
Structuring	27
Direct quotation	11
Lexical need	9
Parenthetical comment	7
Terminological switches	6
Language economy	5
Spelling check	3
Reiteration	1
Total:	312

5.3. Linguistic perspective

Code-switching may occur on different textual levels. The French drafts show language alternation inside a word, inside a sentence, in-between sentences and in-between paragraphs. All these cases will be examined in the three following sections: intra-sentential code-switching, inter-sentential code-switching and word-internal code-switching.

5.3.1 Intra-sentential code-switching.

The most frequent cases of intra-sentential code-switching are one-word switches or switches of a group of words. Examples 1 and 2 illustrate intra-sentential switchings when the alternation happens before or after a colon in the respective coordinate clauses, which nonetheless constitute one utterance, as explained above.

Example 1.

In a first part, ~~the~~ his first election and the New Labour will be addressed. Then, a second part will be dedicated to his reforms concerning education and crimes. Finally, a third part will focus on devolution.

Problématique : to what extent are the education, crime and devolution reforms representative of Blair's era?

devolution

Translation:

Problématique = Problematic

Example 2:

Transit° : *entre les parties*

Translation:

entre les parties = between the sections

A switch may also take place inside a simple sentence represented either by a word or a group of words (Examples 3 and 4).

Example 3:

Black people were considered as *égal* as whites p. but lived in this divided s^t was a ~~hard fight~~ of everyday fight.

Translation:

égal = equal

Example 4:

- *défense de la révolte as a matter of survival*

Translation:

- defence of the revolt *as a matter of survival*

Specific features of drafts are the lack of full grammatically correct sentences and therefore the lack of subordinate clauses. Very often the students use graphic elements such as arrows (\rightarrow , \Rightarrow , \Downarrow , \Leftarrow , etc.) to show dependent relationships between the clauses of the sentence. As a result, there is only one example of code-switching found in a sentence with subordination. Example 5 shows the language alternation in the main clause. The sentence is almost grammatically correct, except for the lack of a subject in the main clause and the lack of a full stop at the end of the subordinate clause.

Example 5:

24 *se rend compte que* the sin is familiar

Translation:

se rend compte que = realizes

5.3.2. Inter-sentential code-switching.

The most frequent case of inter-sentential code-switching is the language alternation between the elements of a list. The switch may be represented by a single word, a group of words or one or a few sentences. The code-switching in listing is shown in Example 1, 2 and 3 below.

Example 1:

sarcasm
satire
comparaison
contradiction
paradox

Translation:

comparaison = comparison

Example 2:

- *Dialogue amoureux.*
- Impossible love
- Greek reference ⇒ associated with his own state.

Translation:

Dialogue amoureux = amorous dialogue

Example 3:

~~both do not have an introduction paragraph, which is interesting to notice, they go straight into the subject.~~

~~—dès les premières lignes on rentre dans le sujet, on apprend les démissions et leurs causes.~~

Translation:

dès les premières lignes on rentre dans le sujet, on apprend les démissions et leurs causes = from the first lines we enter into the subject, we learn about the resignations and their causes.

Code-switching may also occur between the sentences inside one point of a list (Example 4). However, such a combination appears much more rarely than the previous one.

Example 4:

I The Pastoral Genre
A The place & time
B The characters
C. *Legereté du discours*. Flowers music

Translation:

Legereté du discours = lightness (the lack of the diacritics in the original French word "légèreté") of discourse

Besides listing, students often use a new line to separate ideas or to make their subdivision without graphic markers distinctive of listing. Inter-sentential code-switching may occur between new lines, which designate separate utterances. One such example is demonstrated in Example 5.

Example 5:

Pb *Quels sont les arguments?*
Pourquoi: dit-il que / Dans quelle mesure?
How does the journalist justify this gap btw rhetoric and achievement.

Translation:

Problematic (shortening from "problématique") *What are the arguments?*

Why does he say / To what extent?

Another type of inter-sentential code-switching is displayed by the students' language choice for different sections of their writing. A draft may consist of several monolingual paragraphs, each written in a different language (with a few code-switchings inside). Such a rough copy is illustrated in Example 6. Looking the text over, one notices that the introduction is written entirely in French including only the very last phrase in English as if anticipating the future switch (without taking into consideration the sentence in blue, which was probably added later). The next two parts of the student's draft are both in French and English; the first has French as a matrix language and the second has English as its matrix language, both with multiple language alternations. The student's conclusion as their introduction is mostly in French including a few English code-switchings. The draft demonstrates a gradual language switch throughout the three paragraphs from French, then French with multiple code-switchings into English and followed by English with multiple code-switching in French; to come back into French in the final paragraph of the sample.

The sentences in blue seem to be the author's incorporations, which were added while proofreading of the draft and before writing a fair copy, which can simply be determined by changes in the colour of the ink used.

Example 6:

1960's → années très troublées aux EU

montrée d'une forte identité communautaire → chez les blacks, les Natives, les gays (plutôt 70's),...

En ce qui concerne les black : la loi les reconnaît comme des citoyens normaux mais la mentalité des gens n'a pas encore changé : il y a tjs des discriminat° à cause de la couleur de peau. (décalage loi / réalité)

1963 : une année marquante pr la communauté de Noirs : bcp de manifestations violentes. Mais premiers Noirs entrent à l'université (mais doivent ê protégés).

↳ intervention du Psdt ds le privé ⇒ sermonne les américains blancs.

- Paradoxe entre recherche d'une identité américaine collective & revendication de plusieurs identités communautaires.

- ~~Discours~~ Utilisation des medias pr parler aux citoyens → JFK avec son tps.

- que nous apprend JFK sur la société américaine de 1963 ?

What is he trying to do in His text ? **Pq a-t-il besoin de faire un civil right speech en 1963?**

1) Une société divisée → *defining society* (american)

- discriminations : à l'université / ds l'éducation
sur le marché du travail
géographiquement → guettos, quartiers noirs.
- violences
- *are not fully free.* → *not freed from social & economic oppression.*
- remarque les progrès faits (?)

⇒ **paradoxe identité nationale / identité communautaire.**

2) Uniting the society : *rappelle ce que c'est que* « being American »

- all men are equal → *rappel des droits fondamentaux*
- responsibility of all : no discrimination
- paradox : we are committed to a Worldwide struggle to promote and protect the rights of all who wish to be free. (⇒ Cold War)
- a moral issue
- defining americanness.
- we preach freedom around the world ⇒ necessity to be united to be credible on the outside → US need a good image to continue war in Vietnam, etc.
- US : a leading land for rights of MEN & citizens.
- paradox : *obligé de prendre des mesures pr faire respecter les droits fondamentaux.* (nationwide legislation is needed).

~

JKF obligé de rappeler des choses qui devraient être naturelles (mais qui ne le sont pas en réalité). Intervient° ds la sphère privée (essaye de changer la mentalité des gens en leur disant ce qu'ils doivent / devraient faire / penser. ⇒ *Freedom ?*

+ en mettant en place des lois pr les forcer.

↳ paradoxe : *freedom of thoughts* mise de côté, mais pr la bonne cause!

Prend parti des NOIRS

5.3.3. Word-internal switching

The creativity of the French students is represented in a wide range of words, which do not exist in either French or English, but rather show a word-internal code-switching or a lexico-grammatical adaptation of these languages. Ignoring the cases of diacritic omission, letter doubling and the omission or addition of a final "-e", which were mentioned in the methodology (section 4.3.) and are considered as

spelling mistakes, the French samples introduce 11 word-internal code-switches. Each word appears only once in the French data, except for the word "surnatural", which was found in two separate drafts by different authors.

Here is the list of the word-internal switches:

1. plural (plural [Eng] + pluriel [Fr])
2. bruled (bruler [Fr] + Past participle -ed [Eng])
3. favourished (loan translation of "favorisé" [Fr])
4. supperieur (superior [Eng] + supérieur [Fr])
5. surnatural (surnaturel [Fr] + supernatural [Eng])—this word appears twice in two different rough copies.
6. minutious (loan translation of "minutieux" [Fr])
7. inconscious (inconscient [Fr] + unconscious [Eng])
8. économic (économique [Fr] + economic [Eng])
9. theoretical (théorique [Fr] + theoretical [Eng])
10. in filigram (loan translation of "en filigrane")

Only the word "économic" clearly represents a switch inside a lexical item. All the other word-internal code-switches could be equally the ignorance (either conscious or unconscious) of an English equivalent resulting in a spelling mistake and creation of a new lexical item.

5.3.4. Summary table

The table below shows that both intra- and inter-sentential types of code-switching are employed almost equally. Due to the fact that the limit between a word-internal code-switching and ignorance is not explicit, the drafts, whose only switch is within the boundaries of a word were not taken into consideration while counting the drafts with intra-sentential code-switching, which explains that the total given in this table is lower than that in previous discussion.

Table 4. Intra-/inter-sentential code-switching.

	Intra-	Inter- + intra-	Inter-
N ^o of drafts	35	31	33
Total (N ^o of drafts)	99		

5.4. Grading

Students' academic success has always been important for educationalists. Bilingualism is becoming more and more present in education due to bilingual students, bilingual environment, bilingual educational programs, etc. However, the everlasting debate over the relationship between bilingualism and intelligence is still open ended (Hakuta 1986; Bhatia and Ritchie 2006). The present work opens a new direction of the insight into academic personal writing. The collection of examination drafts makes it possible to put to the test the relationship between bilingual drafts and respective marks for the examination. This issue is slightly remote from the main objectives of the present thesis, but it is still of general interest. For this reason, a small survey was conducted to see whether there is a correlation between bilingual personal writing and students' grades. Taking into consideration the previous research attempting to link bilingualism and intelligence (Hakuta 1986; Romaine 1989/1995) my hypothesis is that there is not any such correlation, and that the fact of code-switching in rough copies does not influence the grade or vice versa.

To perform such a survey, the drafts have to contain the students' names, which should be then matched with the respective marks. The collection of non-anonymous drafts necessitates a higher degree of confidence in the researcher. Likewise, the collection of examination marks requires a tight collaboration with professors. The two above-mentioned factors imply a much more limited number of collected data. In the scope of the present work only 38 drafts correspond to these criteria. These rough papers were collected after sociolinguistics examinations. 10 third year students had to give answers to several questions on sociolinguistics. And 28 students in their 4th year of studies were to write a response to the author of the given text.

The total of 38 rough copies contain 12 bilingual samples. The average mark of the students who code-switched in their examination drafts was compared with the average mark of the students whose drafts did not contain code-switching. The French grading system implies a 20-point grading scale. The statistical results were based on t-test calculations for two samples. The results are stated statistically not significant with a probability threshold of $p < 0.05$. Table 4 below provides the averages of the bilingual and monolingual collected drafts.

Table 4. The average marks of the drafts with/without code-switching.

Average of the drafts WITH code-switching	Average of the drafts WITHOUT code-switching
11,141	11,488
The t-value is 0,307907. The p-value is 0,759928. The results are not statistically significant.	

The averages turn to be equal, and thus the results confirm the suggested hypothesis that code-switching, while preparing an examination draft, and grade distribution are not correlated. This survey is only a first sight at the issue of possible relationship between bilingual personal writing and academic performance. Unfortunately, the limited number of the drafts in the present survey allows neither to make a firm conclusion nor to extrapolate it to the general use of code-switching in examination rough copies. For this reason I suggest its further development on a bigger and more various data set.

5.5. Summary and conclusion

The French data is rich in code-switching. It presents 13 pragmatic functions with free switching, triggered switches, stylistic switches, clarification and/or elaboration, structuring and framing being the most frequently used ones. All the three types of switching—intra-/inter-sentential and intra-word code-switching—are illustrated in the French drafts, with almost equal distribution of intra-/inter-sentential types. A closer look at the 6 most frequently employed pragmatic functions from the structural perspective gives support to the equal distribution of intra-/inter-sentential switching. The functions of free and triggered switches admit both intra-/inter-sentential switching; stylistic switches and clarification and/or elaboration suggest mostly intra-sentential code-switching; whereas structuring and framing presuppose mostly inter-sentential switching. The distribution of these six functions also intimates the motivation for code-switching; they illustrate that code-switching is easily used by the French students; the participants switch between languages not to interrupt the process of expression; besides they apply code-switching to give more precise information; and finally code-switching is employed to structure this information and to provide markers for a more functional use of the draft.

Despite a great variety of code-switching functions and types, language alternation goes unnoticed for a reader. Almost a half of the participants visually mark code-switching (45 out of 107); whereas only one rough copy contains all occurrences of code-switching, which have been marked, without any other visual marking in the draft—43 rough copies demonstrate that code-switching is partially or entirely marked in parallel with other content of the draft. This ratio implies that visual marking is used to emphasize particular points of the content of the examination answer rather than draw the reader's attention to a language switch.

Besides, the small survey shows no relationship between code-switching in the rough copies and the marks attributed to the students after correcting their fair copies. The inference, that French data description suggests, is that students employ code-switching both as a methodological tool and an alternative for self-expression. They use all available resources—code-switching being one of them—to prepare the examination draft efficiently with internal and external structure and explicit content which are suitable for them.

Chapter 6.

Russian data

The Russian data comprises 93 drafts, which were collected after final examinations in the English department of Volgograd State University (Russia) during the summer period of 2014–2015. 61 students were enrolled either in Linguistics programme or Philology and applied linguistics programme. These are two 4-year courses. The other 32 students were following a Theoretical and applied linguistics course of 5 years. All the students who participated in the present research had English as their first foreign language.

The examinations consisted either of 4 or 5 parts depending on the programme. All the parts had to be written and then, in addition, some of them were performed orally. Table 1 gives the details of the final examinations of the three above-mentioned programmes.

Table 1. Volgograd State University Examinations' description.

Linguistics (4 years)	Philology and applied linguistics programme (4 years)	Theoretical and applied linguistics (5 years)
1. Translation from English into Russian (written)	1. Annotation of an English text in English (written)	1. Annotation of an English text in English (written)
2. Grammar question in English (written)	2. Reviewing of a Russian text in English (written)	2. Reviewing of a Russian text in English (written)
3. Reviewing of an English text in English (written)	3. Interdisciplinary question in English or Russian (written)	3. A grammar question in English (written and oral)
4. Topic in English (written and oral)	4. Interdisciplinary question in English or Russian (written and	4. Lexicology question in Russian (written)
		5. Topic in English

	oral)	(written and oral)
	5. Topic in English (written and oral)	

The examination prompts were given in Russian, while the content information of each part such as texts, questions and topics were provided in the language in which the task either should have been performed (e.g. annotation, questions and topic) or which the task necessitated for its execution (e.g. translation and reviewing). The Russian data is the only example where the participants had no choice of language for the heading of the examination, i.e. students had to write the title of the examination, their full names and the group title in Russian.

28 out of the total 93 students code-switched in their rough copies, which represents 30.1%. The distribution of bilingual drafts is rather unequal according to the year of studies: the 5th year students employed code-switching almost 2 times as much as the 4th year participants. Table 2 provides information on numbers of collected data and the occurrence of bilingual samples among them.

Table 2. The ratio of Russian/English bilingual drafts to the year of university enrolment.

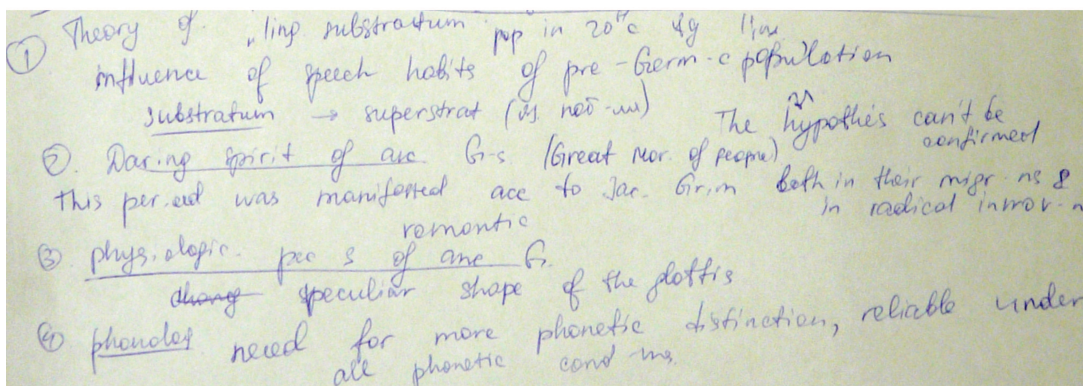
	All collected drafts		Bilingual drafts	
		% (of total)		% (of the respective year)
4 th year	61	65.6%	13	21.3%
5 th year	32	34.4%	15	46.9%
Total	93	100%	28	30.1%

6.1. Visual perspective

The Russian bilingual data presents occurrences of at least two languages—Russian and English—and thus the use of both the Cyrillic and the Latin alphabets. The fact of using the two alphabets raises the expectation of a higher visibility of language alternation; although, it is not always the case. The examples, which are illustrated in Image 1 and 2, show the cases when the language alternation needs an

attentive reading to be noticed. Image 1 presents an excerpt with a single occurrence of code-switching.

Image 1. Extract from the Russian draft.



Despite the extraposition of the code-switched element, it is still difficult to see it. The Russian shortening is inscribed just above the word "hypothe's" on the right of the sample. Its transcription is provided in Example 1 below.

Example 1. Transcription and translation of Image 1.

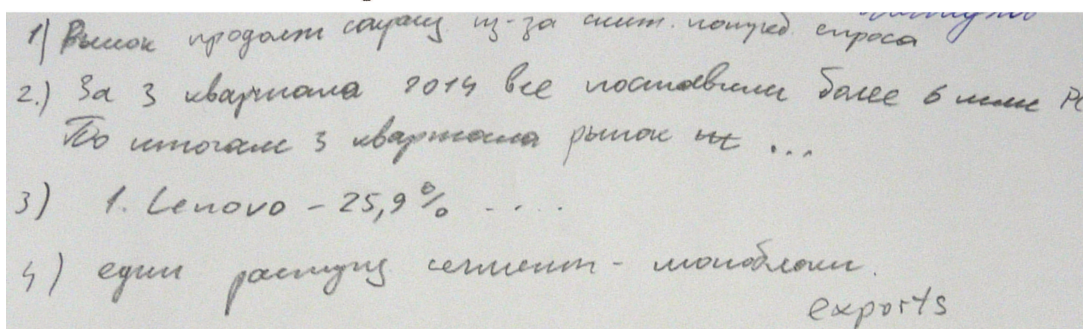
<p>гл</p> <p>The hypothe's can't be confirmed</p>

Translation:

гл = *main* (shortening from "главная/ые")

The excerpt in Image 2 contains two cases of language alternation. One of which is easily noticeable, while the second one requires more attention and understanding of the content to be perceived.

Image 2. Extract from the Russian draft.



The code-switched element, which the reader perceives at once, is the English word "exports" at the end of the extract. However, there is one more occurrence of code-switching at the end of the second line. The abbreviation "PC" could belong both to Russian and English, given that the letters "P" and "C" exist in the Cyrillic and the Latin alphabets. Still, it is not a neutral item; if one reads the whole drafts it becomes clear that by "PC" the student meant "Personal Computer", which is one of the topics of the reviewing task. This word both in its full form and abbreviation is repeated several times in the student's draft. The line 3 is not considered as code-switching, because the word "Lenovo" represents the name of the brand which was given in English in the original text, and which could be often found in its original form in Russian texts. Example 2 provides the transcription and the translation of the above-mentioned excerpt.

Example 2. Transcription and translation of Image 2.

1) Рынок продолж сокращ из-за сниж потреб спроса
2) За 3 квартала 2014 все поставили более 6 млн <i>PC</i>
По итогам 3 квартала рынок ж ...
3) 1. <i>Lenovo</i> - 25,9% ...
4) един растущ сегмент – моноблоки.
<i>exports</i>

Translation:

1) The market continues (shortening from "продолжает/ал") to contract (shortening from "сокращаться") because of a low (shortening from "сниженного") consumer (shortening from "потребительского") demand

2) For the 3 quarters of 2014 all delivered more than 6 ml *PC*

 According to the results of the 3 quarter, the market ...

3) 1. *Lenovo* – 25,9% ...

4) The only (shortening from "единственный") growing (shortening from "растущий") segment – monoblock

exports

Out of 28 bilingual drafts, 12 comprise visually marked code-switching. Table 3 illustrates the number of such drafts according to the amount of marking.

Table 3. Visually marked code-switching.

	All code-switching is visually marked	Some code-switching is visually marked
Monolingual content is also visually marked	5	7
Only code-switched content is marked	0	0
Total (№ of drafts)	5	7

5 samples represent the drafts where all code-switching is marked, while the 7 are the drafts where only some code-switching is marked. There are no samples where marking is an exclusive feature of code-switching; in all 12 drafts visual marking was applied both to the code-switched and monolingual elements. The Russian scripts comprise five types of visual marking, namely the use of parentheses, framing, quotation marks, structural marking and underlining (Example 3).

Example 3. Different types of visually marked code-switching.

1. Parentheses

"Нет никаких сомнений в ~~что~~ том, что ^{сейчас. эконом. ситуац. станов.} экономика —
стабилизируется ^{стабильнее}
по сравнению с прошлым годом (*perhaps*), ~~об~~ этом можно судить по
состоянию финансовых рынков.

2. Framing.

В соответствии с отчетом, опублик net
в сент.
составл, олицетвр
Предст Страны Б7. предст б 66% мирового благосост.

3. Underlining.

основано ⇒ formal | functional | semantic criteria

4. Quotation marks.

вопросы, которые освещаются

В статье рассматр-ся ~~вопросы~~ ~~исчезновения~~ ~~языков~~ в работе Дэвида Кристалла "Language death".

5. Structural marking.

– ultranationalists in the Brit Isles , (use of social network by the gangs/groups to promote ideas

800 знаков.

Eur people are worried about ~~neg~~ increasing imm. rates & are afr.of spreading of islam.

6.2. Pragmatic perspective

6.2.1. Pragmatic functions

Despite the limited number of bilingual samples, the Russian data encompasses a large variety of pragmatic functions of code-switching, 12 in total.

1. Triggered switches.

The triggering word is in bold typeface (added by the author of the present work for an easier visual representation).

In Example 1 the name of a foreign company, which is written with the Cyrillic alphabet, triggers the switch into English. However, after the word "has" the student immediately goes back into Russian.

Example 1:

~~Рынок пр. расп. реч. находится в состоянии стагнации в теч. посл. 10 лет. Недавно~~

– Softw Spee

??? компан **Real Спикер** has заявила о том что собир. разраб.

Translation:

The market of speech (shortening from "речи") recognition (shortening from "распознания") programmes (shortening from "программ") has been stagnating

for the past 10 years. Recently

– ~~Softw Spee~~

??? company (shortening from "компания") **Real Speaker** has announced what they're going to develop (shortening from "разрабатывать").

In the next example the student illustrates the "means that allow to refer to an object without mentioning it". While the first illustration is given in Russian, the second one is in English, and the latter triggers the end of the sentence in English. The code-switching "e.g." is of the language economy type, which will be introduced later.

Example 2:

• средства, позволяющие, не называя предмет, отсылать к нему, e.g. что угодно, **everything**, etc.

Translation:

• means that allow to refer to an object without mentioning it e.g. anything, **everything**, etc.

Example 3 illustrates a vocabulary note with a parenthetical comment. The first word of the latter is triggered by the translation of the word "request".

Example 3:

request - **запрос** (вместо answer)

Translation:

request - *demand* (instead of answer)

2. Terminological switches.

This pragmatic function is found three times in the Russian data. In the first two cases, the students employed the term "Translation memory" in English, while giving the whole answer in Russian. This term denotes a database, which has its

official Russian translation "Память переводов". However, the two students preferred to use the English term, which was probably used in their courses during the semester.

Example 1:

Автоматизир перевод - перевод текстов на ЭВМ с пом. комп.терхнологий. От машинного он отличается тем, что компьютер лишь позволяет ему произвести готовый текст - ~~либо или~~ - работа со словарями, глоссариями, *transl. memory*.

Translation:

Computer-aided (shortening from "автоматизированный") translation - translation of texts on the computer with the help (shortening from "помощью") of computer (shortening from "компьютерных") software. It differs from the machine translation in a way that a computer only allows to produce a finished text - ~~or~~ work with dictionaries, glossaries, *transl.* (shortening from "translation") *memory*.

Example 2:

частично-автоматизир. перевод комп ср-ва в помощь переводчику
М-словари, тезаурусы, корпуса текстов + технология *Translation Memory*.

Translation:

a partially automatic (shortening from "автоматизированный") computer (shortening from "компьютерного средства") translation helps a translator
M-dictionaries, thesauruses, text corpora + technology of *Translation Memory*.

In the third example the student uses the Russian term "Младограмматик" instead of its English equivalent "Neogrammarian".

Example 3:

a-b) 1) Phonetic Theory of simplification of comparatives in 19c.
Младограмматик - Yong Grammarian School, Hernou Paul, Karl Brugmenn, University of Leipzig.

Translation:

Младограмматик = *Neogrammarian*

3. Reiteration.

The Russian scripts demonstrate only the cases of terminological reiteration, i.e. the repetition of specific terms or fixed expressions.

Example 1:

Словарн статья (-*entry*) – основн структурн ед-ца словаря. [...].

Translation:

Diction. (shortening from "словарная") entry (-*entry*) – a basic (shortening from "основная") structural (shortening from "структурная") unity (shortening from "единица") in a dictionary. [...].

Example 2:

В понятие “корпус текстов” входит также система управления текстовыми или лингвистическими данными, которую называют корпусным менеджером (или корпус-менеджером; англ. *corpus manager*).

Translation:

In the notion of “text corpus” also includes a managerial system of texts or linguistic information, which is called a corpus manager (or corpus-manager; in English (shortening from "по-английски") *corpus manager*).

Example 3:

Виды: 1). Морфологич. - *part of speech tagging* - частеречная разметка, включ. не только признак части речи, но и призн. грамматич. категорий, свойств. данной части речи.

Translation:

Types: 1). Morphological (shortening from "морфологическая") - *part of speech tagging* - part of speech marking includes (shortening from "включает") not only an indication of part of speech, but also an indication (shortening from "признак") of grammatical (shortening from "грамматических") category, characteristics of a given part of speech.

Example 4:

Поисковые системы ~~не~~ связаны с поисковой оптимизацией. (*search engine optimization SEO*) - комплекс мер для познания позиций сайта в результатах выдачи поис. сис-м. по определенным запросам пользов.

Translation:

Search engines are related to search optimization. (*search engine optimization SEO*) - a complex of measures to know the position of a web site in the results given by search (shortening from "поисковых") engines (shortening from "систем") according to specific user's (shortening from "пользователей") inquiries.

Example 5:

The stress can be shifting (*подвижн.*)

Translation:

The stress can be shifting (*shifting* (shortening from "подвижный"))

The last example shows the reiteration of the adjective "shifting" ("подвижный"), which is part of a fixed expression "shifting stress"/"подвижное ударение".

4. *Free switching.*

Example 1:

Задание №1

- Название
- Аннотация
- *Key words*

Translation:

Task №1

- Title
- Annotation
- *Key words*

Example 2:

- 1) Рынок продолж сокращ из-за сниж потреб спроса
- 2) За 3 квартала 2014 все поставили более 6 млн *PC*
По итогам 3 квартала рынок ж ...
- 3) 1. Lenovo - 25,9% ...
- 4) един растущ сегмент – моноблоки.
- 5) *exports*

Translation:

- 1) The market continues (shortening from "продолжает") to contract (shortening from "сокращаться") because of a low (shortening from "снижения") consumer (shortening from "потребительского") demand
- 2).For the 3 quarters of 2014 all delivered more than 6 ml *PC*
According to the results of the 3 quarter, the market ...
- 3). 1. Lenovo – 25,9% ...
- 4). The only (shortening from "единственный") grow (shortening from "растущий") segment – monoblock
- 5). *exports*

Example 3:

According HP на HP manager *Николай И* to experts, it happens because of

Translation:

Николай И = *Nikolay I* (the name of the manager)

Example 4:

According to the place of the word stress it may be free or fixed

(eng. rus)

(латын, фран.)

Translation:

латын = *Latin* (shortening from "латинский")

фран = *French* (shortening from "французский")

5. Language economy.

As the previous examples of Russian scripts show, the Russian students shorten Russian words a lot. Besides, they frequently employ foreign abbreviations or even words, which are much shorter than their Russian equivalents.

In Example 1 the English abbreviation "vs." is used in the middle of a Russian phrase. Russian has an equivalent for "versus"—"в сравнении с"—however, there is no equivalent abbreviation for "vs."

Example 1:

Корпус Vs. Эл. Библи.

Translation:

Corpus Vs. Electronic (shortening from "электронная") Library (shortening from "библиотека")

The next example shows the use of an English adjective phrase with the purpose of language economy. The phrase "most popular" is much shorter than its Russian equivalent "наиболее широко использованный".

Example 2:

Развитие лингв. кор. • *most popular is* Британский национ. корпус, [...].

Translation:

The development of the linguistic (shortening from "лингвистический") corpus (shortening from "корпус") • *most popular is* the British national (shortening from "национальный") corpus, [...].

Example 3 illustrates the use of the English words such as "so" and "use" for language economy in a Russian text.

Example 3:

Корп. нового поколения ≈ 100кк слов, *so* выдвигаются принципы разработки систем, которые бы минимизировали вмешат. человека. Автомат. разрешение ~~мор~~ морфолог. или синтаксич. неоднозначн., как правило, основывается на ~~use~~ *use* информац. более высокого явл. (синтаксич., семантич.).

Translation:

Corpus of a new age \approx 100kk words, so the principles of the system development, which could minimize the human interference (shortening from "вмешательство"), are put forward. ??? authorization of ~~мор~~ morphological (shortening from "морфологически") or syntactic (shortening from "синтаксически") is dubious (shortening from "неоднозначно"), as a rule, it is based on the ~~use~~ use of information (shortening from "информации") of a higher event (shortening from "явления") (syntactic (shortening from "синтаксического"), semantic (shortening from "семантического")).

The existing Russian abbreviation for "e.g." is "напр." (shortening from "например"). The student prefers to use the former in their writing (Example 4).

Example 4:

Темат. ориентация

- монотематические (e.g: перевод ~~новое~~ прогноза погоды)
- политематические

Translation:

Thematic (shortening from "тематическая") orientation

- monothematic (e.g: translation of a ~~new~~-weather forecast)
- polithematic

The following example shows the use of such English verbs as "consist of" and "deals" instead of "состоять из" and "имеет дело", respectively.

Example 5:

Цеп. могут *consist of* из яз. элем-ов разн. уровней: словоформы - цепочки морфем; словосоч и предл - цепочки слов. \Rightarrow Ф. гр *deals* с абстракт. объектами.

Translation:

Chains may *consist of* of language (shortening from "языковых") elements (shortening from "элементов") of different (shortening from "разных") levels: word forms - chains of morphemes; word combinations (shortening from "словосочетания") and sentences (shortening from "предложения") - chains of words. \Rightarrow F. gr. *deals* with abstract objects.

The only case of Russian–French code-switching is presented in Example 6. Instead of the Russian noun phrase "прогноз погоды", the student uses the shorter French noun "Meteo" (without diacritics over the letters -e. The correct spelling is "météo").

Example 6:

- монотематические (настроенные на одно предметную область (Meteo/Канада) - метосводки с англ на фр.

Translation:

- monothematic (aiming at one subject area (*Weather forecast/Canada*) - weather forecast from English (shortening from "английского") into French (shortening from "французский").

6. Parenthetical comment.

Example 1:

Annotation.

Key words: (*в конце*)

The article is focused on [...].

Translation:

в конце = *at the end*

Example 2:

"Нет никаких сомнений в ~~что~~ том, что ^{сейчас. эконом. ситуац. станов.} экономика
— ~~стабилизируется~~ ^{стабильнее}
по сравнению с прошлым годом (*perhaps*), ~~об~~ этом можно судить по
состоянию финансовых рынков.

Translation:

"There are no any doubts ~~that~~ ^{now the economic (shortening from "экономическая")}
situation (shortening from "ситуация") is getting (shortening from "становится") ~~the economy is~~
stabilizing

more stable

in comparison with the last year (*perhaps*), it can be stated by financial markets.

Example 3:

В соответствии с отчетом, опубликованным
в сентябре,
составляющие, олицетворяющие net
Представители Страны Б7. представляют более 66% мирового благосостояния.

Translation:

According (shortening from "соответствии") to the report published
in September (shortening from "сентябре").

form (shortening from "составляют/составили"), embody
(shortening from "олицетворяют/олицетворили")

Represent (shortening either from "Представители" or
"Представляют/или") Countries of G7. represent (shortening from
"представляют/представили") more than (shortening from
"более/больше") 66% of world's welfare

net

Example 4:

Черновик
Equally illegible:
1) ФГАОУ ВПО
(Волг. Гос. Ун.)
[...]

Translation:

Draft

Equally illegible:

1). FGAOU VPO

(Volgograd (shortening from "Волгоградский") State
(shortening from "Государственный") University
(shortening from "Университет"))

[...]

7. Lexical need.

Example 1:

In my opinion, the governments of the EU countries should devote ^{more find to} this question ~~more time~~ as consequences ~~of~~ can be *непоправимые*. A lot of people ~~inculcate~~^{inculcate} morals value to citizens values.

Translation:

непоправимые = irretrievable

Example 2:

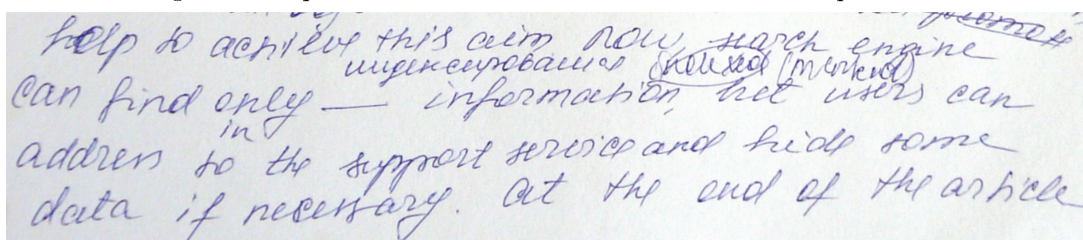
Now search engine can find only ^{*индексированная*} **indexed** (marked) information, net users can address toⁱⁿ the support service and hide some data if necessary.

Translation:

индексированная = indexed

The second example could seem rather ambiguous because the student reiterates the word "indexed" in both languages and even gives a synonym in English. However, I esteem that this example represents a lexical need code-switching. The Russian word "индексированная" appears first, only after it the student proposes its English equivalent. Probably, the student added them some time after. I suggest to have a look at the original draft (Image 1):

Image 1. Excerpt of a Russian draft. Illustration of the Example 1 above.



help to achieve this aim now search engine ^{индексированная (indexed) (marked)}
can find only — information, net users can
address toⁱⁿ the support service and hide some
data if necessary. At the end of the article

The next example is also dubious. It may first appear to belong to the clarification and/or elaboration function. However, the attempt to remember the proper word in English finishes in a misspelling: 'tribe-house effect' instead of "greenhouse effect". The latter makes me consider this example to represent the lexical need (pragmatic) function.

Example 3:

- global warming/climate change → floods, hurricanes, ^{severe} draughts increasing level of carbonhydrates emission → парниковый эффект
tribe-house effect

Translation:

парниковый эффект = greenhouse effect

8. Vocabulary list.

Example 1:

pct - процент

confidence вера

recession - спад

prompted - побуждать

impact воздействие

Example 2:

2. Главным вопросом ^{обсуждения} на предстоящей встрече, которая будет проходить в Лондоне, станут банковские реформы.

emergency rescue - критич. положение спасения

Translation:

2. The main issue ^{of discussion} on the forthcoming meeting, which will take place in London, will be banking reforms.

emergency rescue - critical situation of rescue

The next example of code-switching is situated in the middle of the text.

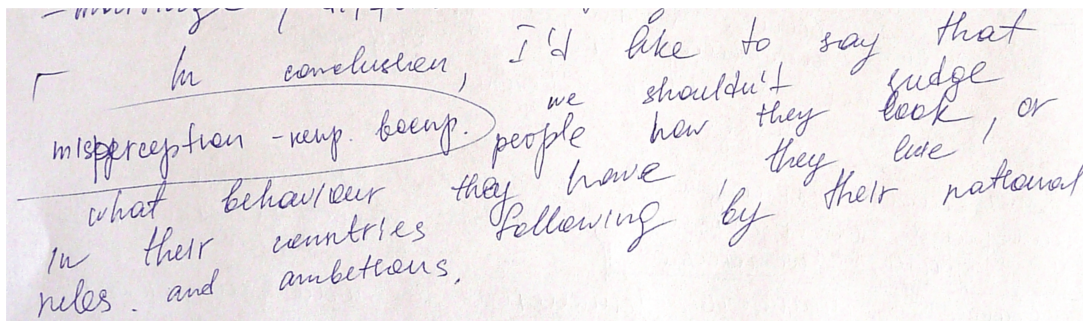
Example 3:

Misperception - непр. воспр.

Example 3 illustrates a piece of vocabulary that was circled to be more visible on the page. The translation for the English word "misperception" is given

with the Russian shortened words: 'непр. воспр.' (*неправильное восприятие* = *false perception*). It probably serves as a reminder. I suggest to have a look at the original script (Image 2):

Image 2. Excerpt of a Russian draft. Illustration of the Example 3 above.



The next example has already been demonstrated in the pragmatic function of triggered switches, where the word "вместо" was discussed. Now, I would like to draw your attention that the example in itself represents a piece of vocabulary list: "request - *запрос*".

Example 4:

request - *запрос* (вместо answer)

Translation:

request - *demand* (instead of answer)

9. Clarification and/or elaboration.

Example 1:

Concept → *осн. из пред.*

Categories. *набор зн.* ⇒ *группы*

Translation:

Concept → *base* (shortening from "основа") of *the object* (shortening from "предмета")

Categories. *a set of signs* (shortening from "знаков") ⇒ *groups*

Example 2:

substratum → superstrat (яз. поб-ля)

Translation:

substratum → superstrat (*language* (shortening from "язык") *of the conqueror* (shortening from "победителя"))

Example 3:

гл
The hypothe's can't be confirmed

Translation:

гл = *main* (shortening from главная)

Example 4:

The second type is usually called "logical necessity" and it is based on the process of deduction (*от общего лица к частному*).

Translation:

The second type is usually called "logical necessity" and it is based on the process of deduction (*from general to particular*).

10. Stylistic switches.

Example 1 shows a stylistic switch into Russian in the last line of the excerpt. The word "заблуждаться" can be translated as "to be mistaken", however, it implies a higher degree of politeness and a lesser degree of directness than its English equivalent, which makes this Russian word a stylistic choice of the student. What follows it represents a triggered switch.

Example 1:

The author gives an example of such extremists activities as the murder of 10 immigrants in Germany. ~~The~~ In Britain the popularity of ultranationalists grow. The experts of ~~the center~~ "Demos" ~~verified~~ ^{figure out by means of social networks.} that many young europeans ~~The author refers to the scientist Jany Bartlett~~ ^{aren't} ~~don't~~ satisfied with the growth of immigrations and ^{are} afraid of ^{islam} spreading ~~of~~ islam The author pays special attention to ^{the} fact that ~~the~~ majority of people *заблуждаться по поводу в отношении мусульман.*

Translation:

заблуждаться по поводу в отношении мусульман = *to be mistaken concerning* (shortening from "по поводу") *with regard to Muslims*.

One case of code-switching, which appears in the next example, has already been discussed in the triggered switches. However, it was an excerpt taken from a bigger abstract; it was the code-switching inside another code-switching (insertion of the word "has"). The example below demonstrates a stylistic function with the help of which the student reformulates in Russian the idea, which they find difficult to express in English.

Example 2:

The market of speech recognition software has been for almost 10 years
The stagnation of speech recognition software market has
*Рынок пр. расп. реч. находится в состоянии стагнации
в теч. посл. 10 лет. Недавно*
– Softw Spee
??? компан Real Спикер has заявила о том что собир. разраб.
The market of speech recognition software has been in stagnation^{decline} for the
past 10 years.

Translation:

The market of speech (shortening from "речи") recognition (shortening from "распознания") programmes (shortening from "программ") has been stagnating

for the past 10 years. Recently

– Softw Spee

*??? company (shortening from "компания") **Real Speaker** has announced what they're going to develop (shortening from "разрабатывать").*

Example 3:

Уже в ранних раб-тах Г. стремился определить общие *tasks of l-cs* в тесной связи с филосо-ей.

Translation:

Already in the early works (shortening from "работах") Humboldt (shortening from "Гумбольд") tended to define general *tasks of l-cs*

(linguistics) in close relation with philosophy (shortening from "философией").

The English word "tasks" in Example 3 is more explicit than the Russian word "цели" or "задачи". In addition to the stylistic use, the word "tasks" triggers the end of the noun phrase to be in English.

Example 4:

space → *космос* + place

Translation:

space → *cosmos* + place

11. *Framing.*

The only case of framing demonstrates the student's remark about the word limit of the task.

Example 1:

– ultranationalists in the Brit Isles , (use of social network by the gangs/groups to promote ideas

800 знаков.

Eur people are worried about neg increasing imm. rates & are afr. of spreading of islam.

Translation:

800 *знаков* = 800 *signs*

12. *Direct quotation.*

Example 1:

вопросы, которые освещаются

В статье рассматр-ся ~~вопросы~~ ~~исчезновения~~ ~~языков~~ в работе Дэвида Кристалла "*Language death*".


Translation:

issues, which are illustrated

The article deals with ~~issues of language disappearing~~ in the David Crystal's work "*Language death*".

In the next example the student quotes (without using quotation marks) an exact word expression from the original text, while searching for the best translation.

Example 2:

<p>12. Рост крупных  <i>emerging markets heavyweights</i> - В связи с ростом и становлением таких крупных рынков как [...]</p>
--

Translation:

~~The growth of big~~

emerging markets heavyweights -

Because of the growth and establishing of such big markets as [...]

6.2.2. Summary table

The occurrences of code-switching in pragmatic functions are distributed rather unequally. For instance, the function of framing has only 1 occurrence of code-switching, while language economy function is represented in 9 occurrences. Table 4 provides the summary of the pragmatic functions of code-switching encountered in the Russian data according to its frequency.

Table 4. Pragmatic functions of code-switching in the Russian drafts.

Pragmatic functions	Code-switching (№ of occurrences)
Language economy	9
Clarification and/or elaboration	9
Free switching	7
Reiteration	6
Vocabulary list	6
Lexical need	6

Triggered switches	5
Stylistic switches	5
Parenthetical comment	4
Terminological switches	3
Direct quotation	2
Framing	1
Total:	63

Framing, direct quotation and terminological switches are the least represented pragmatic functions of the Russian samples. All the other functions range from 4 to 9 occurrences. The total number of occurrences of pragmatic functions in the Russian data goes up to 63 occurrences.

6.3. Linguistic perspective

This section will examine code-switching on intra- and inter-sentential levels. The Russian data do not contain switchings between two parts of either complex or compound sentences. In the Russian drafts switchings occur inside a word, inside a simple utterance, between simple utterances or sentences and before/after a colon.

6.3.1. Intra-sentential code-switching

In the intra-sentential code-switching a switch may occur inside an utterance, surrounded by the matrix language on both sides. It may be a one-word switch (Example 2, the second case of code-switching) or a switch consisting of a group of words or abbreviations (Example 1 and 2).

Example 1:

- катафора - отсылка к объекту, о котором пойдёт речь далее, *e.g.* А вот и он, наш знаменитый чемпион, - Джон Смит.



Translation:

- cataphora - a reference to the object which will be mentioned later, *e.g.* Here he is, our famous champion, - John Smith

Example 2:

Цеп. могут *consist of* из яз. элем-ов разн. уровней: словоформы - цепочки морфем; словосоч и предл - цепочки слов. ⇒ Ф. гр *deals* с абстракт. объектами.

Translation:

Chains may *consist of* of language (shortening from "языковых") elements (shortening from "элементов") of different (shortening from "разных") levels: word forms - chains of morphemes; word combinations (shortening from "словосочетания") and sentences (shortening from "предложения") - chains of words. ⇒ F. gr. *deals* with abstract objects.

A switch may also be situated at the end of the utterance (Example 3). An utterance may start with a code-switching as well (Example 4).

Example 3:

The author pays special attention to ^{the} fact that ~~the~~ majority of people *заблуждаться по-пово в отношении мусульман.*

Translation:

заблуждаться по-пово в отношении мусульман = *to be mistaken concerning* (shortening from "по поводу") *with regard to Muslims.*

Example 4:

Поисковые системы ~~нез~~ связаны с поисковой оптимизацией. (*search engine optimization SEO*) - комплекс мер для познания позиций сайта в результатах выдачи поис. сис-м. по определенным запросам пользов.

Translation:

Search engines are related to search optimization. (*search engine optimization SEO*) - a complex of measures to know the position of a web site in the results given by search (shortening from "поисковых") engines (shortening from "систем") according to specific user's (shortening from "пользователей") inquiries.

An intra-sentential code-switching also occurs in a definition or a translation; it is a switch always before a hyphen or an en dash—no cases of code-switching,

which follow a hyphen or an en dash, were found (Example 5 and 6). Due to a slight visual difference between a hyphen and an en dash in handwritten samples, only hyphens are used in the transcriptions.

Example 5:

pct - процент

Example 6:

a-b) 1) Phonetic Theory of simplification of comparatives in 19c.

Младограмматик - Yong Grammarian School, Hernou Paul, Karl Brugmenn, University of Leipzig.

Translation:

Младограмматик = *Neogrammarian*

Intra-sentential code-switching is also found before and after a colon (Example 7 and 8).

Example 7:

- монотематические (e.g: перевод ~~нового~~ прогноза погоды)

Translation:

- monothematic (e.g: translation of a ~~new~~-weather forecast)

Example 8:

Annotation.

Key words: (*в конце*)

The article is focused on [...].

Translation:

в конце = *at the end*

6.3.2. Inter-sentential code-switching

The Russian samples provide evidence for inter-sentential code-switching happening before a colon; the utterance, which follows the colon is written on a new line (Example 1). The switch occurs only in one direction (a switch before a colon); no reverse order was observed (a switch after a colon).

Example 1:

Черновик
Equally illegible:
1) ФГАОУ ВПО
(Волг. Гос. Ун.)
[...]

Translation:

Draft

Equally illegible:

1). FGAOU VPO

(Volgograd (shortening from "Волгоградский") State
(shortening from "Государственный") University
(shortening from "Университет"))

[...]

Inter-sentential code-switching also occurs between neighbouring utterances, which are not separated by any graphic or punctuation marks (Example 2).

Example 2:

The market of speech recognition software has been for almost 10 years
The stagnation of speech recognition software market has
*Рынок пр. расп. реч. находится в состоянии стагнации
в теч. посл. 10 лет. Недавно*
– Softw Spee
??? компан Реал Спикер has заявила о том что собир. разраб.

The market of speech recognition software has been in stagnation^{decline} for the past 10 years.

Translation:

The market of speech (shortening from "речи") recognition (shortening from "распознания") programmes (shortening from "программ") has been stagnating

for the past 10 years. Recently

– ~~Softw Spee~~

??? *company* (shortening from "компания") **Real Speaker** has announced what they're going to develop (shortening from "разрабатывать").

Intra-sentential code-switching can be also found in lists with either numerical or graphic marking (Example 3).

Example 3:

Задание №1

- Название
- Аннотация
- *Key words*

Translation:

Task №1

- Title
- Annotation
- *Key words*

6.3.3. Word-internal code-switching

The Russian data contain only one word-internal switch, which is presented in the following example: ~~neo-нацисткой~~ group. It is not a typical example of a word-internal code-switching, because here the student switches inside a lexical item, but then separates it into two using a different language for each of its parts. The word "neoNazi" can be written either as a single word or with a hyphen;

however, the student divides it into two words writing the first one in English and the second one in Russian using the Latin and the Cyrillic alphabet, respectively. Probably it is the lack of L2 vocabulary and the student simply did not know or was not sure about its English equivalent.

6.3.4. Summary table

The summarising table (Table 5) shows a clearly unequal distribution of intra-/inter-sentential code-switching. Intra-sentential switches were employed in 20 drafts, while only 1 draft contained inter-sentential code-switching. Even regarding that both intra- and inter-sentential switches were used in 7 rough copies, the distribution remains firmly unequal.

Table 5. Intra-/inter-sentential code-switching.

	Intra-	Inter- + intra-	Inter-
№ of drafts	20	7	1
Total (№ of drafts)	28		

6.4. Summary and conclusion

The description of the Russian data showed that regardless of the limited number of collected drafts, the Russian data contain a significant variety of pragmatic functions (there is a total of 12). Their distribution is gradual; the occurrence of the functions is distributed from 9 to 1. However, from the linguistic perspective, the preference is attributed to the intra-sentential code-switching. The language alternation in the Russian samples is more of an insertional type; usually one word or a group of word is inserted into an utterance with a matrix language. It is supported by the fact that there are no occurrences of code-switching between a main and a subordinate clause, and that the majority of switches is of intra-sentential type.

Eight pragmatic functions, whose frequency varies from 9 to 5, are language economy, clarification and/or elaboration, free switching, reiteration, vocabulary list, lexical need, triggered and stylistic switches. Among them, only the free and

triggered switching function include both intra-/inter-sentential switchings, while the remaining six functions mostly imply intra-sentential switches. This fact reflects the excessive amount of intra-sentential code-switching. At the same time, these six pragmatic functions demonstrate that code-switching constitutes a tool for an efficient composition of a rough copy; they serve to save time and space, and provide a higher degree of accuracy of information.

The visual perspective also shows that code-switching is not marked as it is, but rather constitutes a part of the visually marked content. The Russian students draw their attention to the important elements of their examination answers either to remind them of the key aspects of the answer or to remind them of some doubtful issues, which should be reviewed before/while composing a fair copy. The drafts with visually marked code-switching constitute almost a half of all the bilingual rough copies (12 out of 28). The distribution of all or some code-switching (5/7) visually marked in the samples along with other content of the draft being visually marked, suggests that code-switching comprises a tool for the composition of examination drafts; namely an economical tool applied for the three dimensions: time, space and quality.

Chapter 7.

Bulgarian data

The Bulgarian data is the least homogeneous of all. Half of the data is drafts and another half is fair copies. All the data was collected in 2014–2015 by a professor of Sofia University St. Kliment Ohridski, to whom I am very grateful for making this data collection possible. Then the scripts were sent by postal service.

All the examinations after which the scripts were gathered were carried out in the English department of Sofia University St. Kliment Ohridski. The fair copies were collected after a 4th year written examination in American literature (41 scripts) and after a 1st year written linguistics examination (19 scripts). The literature examination consisted of two parts. In the first one students had to develop in English one of two suggested topics. In the second part they had to discuss a small extract of a literary text according to the following themes:

Identify:

- a) the work
- b) the author

Discuss:

- a) the narrator's problem
- b) the solution offered in the text
- c) the possibilities late twentieth-century literary theory suggests for a solution

The examination questions and given materials were unilingual. All the tasks, questions and literary extracts were written in English and required students to answer in English.

The linguistics examination consisted of explicitly developing a suggested linguistic topic in a form of an essay. Both the examination task and the topic were stipulated in English. Students' answers were also expected in English.

The drafts were collected after a 1st year oral linguistics examination (16 scripts) and the written entrance examination for the Bachelor degree (59 scripts). The linguistics examination represented a linguistic topic, which had to be developed orally. During the time of preparation students could use drafts to

structure their answers. Notably these drafts were collected for the present study. 59 other drafts come after the entrance examination, which required writing an essay on a general topic. The language of the examinations was English for both examination tasks and students' answers.

All the 1st and the 4th year examinations were preceded by courses read in English. Surprisingly, language alternation was present both in drafts and fair copies. However, the students code-switched twice as much in rough copies as in fair copies. Table 1 illustrates the frequency of language alternation in the Bulgarian scripts.

Table 1. The ratio of Bulgarian/English bilingual scripts to the year of university enrolment.

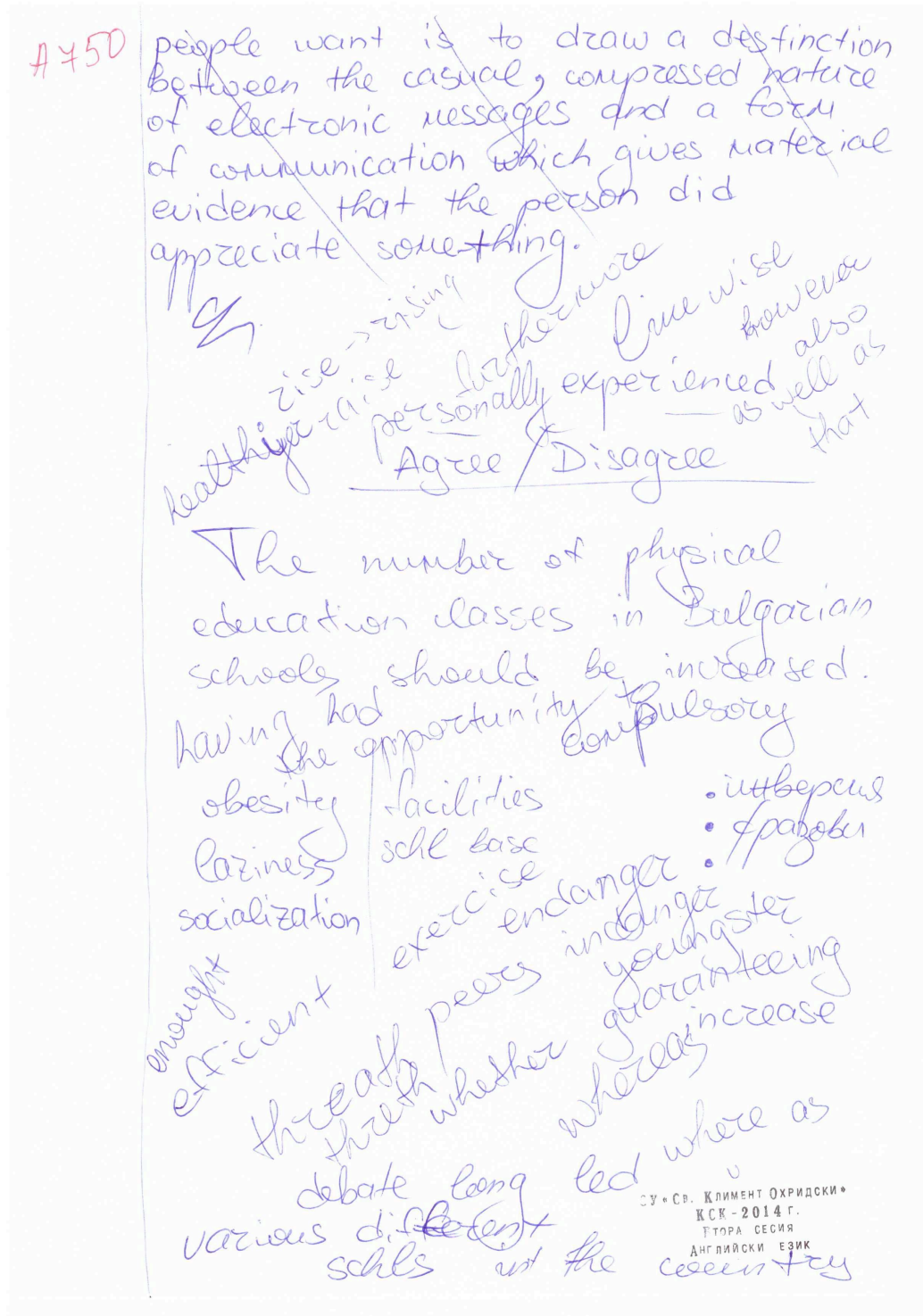
	All collected fair copies		Bilingual fair copies	
		% (of total)		% (of the respective year)
1 st year	19	31.7%	1	5.3%
4 th year	41	68.3%	11	26.8%
Total	60	100%	12	20%
	All collected drafts		Bilingual drafts	
		% (of total)		% (of the 1 st year)
1 st year	75	100%	32	42.7%

Given that the number of collected scripts is rather limited and there is a high diversity among them due to two existing variables (the year of study and the form of scripts, e.g. drafts and fair copies), it would be difficult to make any comparison inside this language group as well as extrapolate the results to the Bulgarian academic environment. However, the collected Bulgarian data merits to be analyzed and its conclusions are worth mentioning.

7.1. Visual perspective

Like the Russian bilingual data, the Bulgarian samples comprise the use of both the Cyrillic and the Latin alphabets. Nevertheless, code-switching is not apparent. Image 1 below shows an example of a Bulgarian draft.

Image 1. Bulgarian draft.



The reader, first, perceives the structural divisions of the script, e.g. the struck out paragraph on top of the page, then the underlined title of the second part and a small listing in the middle-right of the page. Only at second sight, one notices that the listing is given in Bulgarian. Example 1 presents its transcription and translation.

Example 1:

- *инверсия*
- *фразови*
-

Translation:

- inversion
- phrasal
-

Among the bilingual data just a few samples contain visually marked code-switching. To start with the fair copies, none of them show cases of visually marked language alternation. As for the drafts, 8 out 32 bilingual rough copies encompass the occurrences of visually marked switching. Table 2 gives the numbers of bilingual scripts with visually marked code-switching according to the amount of marking.

Table 2. Visually marked code-switching.

	All code-switching is visually marked		Some code-switching is visually marked	
	Fair copies	Drafts	Fair copies	Drafts
Monolingual content is also visually marked	0	3	0	1
Only code-switched content is marked	0	3	0	1
Total (№ of scripts)	0	6	0	2

The rare drafts with marked language alternation represent in their majority marking of all the code-switched elements, either with or without emphasising other content of the scripts. The visual marking encountered in the Bulgarian data constitutes three types: the use of the bold handwriting, underlining and capitalisation. They are illustrated in Example 2.

Example 2. Different types of visually marked code-switching.

1. *Bold handwriting.*

Чернова

2. *Underlining.*

чернова

1. увод = теза

2. selfies show ~~moment~~ feelings at certain moment and...

3. Noone is ever ugly...

3. *Capitalisation.*

ЧЕРНОВА

7.2. Pragmatic perspective

7.2.1. Pragmatic functions

All the cases of language alternation encountered in the fair copies represent one pragmatic function, that of framing. While the whole copy, namely the content information of examination was written in English, the administrative information about the students was given in Bulgarian. 6 students wrote their names and the word "group" ("група") or its shortened form "gr" ("гр") in Bulgarian using the Cyrillic alphabet. If the student mentioned a complete number/title of their student card it was written with Cyrillic alphabet too:

гр. 1, ФН: 12345 (gr. 1, FN: 12345)

2.

Чън X--	12345	Group 3
---------	-------	---------

3.

Виктор	X--	12345	Group 3
--------	-----	-------	---------

4.

Мария X--, gr.3; FN:12345

5.

Тодор Y-- X--
FN 12345 Group 1


6.

Петр X--
F.N. 12345
Group 2

The variety of pragmatic functions slightly increases in the Bulgarian drafts. The code-switching represents the following five functions: free switching, framing, lexical need, clarification and/or elaboration and structuring.

1. *Free switching.*

Example 1:

11:42		
		уседнял застоял
1. Intro 64	rising	In addition
	rise	Furthermore
	raise	постижения
2. For 121		121
		+ 83
3. Against 100 83		+ 64
4. Conclusion 60		_____
		268
		+78

		346

Translation:

уседнял = settled

застоял = stagnant

постижения = achievements

Example 2:

<p><i>начин то у да се спре болеста</i></p> <p><i>начин на живот</i></p> <p>The activists way of life is good me-maner to effects it involved on our physical skill</p> <p>vara variou</p> <p>varaety</p> <p>our variety</p> <p><i>все повече</i></p> <p>education</p> <p>suffer of different</p> <p>illness</p>

Translation:

начин то у да се спре болеста = in a way it can stop a disease (in the original there is a lack of the double "т" in the word "болеста" instead of "болестта")

начин на живот = lifestyle

все повече = increasingly

Example 3:

<ul style="list-style-type: none">• <i>инверсия</i>• <i>фразови</i>•
--

Translation:

инверсия = inversion,

Example 3:

11		- a selfie means	
(10) <u>и 15 мин</u>	10 и 25 мин		"no one gets left behind"
	wanger	as well as	"Y won't get mad
1.	Monalisa		if you put that picture on the
11:20		Internet as appreciate	no one
			based on the principle
			that no one's perfect in a selfie
			<u>consisting of more than one person"</u>

Translation:

(10) и 15 мин = (10) and 15 min = 10.15 a.m.,

10 и 25 мин = 10 and 25 min = 10.25 a.m.

Example 4 with minor modifications such as different visual marking, text alignment and the use of the lower case for the first letter appears 24 times in the Bulgarian data.

Example 4:

<i>Чернова</i>

Translation:

Чернова = *Draft*

3. *Lexical need.*

Example 1:

and help us with the beat to <i>се справим</i> with the problem

Translation:

and help us with the beat to *deal* with the problem

4. Clarification and/or elaboration.

Example 1:

1. discover their real / true identity
2. seek for approval / to be liked
3. Facebook, other social ~~med-net~~ medias affect them
? *влият* ?
4. New technology

Translation:

влият = influence

5. Structuring.

Example 1:

неправилно ползване witness
witnesses
1) ~~This new trend of taking picture of yourself is~~ Taking pictures of yourself is not something new, but the obsession of taking ~~the~~ hundreds of pictures everyday is. [...].

Translation:

неправилно ползване = incorrect use

The last example is the author's meta-discourse, which also structures the attempts to find the correct spelling of the word "witness".

7.2.2. Summary table

Table 3 summarizes the occurrences of all code-switching in the Bulgarian data according to the performed pragmatic functions.

Table 3. Pragmatic functions of code-switching in the Bulgarian scripts.

Pragmatic functions	Code-switching in the fair copies (№ of occurrences)	Code-switching in the drafts (№ of occurrences)
Framing	12	31
Free switching	0	7
Lexical need	0	1
Clarification and/or elaboration	0	1
Structuring	0	1
Total	12	41

Framing turns out to be the most frequent pragmatic function both in the fair copies and in the drafts—12 and 31 occurrences, respectively. The second comes free switching with only 7 occurrences out of 41 in the students' drafts.

7.3. Linguistic perspective

From the linguistic perspective the Bulgarian samples encompass intra-word code-switching and both intra-/inter-sentential types of switching.

7.3.1. Intra-sentential code-switching

Several cases of intra-sentential code-switching appear in the utterances of the same kind: they provide the administrative information about the student (the name, the group number, etc.). The switch introduces one piece of information, which could be separated either by a coma or by a bigger space from the other pieces of information (Example 1).

Example 1:

<i>Виктор</i>	<i>X--</i>	12345	Group 3
---------------	------------	-------	---------

Translation:

Виктор = *Victor* (the student's name)

Intra-sentential code-switching can be also found inside an utterance related to the content of the examination answer (Example 2); however, such cases are less frequent.

Example 2:

and help us with the beat to *се справим* with the problem

Translation;

and help us with the beat to *deal* with the problem

All intra-sentential code-switching in the Bulgarian samples are represented by switches of a group of words. No one-word intra-sentential switches were employed.

7.3.2. Inter-sentential code-switching

Inter-sentential code-switching can be observed in the free-standing utterances, which perform the pragmatic function of framing, and which either give the administrative information about the student (Example 1) or entitle the draft to distinguish it from the fair copy (Example 2, the first case of code-switching).

Example 1:

Десислава X – , I гр., 12345

Translation:

Десислава = *Desislava* (the student's name)

гр. = *group* (shortening from "група")

Example 2:

чернова

1. *увод* = *теза*

2. selfies show ~~moment~~ feelings at certain moment and...

3. Noone is ever ugly...

fact
addition
attractive
~~none~~

Translation:

чернова = *draft*,

увод = *introduction*,

теза = *thesis*

Inter-sentential code-switching also occurs as one of the items of a list, either with numerical and/or graphic markers (Example 2, the second case of code-switching) or being placed on a new line without any marking (Example 3).

Example 3:

In addition
Furthermore
постижения

Translation:

постижения = *achievements*

A less frequent case of inter-sentential code-switching is represented by a switch between complete sentences (Example 4).

Example 4:

1) ~~This new trend of taking picture of yourself is~~ Taking pictures of yourself is not something new, but the obsession of taking ~~the~~ hundreds of pictures everyday is. *Навлизането на новите технологии като смартфони... са част от проблема*

Translation:

Навлизането на новите технологии като смартфони... са част от проблема = *The involvement of new technologies such as smartphones... are only a part of the problem*

7.3.3. Word-internal code-switching

There is one case of word-internal code-switching, which was found in the Bulgarian data and which is illustrated in the example below:

Example 1:

“The use of sonorants, as already mentioned, adds to the *dramaticism* and gothic imagery the sonorants make the words sound more sad”.

The word "dramaticism" does not exist in English. It is an anglicized translation of the Bulgarian word "драматизъм" (transliteration: "dramatism"), which means that something is presented in a dramatic way. Given that this word-internal code-switching appears in a fair copy, it is probably a case of the lack of L2 vocabulary and its use is unconscious.

7.3.4. Summary table

The distribution of intra-/inter-sentential types of code-switching in the Bulgarian data leans towards the latter type. The majority both of the fair copies and the drafts contain mainly inter-sentential code-switching. Table 4 provides the numbers of the intra-/inter-sentential code-switching ratio in the Bulgarian samples.

Table 4. Intra-/inter-sentential code-switching.

	Intra-		Inter- + intra-		Inter-	
	Fair copies	Drafts	Fair copies	Drafts	Fair copies	Drafts
№	4	0	0	2	8	30
Total (№)	Fair copies - 12 / Drafts - 32					

7.4. Summary and conclusion

As observed above, the Bulgarian data is quite heterogeneous due to the variety of the sample types and the year of university enrolment. Whereas the

majority of the fair copies was written by the 4th year students, the drafts were collected either after an entrance examination or the linguistic examination of the 1st year students; the L2 level both of the university candidates and the 1st year students may affect the pragmatic use of code-switching. For this reason, the Bulgarian data description should be regarded as a preliminary study, which gives a first insight into the Bulgarian–English code-switching in students' personal writing; this study should be further developed and conducted on a larger and more homogeneous corpus, which will take into account the limitations and constraints of the present data, to make definite conclusions. However, some summarising observations deserve to be mentioned.

The distribution of both pragmatic functions and intra-/inter-sentential types of code-switching is unbalanced. The most frequently used pragmatic function is framing with 41 out of 49 occurrences, which is followed by free switching with only 7 occurrences. Inasmuch as the pragmatic function of framing mostly suggests the use of inter-sentential code-switching, the ratio of intra-sentential to inter-sentential code-switching is skewed in favour of the latter type (4:38, and 2 samples contained the both). Besides, very few students visually mark code-switching, but if they do mark it, the emphasis is on the utterance and not on code-switching. These results permit the deduction that the Bulgarian students, who participated in this study, use English and Bulgarian for different purposes; whereas the English language is used to elaborate the examination answer, the Bulgarian language is employed to provide the external information, which is not connected with the content of the examination answer. Each language has its own intended use.

This summary and conclusion of the Bulgarian data description closes Part II. The next step will be to view the data of the three language groups (French, Russian and Bulgarian) analytically and to interpret what the results tell us. Part III will provide the comparison and the discussion of the data and its results.

Part III

Comparison and discussion

As the title indicates, this part will take an analytical form. Given that the previous part provided an explicit description of the corpus, now these results will be viewed from an explanatory framework. First, the results described in Part II will be compared internally. Then arguments for and against external comparison will be advanced. The second chapter will be devoted to the discussion, which will attempt to give some explanations of the findings and reason the role of the academic personal writing. Finally, the chapter will be concluded with the contributions of the present work.

Chapter 8.

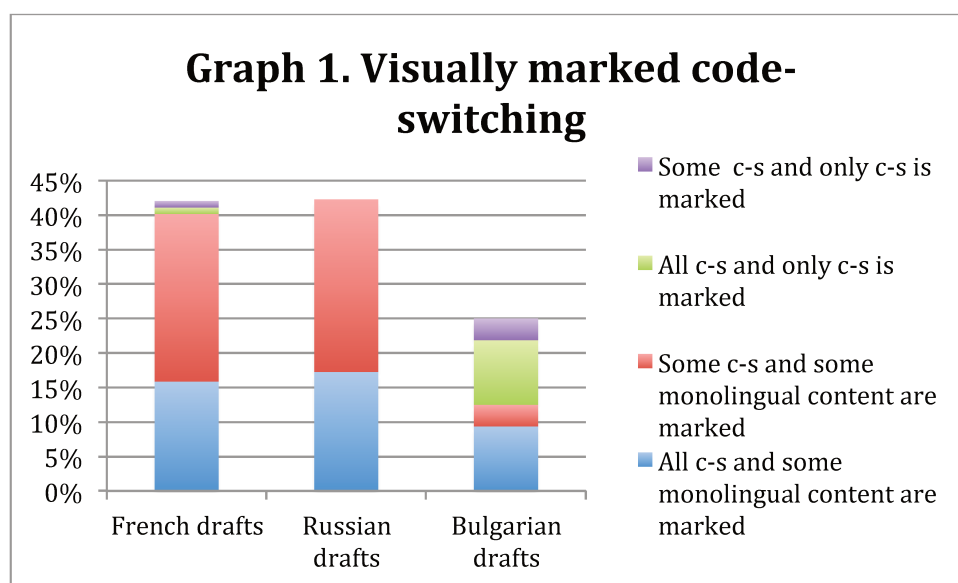
Comparison

8.1. Internal comparison

The three language groups which constitute the corpus of the present work have shown different results in almost all perspectives. The results will be shown on a separate graph for each perspective. They will be viewed both individually and in connection with other results.

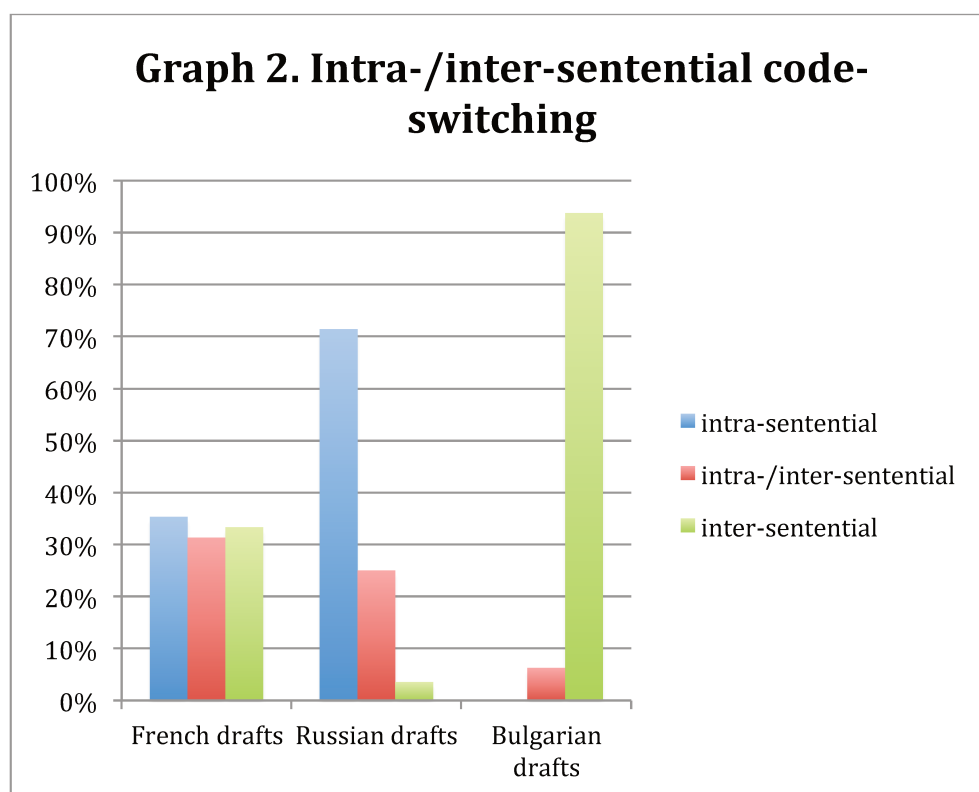
Part II explained the way the data was collected and the specificities of each sub-corpus. Whereas the French and the Russian data differ only in their quantity, the Bulgarian data are of a somewhat different nature in the participants' year of enrolment in the university, the sample type and their amount. However, it would be a regret not to include the Bulgarian data in the comparison. For this reason, the Bulgarian drafts (unlike the Bulgarian fair copies) will be put alongside the French and the Russian data; though, it should be taken into consideration that the year of the participants' enrolment may significantly influence the difference in the results in comparison with the other two sub-corpora. Given the limited number of drafts composing the whole corpus, neither definitive conclusions nor extrapolation to the respective societies can be made, but there are nevertheless a number of interesting pointers.

The least contrasting outcome was produced in the visual perspective (Graph 1). The visually marked code-switching does not exceed 45% in any of the groups.



At first sight the results seem to be different. However, when we take a closer look, this turns out not to be the case. The majority of the French participants, who visually mark code-switching (42%), mark both some of the code-switching and other monolingual content of the draft. As for the Russian participants, they mark either all or some code-switching alongside some monolingual content. The Bulgarian participants visually mark code-switching the least of all; all visual marking constitutes 25%. The green part of the bar "All code-switching and only code-switching is marked" in the Bulgarian results (9.38%) could be explained by the excessive use of the pragmatic function of framing; the majority of the cases of code-switching were represented by the word "чернова" ("draft"), which was often visually marked (underlined, capitalized, or written in a bold handwriting) to emphasize that the sample is a draft and not a fair copy. Therefore, all the three sub-corpora show that the students do not visually mark code-switching as it is, but rather as a part of a semantic or a structural unit, which they intend to visually accentuate.

The linguistic perspective gives evidence of a reverse pattern of code-switching in the French, Russian and Bulgarian data. Whereas the French participants employ intra-/inter-sentential code-switching almost equally, the Russian participants give preference to the intra-sentential type, while the Bulgarian participants put forward the inter-sentential type of switching (Graph 2).



This pattern can be explained by the pragmatic functions, which the participants of each sub-corpus employ to a different extent. The linguistic and the pragmatic perspectives are not so independent. Each pragmatic function implies a certain linguistic pattern: the employment of the intra-sentential, inter-sentential or both types of code-switching (Matras 2009). For example a stylistic switch will tend to occur mainly inside an utterance, when students use another language to explain more efficiently one part of the utterance. Similarly, the pragmatic functions of language economy or lexical need will mainly represent an insertion of (a) lexical item/items inside an utterance; whereas framing is characterized by inter-sentential switching, because the information given in the function of framing is usually spatially separated from the text; however, free switching equally suggests the application of intra- and inter-sentential code-switching. This assignment of intra-/inter-sentential code-switching to pragmatic functions is not absolute, implying the majority of the cases, but not all of them; there will be certainly some counter examples. Table 1 will illustrate all pragmatic functions employed in the corpus with the generalized assignment of the linguistic pattern and functions' distribution according to each language group.

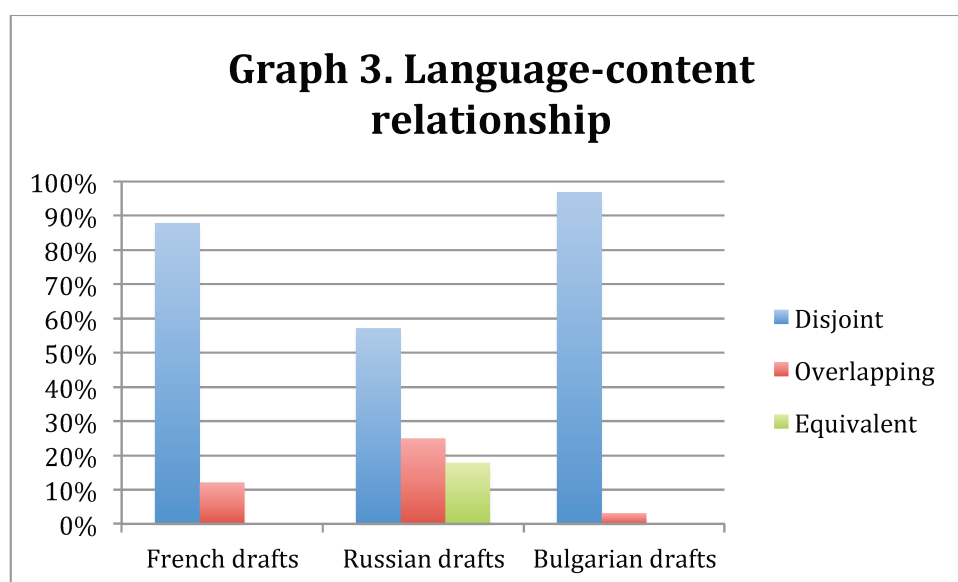
Table 1. Pragmatic functions of code-switching in the French, Russian and Bulgarian drafts.

Pragmatic functions (intra-/inter-sentential code-switching)	Code-switching (% of occurrences)		
	French drafts	Russian drafts	Bulgarian drafts
1. Free switching (intra-/inter-sentential)	30.77%	11.11%	17.07%
2. Triggered switches (intra-/inter-sentential)	14.74%	7.94%	0
3. Stylistic switches (intra-sentential)	13.46%	7.94%	0
4. Clarification and/or elaboration (intra-sentential)	9.94%	14.29%	2.44%
5. Framing (inter-sentential)	8.97%	1.59%	75.61%
6. Structuring (intra/inter-sentential)	8.65%	0	2.44%
7. Direct quotation (intra/inter-sentential)	3.53%	3.17%	0
8. Lexical need (intra-sentential)	2.88%	9.52%	2.44%
9. Parenthetical comment (intra-sentential)	2.24%	6.35%	0
10. Terminological switches (intra-sentential)	1.92%	4.76%	0
11. Language economy (intra-sentential)	1.60%	14.29%	0
12. Spelling check (inter-sentential)	0.96%	0	0
13. Reiteration (intra-sentential)	0.32%	9.52%	0
14. Vocabulary list (intra-sentential)	0	9.52%	0
Total (№ of occurrences of each function):	312	63	41

Regarding the frequency of the pragmatic functions in each sub-corpus and the respective pattern of the intra-/inter-sentential switching, the disparity illustrated in Graph 2 becomes clearer. Besides, the distribution of intra-/inter-sentential code-switching in the three language groups supports Poplack's (1979/1980/2000) assumption that fluent bilinguals tend to switch more inside the boundaries of a sentence, while the bilinguals with less language competence make a greater use of inter-sentential switchings. In the present corpus the highest number of intra-

sentential code-switching is used by the Russian participants—the students of the English department in their 4th and 5th year; the French participants use both intra-/inter-sentential switches almost to the same extent—these are the 3^d and 4th year students of the English department; and the highest amount of inter-sentential code-switching is employed by the Bulgarian participants—the 1st year students or the applicants for the English department. However, the relationship between the participants' language competence and the use of intra-/inter-sentential code-switching requires further research with the evaluation of the participants' level of the language competence and statistical tests of independence.

The distribution of the pragmatic functions also clarifies the inconsistency of language-content relationship in the data (Graph 3).



According to Sebba (2012) bilingual writing may include disjoint texts (when texts have different content), equivalent texts (when the same content is written in both languages) and overlapping texts (when some of the content is produced in both languages, whereas some is written only in one of the languages). This scheme was applied to the present corpus in the following manner: if all code-switched passages constituted new information, the draft was considered to contain disjoint texts; if all code-switched passages were the repetition of some of the content of the draft, the rough copy was deemed to contain equivalent texts; if some code-switched passages reiterated some of the content and some code-switched passages provided new information, the drafts was reckoned to represent overlapping texts. The results show that the French and the Bulgarian participants mainly write disjoint content, while the Russian participants produce disjoint,

overlapping and equivalent texts as much as 57%, 25% and 17.8%, respectively. To understand this discrepancy, it is necessary to go back to Table 1. Only two pragmatic functions imply repetition of the content by default, namely the function of vocabulary list and reiteration. Whereas only the function of reiteration in the French data with 0.32% frequency, and none of these functions is used by the Bulgarian participants, each of these functions makes up to 9.52% of frequency in the Russian data. Besides, four other pragmatic functions—free switching, stylistic switches, clarification and elaboration, and direct quotation—may also produce equivalent content; however, they may be also used for disjoint texts. Thus, the distribution of the six above-mentioned pragmatic functions in each language group elucidates 12% and 3% of overlapping content in the French and Bulgarian data, respectively, and 25% and 17.8% of overlapping and equivalent content, respectively, in the Russian data. The overall predominance of the disjoint texts indicates that the students do not need to reiterate all the content of their drafts, and thus code-switching is used as an alternative tool to produce information and not as a tool of translation or paraphrasing.

There are 14 pragmatic functions which are represented in the corpus of the present PhD thesis. The three language groups contain different amounts of the various pragmatic functions. Whereas the French and the Russian sub-corpus encompass almost the same amount of functions—13 and 12, respectively—the Bulgarian data include only five of them. There is a high probability that this difference is due to the heterogeneity of the corpus—the majority of the Bulgarian participants were either the 1st year students or the university applicants, whereas the French and the Russian students were in their 3^d, 4th and 5th year.

The present study shows that the participants code-switch in their academic personal writing regardless of the language group which they belong to. All of them employ code-switching as a tool to make their work more efficient; code-switching is one of the options along with L1, structural organising of the text, visual marking, etc., which are available for the students. Both intra-/inter-sentential types are present in the corpus. However, the varying degree of the distribution of the variables in the three sub-corpora implies that the French, the Russian and the Bulgarian participants code-switch for different purposes. All the three perspectives—the visual, the pragmatic and the linguistic perspectives—as well as the language-content relationships produced inconsistent results. Whereas all the participants, as was mentioned above, use code-switching as an efficiency tool, the efficiency is multi-faceted and here each language group emphasizes a particular aspect of efficiency, which is most appropriate for their examination-draft writing. Thus, the French participants code-switch to create a functional internal and external

structure of the draft and develop an explicit content. The Russian participants code-switch to save time and space and to provide a higher degree of precision to the given information. As for the Bulgarian participants, they code-switch for a clear external structure: there is a division between the content of the examination answer and the information surrounding it. These are the consequences, which can be drawn from the comparison between the sub-corpora; however, regarding all the limitations and the constraints of this work, a further study is needed to be conducted on a consistent and larger corpus to validate the results.

The linguistic and the pragmatic perspectives, which were applied to the corpus of the students' personal writing, were partially based on the models worked out for code-switching in oral discourse, and which were later applied to literary texts. The intra-/inter-sentential types of code-switching were examined in oral conversations (Poplack 1979/1980/2000; Li, Yu, and Fung 2012, Koban 2013, etc.) and in literature (Pfaff and Chávez 1986; Callahan 2004). All the pragmatic functions, which are applied to the analysis of bilingual literature, are taken from the models elaborated for oral communication (Montes-Alcalá 2000; Jonsson 2005; Montes-Alcalá 2012; Weston and Gardner-Chloros 2015, etc.). However, the pragmatic functions observed in this study illustrate two groups: the functions which can be found in oral code-switching and in bilingual literature, and the functions which are the exclusive characteristics of personal writing. For the reason that the functions, which have been previously developed for oral code-switching, do not cover the whole range of code-switching in the students' drafts, some new functions were added to the existing ones in the analysis of the academic personal writing, which leads us to the following question: to what extent does bilingual personal writing differ from bilingual (oral) communication on the one hand and literature on the other? This and other issues will be dealt with in the next section.

8.2. External comparison

The present work is of a chiefly exploratory character and provides the description of the under-studied field of personal writing from the bilingual perspective. Still it is important to be able to compare the results with already existing studies. Nevertheless, the comparison should be viable; the corpus and the criteria, which are applied to obtain the results, should be akin to those used elsewhere.

One of the possible points of comparison in the present research is the pragmatic functions which are employed. All the research on bilingual literature has

applied the pragmatic functions elaborated for code-switching in oral conversations (Montes-Alcalá 2000; Jonsson 2005; Montes-Alcalá 2012; Weston and Gardner-Chloros 2015, etc.). The studies show that there is at least a partial overlap between the pragmatic functions of code-switching in oral communication and in literature—literature is considered here as a generalized phenomenon without going into detail of literary genres, each having its specificities regarding the use of language alternation (Montes-Alcalá 2012; Gardner-Chloros and Weston 2015). Code-switching both in speech and in literary works portrays the identity of bilingual users and the societies in which they live through the dialogues and monologues. The success of a bilingual literary work is often conditioned by the accuracy of the imitation of the life and the natural discourse of the community (Weston and Gardner-Chloros 2015). However, code-switching in literature has an additional feature of stylistic or literary use:

In sum, CS fulfills creative, artistic, literary and stylistic functions in the plays and can be used to add emphasis to a certain word or passage, to add another level of meaning, to deepen/intensify a meaning, to clarify and to evoke richer images, to instruct the audience about a particular concept, to attempt a more faithful representation of the voice of someone else, to mark closeness, familiarity, to emphasize bonds, and to include or, on the contrary, to mark distance, break bonds and exclude. CS is used to alter expressions and meanings, and, in this aim, it can be used to add humor to expressions and/or situations or, on the contrary to make threatening situations even more threatening and painful situations even more painful, for example through the use of loaded words or reiteration. (Jonsson 2010: 1309)

Now we will lean towards the issue as to whether there is any overlap between personal writing and oral code-switching or bilingual literature. As was mentioned in the methodology, the seven functions were taken from the study on Spanish–English writing by Montes-Alcalá (2000), namely direct quotation, clarification and/or elaboration, parenthetical comment, triggered switch, stylistic switch, lexical need and free switching. These seven functions are representative both of code-switching in oral communication and literature. Two more functions were adopted from Gumperz (1982): reiteration and language economy. These two functions characterize oral discourse, no evidence for their use in literary texts was reported. However, they are employed in online interactions: both reiteration and the principle of economy were performed by Malay–English bilingual users in the Facebook context (Halim and Maros 2014). However, Facebook interaction is closer to oral communication than to literary writing. Five remaining pragmatic functions—framing, structuring, terminological switch, vocabulary list and spelling check—were added to cover specific functions of code-switching in examination

drafts. Regarding the pragmatic functions employed in the present corpus, it appears that the examination drafts show some functions of code-switching, which are common for all types of bilingual language production such as quotation, clarification and/or elaboration, parenthetical comment, triggered switch, stylistic switch, lexical need and free switching, and which they share both with oral discourse and literature. However, at the same time, the corpus has a number of characteristics which owe more to the writing process, one which of course implies a space for graphic characters, something which it hardly needs to be said is missing in oral discourse. This space is visually important—students may organize the information in a way that it will be easy to find, or apply such procedures as vocabulary listing or spelling check. These latter characteristics, in their turn, are alien to literary writing. In addition, the fact that two pragmatic functions found in the corpus (reiteration and language economy) are typical of oral communication, and that the amount of unfinished sentences in the drafts is significant suggests that examination rough copies represent an oralized form of writing.

It would be also interesting to compare the distribution of intra-/inter-sentential types of code-switching in academic personal writing, oral discourse and literature, given that this variable has been widely applied. However, this task seems complicated and difficult to perform for the moment. Numerous studies show that both intra-/inter-sentential code-switching is used in oral communications (Poplack 1979/1980/2000; Muysken 2000; Matras 2009; Kebeya 2013; etc.) and in literary texts (Callahan 2004; Jonsson 2005; Montes-Alcalá 2012; Sebba, Mahootian, and Jonsson 2012; etc.), although not many of them provide quantitative results. The evidence for oral discourse can be found in the research by Li, Yu, and Fung (2012) and Koban (2013). Li, Yu, and Fung (2012) clearly demonstrate that intra-sentential type of code-switching prevails in the formal conversational and project meetings of Chinese–English bilinguals in Hong Kong. Koban (2013) corroborates the same result, attesting that Turkish–English bilingual speakers, who live in New York, tend to employ more intra-sentential code-switching than the inter-sentential type in the face-to-face interviews. It is more difficult to find such data about written practices; only a few of them are available. Li, Yu, and Fung (2012) also examined code-switching in the Chinese articles taken from on-line news, explaining that intra-sentential type is the one which is most used. Schendl (2012) studying early English written texts says that "[...] in late medieval mixed sermons, English-dominant sermons mainly have inter-sentential switches, while the Latin ‘macaronic’ sermons [...] show a high percentage of intra-sentential switches" (Schendl 2012: 29). It is believed that intra-sentential code-switching is representative of oral

communication, whereas inter-sentential type characterizes bilingual writing. However, the evidence is rather implicit.

The issue of defining intra-/inter-sentential code-switching should be also evoked. Researchers often apply slightly different criteria for defining intra-/inter-sentential switchings depending on the corpus under study. Some authors prefer the distinction between intra-/inter-turn or turn-internal and turn-external switches (Bentahila and Davies 1995 for oral code-switching in a Moroccan community; Jonsson 2005 for code-switching in Chicano theatre). Some authors decide whether they will count the code-switch between the main and the subordinate clause as an intra-sentential code-switching (Treffers-Daller 1992), while others as inter-sentential switch (Romaine 1989/1995). Consequently, the results may vary due to the difference of the applied criteria.

Besides, the present study shows a high disparity considering the use of intra-/inter-sentential code-switching between the sub-corpora. I believe, therefore, that the comparison of the distributions of intra-/inter-sentential types of code-switching will not be functional; at the present moment the results of such a comparison will not be able to give any reasonable explanations. First, more research is needed on personal writing. The present thesis is one of the first steps in studying personal writing. The corpus has several limitations, which seriously complicate the comparison. Moreover, this work is representative only of academic personal writing, which may greatly differ from other types of personal writing such as personal diaries or personal notes. It is too early to make thorough comparisons with other types of data. Second, written code-switching in general remains understudied, especially considering the linguistic perspective. The great variety of written practices and texts requires much time to be explored. Besides, to be able to make any generalisations about code-switching in writing or even in literature, a greater amount of studies both of each single specific "genre" and of different language pairs is necessary.

Finally, this section presented a general comparison of academic personal writing with oral communication and literature on the basis of the pragmatic functions which were used. The pragmatic functions were not compared directly because their distribution varies significantly both in oral discourse due to different social contexts (Gumperz 1982; Myers-Scotton 1999; Gardner-Chloros 2009; Walters 2005) and in the literary texts depending on the literary genres and the origin of the literary work (Jonsson 2005; Jonsson 2010; Montes-Alcalá 2012). This variation is also observed within the corpus of the present work, probably because of

the difference in the language proficiency of the participants and the country where the students were studying.

Chapter 9

Discussion

9.1. Discussion

This work has presented the corpus, which consists of the students' examination drafts. These writing samples have one specific feature which distinguishes them from the majority of the writing data analysed previously. This feature is the lack of a third-person addressee. Whereas all literary and theatre works, correspondence both postal and electronic, online chats, sermons and judicial reports, all imply an external reader; the present corpus contains writing samples which were initially intended for their authors only, no external readers were meant to exist. The fact that nobody will read the samples except for their authors means that nobody will judge, comment or give their opinion about the works. That manifests itself in the freedom, that writing may take any form regarding structure, grammar, syntax, orthography, language choice, etc.—there are no limitations imposed on the writers because nobody will see what they write. I suppose that exactly for this reason, some of the students did not give me their drafts and preferred to throw them away rather than let a third person access their private space.

The students' drafts represent academic personal writing, which is only one of the multiple types of personal writing. For the time being, it is unreasonable to extrapolate the results and conclusions to personal writing in general, because the impact of such constraints as limited time and stress could not be measured regarding the structure of the corpus. The participants were limited in time to prepare the examination answer and a draft if the latter was needed. Moreover, the time limitation in addition to the necessity to be concentrated to succeed in the examination exposes the students to stress. These two factors influence the output: both the rough copy and the draft. Further research is needed to apply a measurable assessment of the external factors influencing academic personal writing. Besides, it would be of interest to study other types of personal writing, for example personal diaries, personal notes (both being the least influenced by the external factors) and note-taking during academic courses and business meetings. Only after further

research on various types of personal writing, will it be possible to say whether there is a general tendency and analogy between them.

Regarding the notion of personal writing in general, what is its niche on the multilingual arena? Much attention has always been given to oral discourse, which is representative of the community or society where it is collected. Subsequently, the flow of information easily available online gave rise to researchers' interest in electronic communication. The oral and the written modes have always been distinguished, but both of them have been seen as equally valuable resources of multilingualism. The status of literature in bilingual studies has long been debated as to whether it could be considered as naturalistic data. A recent study by Weston and Gardner-Chloros (2015) infers that multilingual literature is a complex and multi-layered resource, the study of literary code-switching "provides a useful complement to studies based on 'natural' speech samples, helping us to understand patterns of multilingual choice which are not only distilled by the writer, but placed in a meaningful, interpretable context" (Gardner-Chloros 2015: 209). For the present, multilingualism distinguishes three sources: speech, written communication and literature. Can personal writing be placed in the same line? The answer, as this study goes some way to demonstrating, is yes. Personal writing provides truly naturalistic data, which is produced by bilinguals in their daily life. Moreover, as the present work shows, personal writing is representative of society. At first, one may think that each individual has their own way of writing, structuring, using visual marking, choosing languages, etc., especially in personal writing. However, as the results show, each individual lives in a society, which affects their way of writing in the same way as it affects their way of living and thinking. The corpus of this thesis presented the academic personal writing of French, Russian and Bulgarian students. The conditions in which the drafts were collected were similar with some minor different characteristics of the examinations. The results showed inconsistency among the French, Russian and Bulgarian sub-corpora on all the levels of analysis. The only explanation is that the social environments where the participants live and study indirectly influence their academic personal writing. Besides the micro-non-linguistic factors such as the type of the examination, personal ability to manage the stress, the degree of participants' adaptation to the academic system, etc., there are probably macro-non-linguistic factors, which are of influence: educational system and pedagogical traditions, norms of writing, society's attitudes towards English and code-switching, and so on. This hypothesis, like possible linguistic reasons, requires a thorough empirical examination. Even if the link between society and personal writing is opaque and difficult to see, it should not be denied. Personal writing represents a rich and under-studied ground for sociolinguistic research.

The corpus of the present work demonstrated the presence of code-switching in academic personal writing. 26.5% of the French participants, 30.1% of the Russian participants and 42.7% of the Bulgarian participants code-switched in their examination drafts. It represents almost a third of the participants in each sub-corpus. It is therefore impossible to disapprove that the academic personal writing does comprise a source of bilingual writing. It is a unique naturalistic bilingual writing data, which is not constrained by interlocutors and the language proficiency of a reader. As has been discussed earlier, academic personal writing corresponds to an oralized form of writing, but not only. According to the pragmatic functions found in the corpus—some of which are common both to oral conversations and literature, some only to oral discourse with the remaining ones being exclusive to academic personal writing—it is conceived of as an amalgam of a monologue with oneself and visualising of language processing. The drafts manifest the written narrative in the form of a monologue addressed to their authors; they are usually read and used for the composition of a fair copy of a work by the same person. Besides, they often contain unfinished and/or ungrammatical sentences, unfinished and/or shortened lexical items and nonlinear spatial layout, which provide a sort of insight into the participants' language processing while creating the draft. Thus, academic personal writing is a unique writing practice, which merits to be studied as such. Besides, the present work has shown that code-switching in such writing does not completely mirror the functions of oral code-switching, but also introduces new ones. Code-switching in personal writing is almost a new phenomenon for social sciences, which requires a thorough examination. However, the existing models for oral and literary code-switching are not ample for it. Code-switching in personal writing should be studied intrinsically with an appropriate multimodal approach.

The present work proposes a comprehensive methodology which allows viewing code-switching from several perspectives. The academic personal writing encompasses much information, which is important to take into account. First, it contains a tangible two-dimensional space. The students fill it not only with lexical items, but also with graphic symbols; colour it with different felt-tip markers and pens of different colours; and use this space to organize the information in different ways: by grouping some elements, by leaving some space, by filling different parts of a sheet, etc. Second, rough copies have a particular syntax, where the classic rules of grammar are not always followed, namely the sentences can be unfinished, punctuation marks missing, words can be shortened or abbreviated, etc. Third, both the external and internal factors affect the writing. The external factors can be represented by time constraints, examination requirements, linguistic factors regarding the languages in use and so on; the internal factors manifest themselves in

the participants' habits, language proficiency, language processing, etc. The code-switching is only one of the elements of the bilingual academic personal writing, and should not be studied in isolation, but rather as an element of a larger system. For this reason, all the three above-mentioned points, being essential characteristics of academic personal writing, should be taken into consideration in the examination of the code-switching. If academic personal writing is conceived of as a system, code-switching will be one of its sub-systems. The components both of the main system and its subsystems interact between the sub-systems and inside them. Thus, code-switching should be studied as a phenomenon of language alternation and as an element of the writing system. One perspective of code-switching reveals its one facet, which is not enough to explain the whole phenomenon in its context. The fact that in the present corpus the distribution of intra-/inter-sentential code-switching is reflected by the distribution of pragmatic functions (as shown in Table 1) and the latter also clarifies the application of visual marking in the respective language groups, emphasizes that both structural and sociolinguistic perspectives are interrelated. As Backus argues: "what we really need is an approach that doesn't favor one over the other but instead sees form and meaning (including function) as the parts that any linguistic unit is made up of" (Backus 2015: 20). However, not only form and meaning should be regarded as the parts of a whole, but all the aspects as well as its positioning in the context.

The present work conducted research on code-switching from the visual, pragmatic and linguistic perspectives being limited by the spheres of its author's competence. It would be interesting to add a psycholinguistic and a diachronic perspective, which implies interdisciplinary collaboration. An overall approach is able to give a complete description of the phenomenon of code-switching in its context as well as incorporate it into a general account of contact phenomena. The research conducted in this thesis raises a chain of issues for further development. First, given that there are oral, written, literary code-switching and code-switching in personal writing, how many more other types are still undiscovered or under-studied as for example mental code-switching when one thinks, dreams, etc. Second, can one of the types of code-switching be considered as a primary one, implying that it appeared first and then influenced the emergence of the others? Or did all the types of code-switching develop independently under the influence of different external and internal factors? Third, what place does code-switching occupy among language contact phenomena? All these issues suggest an overall account of code-switching and its integration into a general account of language. An alternative methodology resides in the usage-based approach to code-switching (Bybee and Beckner 2010; Backus 2015). "USAGE-based theory takes language to be an embodied and social

human behavior and seeks explanations in that context" (Bybee and Beckner 2010: 827). This approach does not exclude any appropriate methods and encompasses all the available perspectives on code-switching delving into the evidence of its usage. It considers the relationship between the synchronic and diachronic modes, between the structural, functional and cognitive perspectives, thus allowing an integral view of code-switching. The usage-based approach requires the convergence of the sciences such as linguistics, anthropology, sociology, psychology, neurosciences and even politics.

9.2. Contribution of the study

The scientific contribution of this work cannot be denied. The absolute novelty of this thesis is in its study of code-switching in academic personal writing. First, a new type of code-switching, which stands along with the oral written and literary code-switching, has been recognized. Second, the bilingual personal writing of the French, the Russian and the Bulgarian participants was analysed both qualitatively and quantitatively. Third, this work demonstrates the usefulness of a multimodal methodology, proving the interrelation of the perspectives and showing the necessity of an overall view of the phenomenon of code-switching. Fourth, this study has opened up new directions, which have been unnoticed, for different social sciences both by indicating a rich field of bilingual naturalistic data and by showing the significant influence of social factors on personal writing.

The practical contribution is also valuable. This work manifested the existence of code-switching in academic personal writing. Almost a third of the participants code-switched in all the three sub-corpora. Code-switching represents one of the resources available for bilingual students. Thus, it should be commonly accepted, especially in academic institutions, that the students who are in contact with several languages may use code-switching as one of the tools at their disposal. This action should not be neither denied, nor forbidden, nor punished to avoid the mental block and the possible negative influence on the students' academic performance; of course the effects of the professors' opposition to the students' use of code-switching should previously be thoroughly examined. Moreover, during the methodological courses for the 1st year students it could be even emphasized that code-switching can be a part of their note-taking system, which they will elaborate during their years of studies. Code-switching can be employed for such purposes as language economy, for the precision of the information, for time economy, for structuring purposes, etc. Some exercises could be performed in class. For example

to take notes as fast as possible by timing students' note-taking and then comparing their time performance and the techniques they used; code-switching should be emphasized to be one of such techniques introduced to students before the exercise. Another exercise could consist of taking the most precise notes of a video (the video could be in a foreign language to facilitate the use of code-switching while note-taking) with a following discussion of the techniques employed. The group may also work on the structure of notes. A creative task can be given to the class: to take notes with a highly clear structure so that a person who has not read the original document could make his or her presentation based on the notes of another student. Two or three different texts (in a foreign language) should be selected by the lecturer, each having a complex internal structure. The students should take notes of a given document, then exchange their notes with a neighbour and prepare an oral presentation of the text that they have not seen with the help of their neighbour's notes. The analysis and the discussion of the structures of students' notes should be performed afterwards. Of course, note-taking techniques are individual and personal and students will create their own system using the given instructions and advice, but basing on their personal experience, strong and weak points and their aims. However, they should be initiated into different possibilities they have before being able to choose what suits them best.

Another practical contribution of the present thesis is a didactic suggestion to use the students' academic personal writing as a material to introduce and explore the phenomenon of code-switching in class during sociolinguistic courses. If the group is monolingual, the samples of other students such as the drafts of the present corpora, can be demonstrated in class and used as hand-outs. If the group is bilingual (with low or high level of the language proficiency) the experiment can be conducted during the lesson by asking the students either to look through their notes and to find whether they use two or more languages or to ask them to prepare a rough copy (either at once for a test, for example, or at home for a given homework task and to take it with them for the class work). Then through the vote system, the professor can count the amount of the students who code-switch in their personal writing in the given group. Afterwards, a discussion can begin on introducing code-switching with the examples taken from the group; later it can be developed into exploration of different linguistic types of code-switching, the pragmatic functions and the visual marking also basing either on the samples taken from the same group or adding some external examples of other students.

Conclusion

The review of literature presented in this work unveiled a theoretical delineation of bilingual personal writing. Naturalistic written data find themselves at the intersection of general notions of bilingualism, language alternation and note-taking, and yet they have so far been very much under-studied. Only a few works have examined this field such as the sociolinguistic research by Montes-Alcalá (2000), who focused on the pragmatic functions of code-switching in a personal diary and the psychological studies by Faraco (1997) and Barbier et al. (2006) devoted to note-taking technique in L2. Taking inspiration from the above-mentioned studies and aiming at conducting both the qualitative and quantitative research on code-switching in personal writing, this PhD thesis has provided a complete multimodal description of code-switching in academic personal writing. The novelty of this research resides both in its corpus and methodology. It is the first study to have collected a corpus of bilingual academic drafts (composed of the three sub-corpora: French, Russian and Bulgarian). Such a corpus requires an original approach. Thus after the description of the data from the visual, pragmatic and linguistic perspectives, the findings were discussed in their interrelation to each other emphasising the composite character of the code-switching. Code-switching represents the tool which students skillfully use in their academic personal writing. The purposes of use differ not only among individuals but also among societies, showing the tendency of the French participants to employ code-switching for external and internal structuring of the draft as well as to maintain the continuity of the expression characterized by its explicit content. The Russian participants tend to use code-switching as a tool of efficiency and economy for time, space and quality of their personal writing. The Bulgarian participants are inclined to code-switch to provide the external information, which is not linked with the content of their examination answers. Regarding the limitations and the constraints of the present corpus, further research is needed to attest or contest the findings. Nonetheless, the inconsistency of the results highlights the influence of external factors on personal writing. Hence, this study opens two sociolinguistic perspectives. First, to explore personal writing as synchronic sociolinguistic evidence; to determine the social factors that affect code-switching and which make it different across various countries/cultures/communities. Second, to delve into personal writing to have an insight into society from the side, which has not been studied yet, and which may even be called the hidden "personal" side of the society.

This work has placed code-switching in academic personal writing in the same line with oral, written and literary code-switching both emphasising its social features and bringing the recognition to code-switching in personal writing which it deserves. This writing practice represents a naturalistic bilingual written data, which merits to be studied as such, and which is a potentially fruitful area of research in linguistics, sociolinguistics, psychology, psycholinguistics, neurosciences, educational sciences, etc. Moreover, this study has argued for the necessity of an interdisciplinary approach to provide a comprehensive understanding of the phenomenon of code-switching and its role on the language arena. Code-switching should be viewed both as a component of a bigger whole and as a system itself comprised of inter-connected elements.

References

- Altarriba, Jeanette, and Roberto R. Heredia, eds. 2008. *An Introduction to Bilingualism : Principles and Processes*. New York NY: London.
- Altemeier, Leah, Janine Jones, Robert D. Abbott, and Virginia W. Berninger. 2006. 'Executive Functions in Becoming Writing Readers and Reading Writers: Note Taking and Report Writing in Third and Fifth Graders'. *Developmental Neuropsychology* 29 (1): 161–73.
- Alvarez-Cáccamo, Celso. 1998. 'Codes'. In AAA98, session "Language Matters: A Lexicon for the Millennium", 1–5. Philadelphia.
- . 2002. 'From "switching Code" to "code-Switching": Towards a Reconceptualisation of Communicative Codes'. In *Code-Switching in Conversation: Language, Interaction and Identity*, edited by Peter Auer, 3d edition, 29–49. London, New York: Routledge.
- Anchimbe, Eric A. 2015. 'Code-Switching: Between Identity and Exclusion'. In *Code-Switching between Structural and Sociolinguistic Perspectives*, 139–61. *Linguae & Litterae* 43. Berlin, Boston: De Gruyter.
- Androutsopoulos, Jannis. 2013. 'Code-Switching in Computer-Mediated Communication'. In *Handbook of the Pragmatics of CMC*, edited by Susan Herring, Dieter Stein, and Tuija Virtanen, 667–94. Berlin, Boston: Mouton de Gruyter.
- Atawneh, Ahmed. 1992. 'Code-Mixing in the Speech of Arabic-English Bilinguals'. In *Perspectives on Arabic Linguistics IV: Papers from the Fourth Annual Symposium on Arabic Linguistics*, edited by Ellen Broselow, Mushira Eid, and John McCarthy, 219–45. Amsterdam: John Benjamins.
- Auer, Peter. 1984. 'On the Meaning of Conversational Code-Switching'. In *Interpretive Sociolinguistics*, by Peter Auer, 87–112. Tübingen: Narr.
- . 1988. 'A Conversation Analytic Approach to Code-Switching and Transfer'. In *Codeswitching: Anthropological and Sociolinguistic Perspectives*, by Monica Heller, 187–213. Berlin: Mouton de Gruyter.
- Auer, Peter, and Li Wei, eds. 2007. *Handbook of Multilingualism and Multilingual Communication*. Handbooks of Applied Linguistics Communication Competence, Language and Communication Problems, Practical Solutions Vol. 5. Berlin, New York NY: Mouton de Gruyter.

- Backus, Ad. 2015. 'A Usage-Based Approach to Code-Switching: The Need for Reconciling Structure and Function'. In *Code-Switching between Structural and Sociolinguistic Perspectives*, edited by Gerald Stell and Kofi Yakpo, 19–37. *Linguae & Litterae* 43. Berlin, Boston: De Gruyter.
- Baddeley, Alan. 1998. 'Working Memory'. *Comptes Rendus de l'Académie Des Sciences - Series III - Sciences de La Vie* 321 (2–3): 167–73.
- . 2003. 'Working Memory and Language: An Overview'. *Journal of Communication Disorders*, ASHA 2002, 36 (3): 189–208.
- . 2010. 'Working Memory'. *Current Biology* 20 (4): R136–40.
- Baetens Beardsmore, Hugo. 1982. *Bilingualism: Basic Principles*. Multilingual Matters 1. England, Clevedon, Avon: Tieto.
- . 1986. *Bilingualism: Basic Principles*. 2nd edition. 2. Multilingual Matters.
- Bailey, Benjamin. 2007. 'Multilingual Forms of Talk and Identity Work'. In *Handbook of Multilingualism and Multilingual Communication*, edited by Peter Auer and Li Wei, 341–70. Berlin: Mouton de Gruyter.
- Baker, Colin. 2011. *Foundations of Bilingual Education and Bilingualism: 5th Edition*. Multilingual Matters.
- Baker, Colin, and Sylvia Prys Jones. 1998. *Encyclopedia of Bilingualism and Bilingual Education*. Multilingual Matters.
- Baker, Colin Ronald. 1993. *Foundations of Bilingual Education and Bilingualism*. Multilingual Matters 95. Clevedon, Philadelphia, Adelaide: Multilingual matters Ltd.
- Barbier, Marie-Laure, Martine Faraco, Annie Piolat, Jean-Yves Roussey, and Tsu Kida. 2003. 'Comparaison de La Prise de Notes D'étudiants Japonais et Espagnol Dans Leur Langue Native et En Français L2'. *Arob@se*, *Www.arobase.to* 1–2: 180–203.
- Barbier, Marie-Laure, and Annie Piolat. 2005. 'L1 and L2 Cognitive Effort during Note Taking and Writing'. *Proceedings Writing 2004 Geneva*.
- Barbier, Marie-Laure, Jean-Yves Roussey, Annie Piolat, and Thierry Olive. 2006. 'Note-Taking in Second Language: Language Procedures and Self Evaluation of the Difficulties'. *Current Psychology Letters. Behaviour, Brain & Cognition*, no. 20, NaN, 2006 (December).
- Beare, Sophie, and Johanne S. Bourdages. 2007. 'Skilled Writers' Generating Strategies in L1 and L2: An Exploratory Study'. In *Writing and Cognition:*

- Research and Applications*, edited by Mark Torrance, Luc Van Vaes, and David Galbraith, 20:151–61. Studies in Writing. Amsterdam: Emerald Group Publishing Limited.
- Bednarek, Adam. 2012. *Interdisciplinary Perspectives in Cross-Cultural Communication*. Muenchen: Lincom Europa.
- Bell, Roger Thomas. 1976. *Sociolinguistics: Goals, Approaches, and Problems*. London: Bloomsbury.
- Bentahila, Abdelâli, and Eirlys E. Davies. 1995. 'Patterns of Code-Switching and Patterns of Language Contact'. *Lingua* 96: 75–93.
- Bernard, Roger, and James McLellan. 2014. *Codeswitching in University English-Medium Classes: Asian Perspectives*. Bristol, Buffalo, Toronto: Multilingual Matters.
- Bhatia, Tej K., and William C. Ritchie, eds. 2004. *The Handbook of Bilingualism*. Blackwell Publishing.
- Bhatt, Rakesh M. 2008. 'In Other Words: Language Mixing, Identity Representations, and Third space'. *Journal of Sociolinguistics* 12 (2): 177–200.
- Biber, Douglas, and Edward Finegan. 1994. *Sociolinguistic Perspectives on Register*. Oxford University Press.
- Blom, Jan-Petter, and John Joseph Gumperz. 2000 (reprint of 1972). 'Social Meaning in Linguistic Structure: Code-Switching in Norway'. In *The Bilingualism Reader*, edited by Li Wei, 111–36. London: Routledge.
- Bloomfield. 1933. *Language*. 1956 edition. New York: Holt and Company.
- Boch, Françoise, and Annie Piolat. 2005. 'Note Taking and Learning: A Summary of Research'. *The WAC Journal* 16: 101–13.
- Bokamba, Eyamba. 1988. 'Code-Mixing, Language Variation, and Linguistic Theory: Evidence from Bantu Languages'. *Lingua* 76: 21–62.
- Broersma, Mirjam, and Kees De Bot. 2006. 'Triggered Codeswitching: A Corpus-Based Evaluation of the Original Triggering Hypothesis and a New Alternative'. *Bilingualism: Language and Cognition* 9 (1): 1–13.
- Bullock, Barbara E., and Almeida Jacqueline Toribio, eds. 2009. *The Cambridge Handbook of Linguistic Code-Switching*. Cambridge University Press.

- Butler, Yuko G., and Kenji Hakuta. 2004. 'Bilingualism and Second Language Acquisition'. In *The Handbook of Bilingualism*, edited by Tej K. Bhatia and William C. Ritchie, 114–44. Malden, Oxford, Victoria: Blackwell Publishing.
- Bybee, Joan L., and Clay Beckner. 2010. 'Usage-Based Theory'. In *The Oxford Handbook of Linguistic Analysis*, edited by Bernd Heine and Heiko Narrog, First edition. Oxford University Press.
- Callahan, Laura. 2004. *Spanish/English Codeswitching in a Written Corpus*. John Benjamins Publishing.
- Cantone, Katja F. 2007. *Code-Switching in Bilingual Children*. Studies in Theoretical Psycholinguistics Vol. 37. Dordrecht Springer.
- Carpenter, Dora. 2011. 'Harissa.com: Narrative(s), Code-Switching and Nostalgia – a Case Study'. In *Code-Switching, Languages in Contact and Electronic Writings*, by Foued Laroussi, 15–32. Peter Lang.
- Chan, Brian Hok-Shing. 2003. *Aspects of the Syntax, the Pragmatics, and the Production of Code-Switching: Cantonese and English*. Berkeley Insights in Linguistics and Semiotics 51. New York Peter Lang.
- Clyne, Michael. 1970. 'Bilingual Speech Phenomena (with Special Reference to German–English Bilinguals in Victoria)'. *Kivung* 3 (2): 99–111.
- . 1980. 'Triggering and language processing.' *Canadian Journal of Psychology/Revue canadienne de psychologie* 34 (4): 400–406.
- . 2000 (reprint of 1987). 'Constraints on Code-Switching: How Universal Are They?' In *The Bilingualism Reader*, edited by Li Wei, 257–80. London: Routledge.
- Crystal, David. 2006. *Language and the Internet*. 2nd edition. Cambridge University Press.
- Cumming, Alister. 1990. 'Metalinguistic and Ideational Thinking in Second Language Composing'. *Written Communication* 7 (4): 482–511.
- Datta, Manjula. 2000. *Bilinguality and Literacy: Principles and Practice*. London New York NY Continuum.
- Delamotte, Régine, and Cecile Desoutter. 2011. 'Languages in Contact in Business E-Mails and Student Forums'. In *Code-Switching, Languages in Contact and Electronic Writings*, by Foued Laroussi, 53–72. Peter Lang.
- Deppman, Jed, Daniel Ferrer, and Michael Groden. 2004. *Genetic Criticism: Texts and Avant-Textes*. University of Pennsylvania Press.

- Dingfelder Stone, Maren. 2015. 'The Theory and Practice of Teaching Note-Taking'. In *To Know How to Suggest...: Approaches to Teaching Conference Interpreting*, edited by Dörte Andres and Martina Behr, 145–69. Frank & Timme.
- Edwards, John. 2004. 'Foundations of Bilingualism'. In *The Handbook of Bilingualism*, edited by Tej K. Bhatia and William C. Ritchie, 7–31. Malden, Oxford, Victoria: Blackwell Publishing.
- Escamilla, Kathy, and Susan Hopewell. 2007. 'The Role of Code-Switching in the Written Expression of Early Elementary: Simultaneous Bilinguals'. American Educational Research Association.
- Faraco, Martine. 1997. 'Étude longitudinale de la prise de notes d'un cours universitaire français : le cas d'étudiants étrangers d'un cursus européen'. *ASp. la revue du GERAS*, no. 15–18 (December): 41–54.
- Faraco, Martine, Marie-Laure Barbier, Achille Falaise, and Sonia Branca-Rosoff. 2003. 'Codage et Traitement Automatique de Corpus Pour L'étude Des Prises de Notes En Français Langue Première et Langue Seconde'. *Arob@se, Www.arobase.to* 1–2: 97–117.
- Ferguson, Charles A. 2000 (reprint of 1959). 'Diglossia'. In *The Bilingualism Reader*, edited by Li Wei, 65–80. London, New York: Routledge.
- Gafaranga, Joseph. 2007. 'Code-Switching as a Conversational Strategy'. In *Handbook of Multilingualism and Multilingual Communication*, edited by Peter Auer and Li Wei, 279–314. Berlin: Mouton de Gruyter.
- Gafaranga, Joseph, and Maria-Carme Torras. 2002. 'Interactional Otherness: Towards a Redefinition of Codeswitching'. *International Journal of Bilingualism* 6 (1): 1–22.
- Ganascia, Jean-Gabriel, Irène Fenoglio, and Jean-Louis Lebrave. 2004. 'Manuscripts, genèse et documents numérisés. EDITE: une étude informatisée du travail de l'écrivain.' *Document numérique* 2004/4 (8): 91–110.
- Gardner-Chloros, Penelope. 2008. 'Bilingual Speech Data: Criteria for Classification'. In *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*, edited by Li Wei and Melissa G. Moyer, 53–72. Oxford: Blackwell Publishing.
- . 2009. *Code-Switching*. Cambridge: Cambridge University Press.

- Gardner-Chloros, Penelope, and Malcolm Edwards. 2004. *Assumptions behind Grammatical Approaches to Code-Switching: When the Blueprint Is a Red Herring*. London: Birkbeck ePrints.
- Gardner-Chloros, Penelope, and Daniel Weston. 2015. 'Code-Switching and Multilingualism in Literature'. *Language and Literature* 24 (3): 182–93.
- Georgieva, Maria. 2011. *Global English in Bulgarian Context: A Sociolinguistic Perspective*. Varna: Silueti Publishing House.
- Gile, Daniel. 2009. *Basic Concepts and Models for Interpreter and Translator Training*. Revised edition. John Benjamins Publishing.
- Gimenez, Guy, and Sophie Barthélémy. 2011. 'Le temps de la prise de notes'. *Revue de psychothérapie psychanalytique de groupe* 56 (1): 171–85.
- Gómez-Ruiz, M. I. 2010. 'Bilingualism and the Brain: Myth and Reality'. *Neurología (English Edition)* 25 (7): 443–52.
- Green, David William. 1998. 'Mental Control of the Bilingual Lexico-Semantic System.' *Bilingualism: Language and Cognition* 1 (2): 67–81.
- . 2011. 'Language Control in Different Contexts: The Behavioral Ecology of Bilingual Speakers'. *Frontiers in Psychology* 2 (May).
- Grésillon, Almuth. 1994. *Éléments de critique génétique: lire les manuscrits modernes*. Paris Presses universitaires de France.
- . 2008. *La mise en oeuvre: itinéraires génétiques*. Textes & manuscrits. Paris CNRS éd.
- Griggs, Peter. 2007. *Perspective sociocognitive sur l'apprentissage des langues étrangères: situation naturelle ou guidée*. Langue & parole. Paris: l'Harmattan.
- Groot, Annette M. B. de, and Judith F. Kroll. 1997. *Tutorials in Bilingualism: Psycholinguistic Perspectives*. Mahwah, N.J: Lawrence Erlbaum Associates.
- Grosjean, François. 1982. *Life with Two Languages: An Introduction to Bilingualism*. Cambridge, Massachusetts, London: Harvard University press.
- . 1989. 'Neurolinguists, Beware! The Bilingual Is Not Two Monolinguals in One Person'. *Brain and Language* 36 (1): 3–15.
- Guichon, Nicolas. 2006. 'Le brouillon: aide mémorielle et trace de la compréhension de l'oral'. *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Apliut*, no. Vol. XXV N° 2 (June): 61–76.

- Gumperz, John Joseph. 1982. *Discourse Strategies*. Studies in Interactional Sociolinguistics 1. Cambridge London New York: Cambridge University Press.
- . 1984. *Discourse Strategies*. Studies in Interactional Sociolinguistics 1. Cambridge: Cambridge University Press.
- Hakuta, Kenji. 1986. *Mirror of Language: The Debate on Bilingualism*. New York: Basic Books, Inc., Publishers.
- Halim, Nur Syazwani, and Marlyna Maros. 2014. 'The Functions of Code-Switching in Facebook Interactions'. *Procedia - Social and Behavioral Sciences* 118: 126–33.
- Hamers, Josiane F., and Michel H. A. Blanc. 1989. *Bilinguality and Bilingualism*. Cambridge, New York (N.Y.), Melbourne: Cambridge University Press.
- . 2000. *Bilinguality and Bilingualism*. 2nd ed. Cambridge, New York (N.Y.), Melbourne: Cambridge University Press.
- Haugen, Einar. 1950. 'The Analysis of Linguistic Borrowing'. *Language* 26: 210–31.
- Heller, Monica, ed. 1988. *Codeswitching: Anthropological and Sociolinguistic Perspectives*. Berlin: Mouton de Gruyter.
- . 1997. 'Autonomy and Interdependence: Language in the World'. *International Journal of Applied Linguistics* 7 (1): 79–85.
- . 2002. 'Globalization and the Commodification of Bilingualism in Canada'. In *Globalization and Language Teaching*, edited by David Block and Deborah Cameron, 47–64. Routledge.
- . 2003. 'Globalization, the New Economy, and the Commodification of Language and Identity'. *Journal of Sociolinguistics* 7 (4): 473–92.
- . 2008. 'Language and the Nation-State: Challenges to Sociolinguistic Theory and Practice'. *Journal of Sociolinguistics* 12 (4): 504–24.
- Hlavac, Jim. 1999. '32 Years on and Still Triggering: Psycholinguistic Processes as Motivation for Switching amongst Croatian-English Bilinguals'. *Monash University Linguistic Papers* 2 (1): 11–24.
- Isurin, Ludmila, Donald Winford, and Kees De Bot. 2009. *Multidisciplinary Approaches to Code Switching*. Studies in Bilingualism v. 41. Amsterdam, Philadelphia: John Benjamins Publishing Company.

- Jackendoff, Ray. 2012. *A User's Guide to Thought and Meaning*. Second edition. Oxford GB, New York, Auckland, etc.: Oxford University Press.
- Jacobson, Rodolfo. 2001. *Codeswitching Worldwide II*. Trends in Linguistics. Studies and Monographs. Mouton de Gruyter.
- Jonsson, Carla. 2005. 'Code-Switching in Chicano Theater: Power, Identity and Style in Three Plays by Cherríe Moraga'. Umeå: Umeå universitet.
- . 2010. 'Functions of Code-Switching in Bilingual Theater: An Analysis of Three Chicano Plays'. *Journal of Pragmatics* 42 (5): 1296–1310.
- . 2012. 'Making Silenced Voices Heard: Code-Switching in Multilingual Literary Texts in Sweden.' In *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*, edited by Mark Sebba, Shahrzad Mahootian, and Carla Jonsson, 212–32. New York, London: Routledge.
- Kebeya, Hilda. 2013. 'Inter- and Intra-Sentential Switching: Are They Really Comparable?' *International Journal of Humanities and Social Science* 3 (5): 225–33.
- Koban, Didem. 2013. 'Intra-Sentential and Inter-Sentential Code-Switching in Turkish-English Bilinguals in New York City, U.S'. *Procedia - Social and Behavioral Sciences*, Akdeniz Language Studies Conference, May, 2012, Turkey, 70 (January): 1174–79.
- Kootstra, Gerrit Jan. 2015. 'A Psycholinguistic Perspective on Code-Switching: Lexical, Structural, and Socio-Interactive Processes'. In *Code-Switching between Structural and Sociolinguistic Perspectives*, 39–64. *Linguae & Litterae* 43. Berlin, Boston: De Gruyter.
- Krashen, Stephen. 1982. *Principles and Practice in Second Language Acquisition*. First Internet edition July 2009. Pergamon Press Inc.
- . 2004. 'Free Voluntary Reading: New Research, Applications, and Controversies'. In . Singapore.
- Krashen, Stephen D. 1993. *The Power of Reading: Insights from the Research*. Libraries Unlimited.
- Labov, William. 1972. *Sociolinguistic Patterns*. Conduct and Communication. Philadelphia: University of Pennsylvania Press.
- Lagarde, Christian. 2001. *Des écritures bilingues : sociolinguistique et littérature*. Sociolinguistique. Paris, Montréal, Budapest, etc.: l'Harmattan.

- Lambert, Wallace E. 1973. 'Culture and Language as Factors in Learning and Education'. In Annual Learning Symposium on "Cultural Factors in Learning" (5th, Western Washington State College, Bellingham, Washington, November 1973), 54. Washington.
- Laroussi, Foued. 2011. *Code-Switching, Languages in Contact and Electronic Writings*. Sprache, Mehrsprachigkeit Und Sozialer Wandel Bd. 14. Frankfurt am Main, Berlin, Bern, etc: Peter Lang.
- Lebrave, Jean-Louis. 1983. 'Lecture et Analyse Des Brouillons'. *Langages* 17 (69): 11–23.
- . 1991. 'L'hypertexte et L'avant-Texte'. *Linx* 4 (1): 103–19.
- . 2006a. 'Hypertextes — Mémoires — Ecriture'. Article. <http://www.item.ens.fr>. October 16.
- . 2006b. 'L'édition critique au XXIe siècle'. Article. <http://www.item.ens.fr>. October 31.
- Lewis, W. Gwyn, and Hunydd Andrews. 2014. 'Bilingualism in Higher Education'. In *Advances in the Study of Bilingualism*, 171–97. Bristol, Buffalo, Toronto: Multilingual Matters.
- Li Wei. 2000a. 'Dimensions of Bilingualism'. In *The Bilingualism Reader*, 3–25. London, New York: Routledge.
- . , ed. 2000b. *The Bilingualism Reader*. London, New York: Routledge.
- Li Wei, and Melissa G. Moyer, eds. 2008. *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. Oxford: Blackwell Publishing.
- Li, Ying, Yue Yu, and Pascale Fung. 2012. 'A Mandarin-English Code-Switching Corpus'. In 8th International Conference on Language Resources and Evaluation (LREC), Istanbul, 2515–19.
- Lienard, Fabien, and Marie-Claude Penloup. 2011. 'Language Contacts and Code-Switching in Electronic Writing: The Case of the Blog'. In *Code-Switching, Languages in Contact and Electronic Writings*, by Foued Laroussi, 73–86.
- Lyons, John. 1968. *Introduction to Theoretical Linguistics*. Cambridge University Press.
- Mackey, William F. 1970. 'Interference, Integration and the Synchronic Fallacy.' In Twenty-first Round Table Meeting on Linguistics and Language Studies. Georgetown University, Washington, D.C.

- . 2000 (reprint of 1962). 'The Description of Bilingualism'. In *The Bilingualism Reader*, 26–54. London, New York: Routledge.
- Macnamara, John. 1967. 'The Bilingual's Linguistic Performance - a Psychological Overview' 23 (2): 58–77.
- MacSwan, Jeff. 2009. 'Generative Approaches to Code-Switching'. In *The Cambridge Handbook of Linguistic Code-Switching*, edited by Barbara E. Bullock and Almeida Jacqueline Toribio, 309–35. Cambridge University Press.
- Mallinson, Christine, Becky Childs, and Gerard Van Herk. 2013. *Data Collection in Sociolinguistics: Methods and Applications*. New York, London: Routledge.
- Marley, Dawn. 2011. 'Code-Switching in Websites for the Moroccan Diaspora'. In *Code-Switching, Languages in Contact and Electronic Writings*, by Foued Laroussi, 33–52. Peter Lang.
- Matras, Yaron. 2009. *Language Contact*. Cambridge Textbooks in Linguistics. Cambridge University Press.
- Mbodj-Pouye, Aïssatou, and Cécile Van Den Avenne. 2012. 'Vernacular Literacy Practices in Present-Day Mali: Combining Ethnography and Textual Analysis to Understand Multilingual Texts'. In *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*, by Mark Sebba, Shahrzad Mahootian, and Carla Jonsson, 170–91. University of Birmingham, UK: Routledge Critical Studies in Multilingualism.
- McClure, Erica. 1981. 'Formal and Functional Aspects of the Codeswitched Discourse of Bilingual Children'. In *Latino Language and Communicative Behavior*, edited by Richard Durán, 69–92. Norwood, NJ: Ablex.
- McCormick, Kay. 2002. *Language in Cape Town's District Six*. Oxford: Oxford University Press.
- Meisel, Jürgen M. 1994. 'Code-Switching in Young Bilingual Children: The Acquisition of Grammatical Constraints'. *Studies in Second Language Acquisition* 16 (4): 413–39.
- Montes-Alcalá, Cecilia. 2000. 'Two Languages, One Pen: Socio-Pragmatic Functions In Written Spanish-English Code-Switching'. Santa Barbara: University of California.
- . 2012. 'Code-Switching in US-Latino Novels'. In *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*, edited by Mark Sebba, Shahrzad Mahootian, and Carla Jonsson,

- 68–88. University of Birmingham, UK: Routledge Critical Studies in Multilingualism.
- Muysken, Pieter. 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge, New York, Melbourne, etc: Cambridge University Press.
- . 2004. ‘Two Linguistic Systems in Contact: Grammar, Phonology and Lexicon’. In *The Handbook of Bilingualism*, edited by Tej K. Bhatia and William C. Ritchie, 147–68. Blackwell Publishing.
- . 2007. ‘Mixed Codes’. In *Handbook of Multilingualism and Multilingual Communication*, edited by Peter Auer and Li Wei, 315–40. Berlin: Mouton de Gruyter.
- Myers-Scotton, Carol. 1989. ‘Codeswitching with English: Types of Switching, Types of Communities’. *World Englishes* 8 (3): 333–46.
- . 1999. ‘Explaining the Role of Norms and Rationality in Code-Switching’. *Journal of Pragmatics* 32: 1259–71.
- . 2000 (reprint of 1988). ‘Code-Switching as Indexical of Social Negotiations’. In *The Bilingualism Reader*, edited by Li Wei, 137–65. London: Routledge.
- . 2001. ‘The Matrix Language Frame Model: Developments and Responses’. In *Codeswitching Worldwide II*, edited by Rodolfo Jacobson, 23–58. Trends in Linguistics. Studies and Monographs. Mouton de Gruyter.
- . 2006. *Multiple Voices: An Introduction to Bilingualism*. Malden (MA), USA: Blackwell Publishing.
- Myers-Scotton, Carol, and Janice L. Jake. 2000a (reprint of 1995). ‘Matching Lemmas in a Bilingual Language Competence and Production Model: Evidence from Intrasentential Code-Switching’. In *The Bilingualism Reader*, edited by Li Wei, 281–320. London: Routledge.
- . 2000b. ‘Testing the 4-M Model: An Introduction’. *International Journal of Bilingualism* 4 (1): 1–8.
- . 2009. ‘A Universal Model of Code-Switching and Bilingual Language Processing and Production’. In *The Cambridge Handbook of Linguistic Code-Switching*, edited by Barbara E. Bullock and Almeida Jacqueline Toribio, 336–57. Cambridge University Press.
- Namba, Kazuhiko. 2005. ‘What Is Code-Switching?’.

- Niemiec, Ewa. 2010. 'Between Us Bilinguals. A Fairly Unbiased Dissertation on Monolingual and Bilingual Views on Code- Switching'. Warsaw University.
- Nilep, Chad. 2006. "'Code Switching" in Sociocultural Linguistics'. *Colorado Research in Linguistics* 19 (June).
- Olive, Thierry, Jean-Louis Lebrave, Jean-Michel Passerault, and Nathalie Le Bigot. 2010. 'La Dimension Visuo-Spatiale de La Production de Textes: Approches de Psychologie Cognitive et de Critique Génétique'. *Langages* 2010/1 (177): 29–55.
- Omer, Danielle. 2003. 'La prise de notes à la française pour des noteurs non natifs'. *Arob@se*, www.arobase.to 1–2: 141–51.
- Onysko, Alexander. 2007. *Anglicisms in German: Borrowing, Lexical Productivity, and Written Codeswitching*. Walter de Gruyter.
- Osmond, Humphry. 1955. 'Notes on Note-Taking'. *The Lancet*, Originally published as Volume 2, Issue 6887, 266 (6887): 415–17.
- Paradis, Michel. 2004. *A Neurolinguistic Theory of Bilingualism*. Amsterdam: Philadelphia.
- Parafita Couto, M. Carmen, Peredur Davies, Diana Carter, and Margaret Deuchar. 2014. 'Factors Influencing Code-Switching'. In *Advances in the Study of Bilingualism*, edited by Enlli Môn Thomas and Ineke Mennen, 111–38. Bristol, Buffalo, Toronto: Multilingual Matters.
- Pfaff, Carol. 1979. 'Constraints on Language Mixing: Intrasentential Codeswitching and Borrowing in Spanish/English'. *Language* 55: 291–318.
- Pfaff, Carol, and Laura Chávez. 1986. 'Spanish/English Code-Switching: Literary Reflections of Natural Discourse'. In *Missions in Conflict: Essays on U.S. - Mexican Relations and Chicano Culture*, edited by Renate von Bardeleben, 229–54. Tübingen: Gunter Narr Verlag.
- Piolat, Annie. 2003. 'Introduction : Pourquoi Étudier La Prise de Notes En Langue Seconde?' *Arob@se*, Www.arobase.to 1–2: 1–5.
- . , ed. 2004. *Écriture: approches en sciences cognitives*. 1 vols. Langues et écritures, Aix-en-Provence, France: Publications de l'Université de Provence.
- . 2006. *La prise de notes*. 1 vols. Que sais-je?, Paris, France: Presses universitaires de France.

- Piolat, Annie, and Marie-Laure Barbier. 2007. 'De l'écriture elliptique estudiantine: analyse descriptive de prise de note et de brouillons'. *Langue française* 3 (155): 84–100.
- Piolat, Annie, Thierry Olive, and Ronald T. Kellogg. 2005. 'Cognitive Effort during Note Taking'. *Applied Cognitive Psychology* 19 (3): 291–312.
- Pohl, J. 1965. 'Bilinguismes'. *Revue Roumaine de Linguistique* 10 (N-os 1-3): 343–49.
- Poplack, Shana. 2000 (reprint of 1979/80). 'Sometimes I'll Start a Sentence in Spanish Y Termino En Español: Toward a Typology of Code-Switching'. In *The Bilingualism Reader*, 221–56. London, New York: Routledge.
- Poplack, Shana, and David Sankoff. 1984. 'Borrowing: The Synchrony of Integration'. *Linguistics* 22 (1): 99–135.
- Ransdell, Sarah, and Marie-Laure Barbier, eds. 2002. *Studies in Writing*. Vol. 11. New Directions for Research in L2 Writing. Netherlands: Kluwer Academic Publishers.
- Romaine, Suzanne. 1986. 'The Notion of Government as a Constraint on Language Mixing: Some Evidence from the Code-Mixed Compound Verb in Panjabi.' In *Linguistics and Language in Context. The Interdependence of Theory, Data and Application*, edited by Deborah Tannen. Washington: Georgetown University Press.
- . 1989/1995. *Bilingualism*. 2nd ed. Language in Society 13. Oxford, Cambridge, Mass.: Blackwell Publishing.
- . 1994/2000. *Language in Society: An Introduction to Sociolinguistics*. 2nd edition. Oxford University Press.
- Ronjat, Jules. 1913. *Le développement du langage observé chez un enfant bilingue*. Paris: Librairie Ancienne H. Champion.
- Rudin, Ernst. 1996. *Tender Accents of Sound: Spanish in the Chicano Novel in English*. Bilingual Press.
- Sarkar, Mela, and Lise Winer. 2006. 'Multilingual Codeswitching in Quebec Rap: Poetry, Pragmatics and Performativity'. *International Journal of Multilingualism* 3 (3): 173–92.
- Saunders, George. 1988. *Bilingual Children: From Birth to Teens*. Clevedon, Philadelphia: Multilingual Matters.

- Schendl, Herbert. 2012. 'Literacy, Multilingualism and Code-Switching in Early English Written Texts'. In *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*, edited by Mark Sebba, Shahrzad Mahootian, and Carla Jonsson, 27–43. New York, London: Routledge.
- Schneider, Edgar. 2013. 'Written Data Sources'. In *Data Collection in Sociolinguistics: Methods and Applications*, edited by Christine Mallinson, Becky Childs, and Gerard Van Herk. 169–178. Routledge.
- Scotton, Carol Myers. 1983. 'The Negotiation of Identities in Conversation: A Theory of Markedness and Code Choice'. *International Journal of the Sociology of Language* 44: 115–36.
- Sebba, Mark. 2009. 'Sociolinguistic Approaches to Writing Systems Research'. *Writing Systems Research* 1 (1): 35–49.
- . 2012. 'Multilingualism in Written Discourse: An Approach to the Analysis of Multilingual Texts'. *International Journal of Bilingualism* 17 (1): 97–118.
- Sebba, Mark, Shahrzad Mahootian, and Carla Jonsson, eds. 2012. *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*. New York: Routledge.
- Sebba, Mark, and Tony Wootton. 2013. 'We, They and Identity: Sequential versus Identity-Related Explanation in Code-Switching'. In *Code-Switching in Conversation: Language, Interaction and Identity*, edited by Peter Auer, 262–89. Routledge.
- Sommer, Doris, ed. 2003. *Bilingual Games: Some Literary Investigations*. New Directions in Latino American Cultures. New York: Palgrave Macmillan.
- Stell, Gerald. 2015. 'Towards an Integrated Approach to Structural and Conversational Code-Switching through Macrosociolinguistic Factors'. In *Code-Switching between Structural and Sociolinguistic Perspectives*, 117–38. *Linguae & Litterae* 43. Berlin, Boston: De Gruyter.
- Stell, Gerald, and Kofi Yakpo, eds. 2015a. *Code-Switching between Structural and Sociolinguistic Perspectives*. *Linguae & Litterae* 43. Berlin, Boston: De Gruyter.
- . 2015b. 'Elusive or Self-Evident? Looking for Common Ground in Approaches to Code-Switching'. In *Code-Switching between Structural and Sociolinguistic Perspectives*, 1–16. *Linguae & Litterae* 43. Berlin, Boston: De Gruyter.

- Szabó, Csilla. 2006. 'Language Choice in Note-Taking for Consecutive Interpreting'. *Interpreting* 8 (2): 129–47.
- Teng, Huei-Chun. 2011. 'Exploring Note-Taking Strategies of EFL Listenerst'. *Procedia - Social and Behavioral Sciences*, 3rd World Conference on Educational Sciences - 2011, 15: 480–84.
- Thomas, Enlli Môn, and Ineke Mennen, eds. 2014. *Advances in the Study of Bilingualism*. Bristol, Buffalo, Toronto: Multilingual Matters.
- Timm, Lenora. 1975. 'Spanish–English Code-Switching: El Porque´ Y How-Not-To'. *Romance Philology* 28 (4): 473–82.
- . 1978. 'Code Switching in War and Peace'. In *The Fourth LACUS Forum 1977*, edited by Michel Paradis, 236–49. Columbia: Hornbeam Press.
- Treffers-Daller, Jeanine. 1992. 'French-Dutch Codeswitching in Brussels: Social Factors Explaining Its Disappearance'. In *Codeswitching*, edited by Carol Eastman, 143–56. Multilingual Matters 89. Clevedon, Philadelphia, Adelaide: Multilingual Matters Ltd.
- Turner, Kristen Hawley. 2009. 'Flipping the Switch: Code-Switching from Text Speak to Standard English'. *English Journal* 98 (5): 60–65.
- Uzawa, Kozue. 1996. 'Second Language Learners' Processes of L1 Writing, L2 Writing, and Translation from L1 into L2'. *Journal of Second Language Writing* 5 (3): 271–94.
- Valdés-Fallis, Guadalupe. 1976. 'Codeswitching in Bilingual Chicano Poetry'. *Hispania*, no. 59: 877–86.
- Walsh, John, Bernadette O'Rourke, and Hugh Rowland. 2015. *Research Report on New Speakers of Irish*. Dublin: Foras na Gaeilge.
- Walters, Joel. 2005. *Bilingualism : The Sociopragmatic-Psycholinguistic Interface*. Mahwah (NJ), London: Lawrence Erlbaum Associates.
- Wang, Lurong. 2003. 'Switching to First Language among Writers with Different Second-Language Proficiency'. *Journal of Second Language Writing*, no. 12: 347–75.
- Wang, Wenyu, and Qiufang Wen. 2002. 'L1 Use in the L2 Composing Process: An Exploratory Study of 16 Chinese EFL Writers'. *Journal of Second Language Writing*, no. 11: 225–46.
- Wardhaugh, Ronald. 2006. *An Introduction to Sociolinguistics*. 5th edition. Malden (MA), USA: Blackwell Publishing.

- Weinreich, Uriel. 1970. *Languages in Contact: Findings and Problems*. 7th edition. The Hague, Paris: Mouton.
- Weston, Daniel, and Penelope Gardner-Chloros. 2015. 'Mind the Gap: What Code-Switching in Literature Can Teach Us about Code-Switching'. *Language and Literature* 24 (3): 194–212.
- Wetzels, Sandra A. J., Liesbeth Kester, Jeroen J. G. van Merriënboer, and Nick J. Broers. 2011. 'The Influence of Prior Knowledge on the Retrieval-Directed Function of Note Taking in Prior Knowledge Activation'. *British Journal of Educational Psychology* 81 (2): 274–91.
- Woodall, Billy R. 2002. 'Language-Switching: Using the First Language While Writing in a Second Language'. *Journal of Second Language Writing*, no. 11: 7–28.
- Yakpo, Kofi. 2015. 'Code-Switching and Social Change: Convergent Language Mixing in a Multilingual Society'. In *Code-Switching between Structural and Sociolinguistic Perspectives*, 259–87. *Linguae & Litterae* 43. Berlin, Boston: De Gruyter.
- Zentella, Ana Celia. 1997. *Growing up Bilingual: Puerto Rican Children in New York*. Blackwell Publishing.

Web resources

English Oxford Living Dictionary [online], [retrieved 27/07/2016], from http://www.oxforddictionaries.com/fr/definition/anglais_america/utterance

First results of the language, religion and culture survey 2014 – Language [online], website The Federal Council: The portal of the Swiss government, [retrieved 9/11/2016], from <https://www.bfs.admin.ch/bfs/en/home/statistics/population.assetdetail.1000408.html>

Linguistic Characteristics of Canadians [online], website Statistics Canada, 22/12/2015 [retrieved July 2015], from <https://www12.statcan.gc.ca/census-recensement/2011/as-sa/98-314-x/98-314-x2011001-eng.cfm>

Merriam-Webster [online], [retrieved 24/10/2016], from <http://www.merriam-webster.com/dictionary/utterance>

Szablya Consultants, Inc. Translators and interpreters [online], [retrieved July 2015], from www.szablya.com

Webster Dictionary [online], [retrieved 19/09/2016], from <http://www.webster-dictionary.org/definition/bilingual>

Résumé

Introduction

La globalisation, le réseau internet qui donne accès à une multitude de textes multilingues, les exigences croissantes du monde de travail et parfois des raisons d'ordre plus personnel font que les uns et les autres se trouvent de plus en plus souvent au cœur d'un environnement multilingue (que l'on soit immigré, un professionnel qui travaille à l'étranger, un étudiant effectuant un séjour de mobilité, un enfant de couple mixte, etc.). Par la même occasion, nous devons de manière plus régulière nous rapprocher les uns des autres et nous adapter à des conditions changeantes, ce qui explique une explosion dans le nombre de personnes qui apprennent une seconde langue, par exemple.

L'alternance codique est un phénomène linguistique qui est fréquent, voire omniprésent, dans un environnement, comme celui que l'on définit, où au moins deux langues sont en contact. Des personnes bi-/multilingues peuvent produire de multiples alternances codiques à l'intérieur d'une phrase, comme entre deux phrases ou deux tours de conversations. Ce phénomène est de mieux en mieux compris, et de nombreux modèles et approches ont été proposés pour rendre compte de ce phénomène d'alternance codique dans le discours.

L'une des préoccupations des chercheurs en sociolinguistique en générale, et en l'étude de l'alternance codique en particulier, est la manière dont les données bilingues sont collectées. Selon Labov, la linguistique doit chercher à étudier la langue parlée dans des conditions « naturelles » (Labov 1972, 205). Pour cette raison, beaucoup d'études ont été axées sur l'alternance codique dans la communication orale de la vie réelle. Par conséquent, ces différents modèles, basés sur des considérations grammaticales, phonologiques et sociales fort divergentes, ont été testés sur des dizaines de paires de langues, parfois très proches, typologiquement parlant, et parfois très éloignées. Il y a eu un véritable foisonnement d'études, au cours des années.

Mais l'alternance codique n'est pas qu'un phénomène oral. Elle existe bel et bien dans des formes écrites. Avec le progrès technologique, beaucoup de données multilingues écrites sont devenues accessibles pour les chercheurs sur Internet, ce qui a donné lieu à des études sur l'alternance codique dans la correspondance par courriel, les blogs, les journaux en ligne, etc. En outre, l'alternance codique occupe une place spéciale dans les œuvres littéraires, ce qui n'a pas manqué d'attirer

également l'attention de nombreux chercheurs. Différente à la fois de l'alternance codique dans les communications orales et écrites « naturelles », l'alternance codique littéraire a seulement récemment gagné une forme de reconnaissance.

Les travaux dont cette thèse rend compte ont pris leur source dans une observation intéressante que je fis pendant une conférence sur la psycholinguistique, me permettant de m'ouvrir à un champ d'études qui est, pour l'instant en tout cas, très sous-étudié : celui des données bilingues écrites « naturelles ». Celui qui donnait la conférence était un universitaire invité d'origine espagnole. Sa langue maternelle était l'espagnol, le diaporama qui servait de base à sa présentation était en anglais, mais comme nous étions en France et comme il avait déjà salué le public en français, il a décidé de poursuivre son discours en français. Même si le conférencier parlait français, parfois on entendait des phrases ou des mots dans sa langue maternelle. Par exemple, à un moment donné, quand il expliquait quelque chose, tout à coup l'écran avec le diaporama est devenu noir. Sa réaction spontanée a été : « Qu'est-ce qui se passe aquí? ». Au cours de sa présentation, les auditeurs, même ceux qui étaient unilingues, ont été immergés dans un milieu bilingue ou même trilingue. Et c'est là que j'observai une chose intéressante : les notes que je prenais étaient en anglais, français et russe. Je me demandais alors si la nature multilingue de la situation dans laquelle je me trouvais à ce moment-là pouvait expliquer ma propre alternance codique à l'écrit, et surtout pourquoi j'alternais ainsi, alors même que les notes que je prenais n'auraient pas d'autre lecteur que moi-même, à savoir l'auteur même.

Les notes bi/multilingues prises au cours d'une conférence représentent un type de pratique écrite qui diffère à la fois des textes littéraires et de la communication écrite. L'exemple ci-dessus démontre que les locuteurs bi/multilingues utilisent plusieurs langues non seulement pour communiquer des informations à quelqu'un, mais aussi pour noter quelques informations pour eux-mêmes. Une telle activité d'écriture peut être appelée écriture personnelle, parce qu'elle n'implique pas un lecteur externe, mais uniquement son auteur. Ainsi, en ce qui concerne l'écriture personnelle, je commençais à remarquer que les locuteurs bilingues appliquent l'alternance codique quand ils font des listes de tâches ou écrivent des journaux intimes, mes étudiants utilisent plusieurs langues pendant la prise de notes en classe de langue, j'écris en plusieurs langues en préparant mes cours. Ainsi, l'écriture personnelle semble être une riche source de données bilingues. Quelque répandue que soit cette activité d'écriture parmi les bilingues, l'alternance codique dans l'écriture personnelle reste nettement sous-étudiée. Le seul travail sur le journal intime bilingue a été réalisé par Montes-Alcalá (2000), où l'auteur a examiné l'alternance codique dans son propre journal personnel.

Une fois mon attention tournée vers l'écriture personnelle bilingue, je visais à recueillir un corpus qui pourrait offrir la possibilité de faire une analyse qualitative et quantitative de l'alternance codique. Pour plusieurs raisons (dont il sera question dans la deuxième partie de ce travail), les brouillons d'examen des étudiants ont été sélectionnés pour composer trois sous-corpus, à savoir des textes produits par des apprenants français, russes et bulgares en langue seconde. Par conséquent, cette thèse présentera les recherches sur l'alternance codique dans l'écriture personnelle académique et donnera une description multimodale de ce phénomène. Elle s'attèlera également à situer l'alternance codique dans l'écriture personnelle académique sur l'arène bilingue.

La première partie de cette thèse donne un aperçu approfondi de l'état de la littérature dans les domaines idoines, à savoir le bilinguisme, l'alternance codique et la prise de notes. Le chapitre 1 est ainsi consacré à la question du bilinguisme. En plus de passer en revue les définitions et la terminologie employée dans la littérature sur la question, ce chapitre décrit plusieurs typologies de bilinguisme pour plus tard situer les participants du corpus sur le continuum du bilinguisme.

Partie I

Chapitre 1. Bilinguisme

Le bilinguisme est un vaste champ qui fait appel à divers domaines d'études, tels que les neurosciences et de la psychologie des bilingues, l'intelligence artificielle, des études en contact entre les langues, l'éducation bilingue, etc. Le bilinguisme en tant que tel dévoile des questions et des polémiques intéressantes. Ce chapitre va tenter de répondre aux questions suivantes : Qui est une personne bilingue ? Comment les bilingues peuvent différer les uns des autres ? Quels sont les différents niveaux de bilinguisme ?

1.1. Définition

Donnant une image assez générale d'une personne bilingue, les définitions les plus souvent rencontrées dans la littérature s'associe généralement à l'un ou à l'autre des deux pôles d'un spectre, à savoir les maximalistes et les minimalistes. Cette tendance persiste dans des définitions plus détaillées d'un locuteur bilingue. Une interprétation maximaliste pourrait être trouvée dans le dictionnaire Webster de

1961 selon laquelle une personne bilingue est décrite comme « ayant ou utilisant deux langues avec l'aisance caractéristique d'un locuteur natif ; une personne qui utilise deux langues en particulier habituellement et avec une maîtrise semblable à celle d'un locuteur natif »¹⁴ (Hamers et Blanc 1989: 6). L'entrée du dictionnaire Webster (<http://www.webster-dictionary.org/definition/bilingual>) d'aujourd'hui (2015) fait référence à un bilingue comme « une personne qui parle deux langues couramment »¹⁵. Au pôle opposé John Macnamara propose une définition minimaliste en utilisant le terme « bilingue » pour « les personnes qui possèdent au moins une des compétences linguistiques [parler, écrire, écouter ou lire], même à un degré minimal dans leur langue seconde »¹⁶ (Macnamara 1967 : 59–60). Toutes ces définitions manquent de précision.

Les deux Mackey (1962/2000) et Baetens Beardsmore (1982) évoquent la notion de relativisme, montrant l'impossibilité de fournir une définition universelle du bilinguisme. Ils donnent au moins deux raisons qui représente le relativisme.

Tout d'abord, ces définitions ne suggèrent pas la limitation stricte à un certain nombre de langues concernées. Le bilinguisme, et cela sans égard à l'étymologie du terme, devrait être considéré comme un terme qui embrasse tout, y compris les cas de multi/plurilinguisme. « Nous allons donc considérer le bilinguisme comme l'utilisation alternée de deux ou plusieurs langues par le même individu »¹⁷ (Mackey 1962/2000: 27).

Deuxièmement, étant donné que le bilinguisme est perçu comme un concept relatif, il représente un continuum (Hamers et Blanc, 1989). Les bilingues peuvent varier dans leur compétence ; quel est le degré de la maîtrise la langue de la personne ? Les bilingues peuvent utiliser les langues pour des fonctions différentes ; dans quel but la personne utilise-t-elle la langue ? Les bilingues peuvent connaître les mêmes langues, mais les raisons et le détail des commutations peuvent différer ; chaque cas d'alternance codique est différent.

Certains chercheurs (Weinreich 1953/1970; Mackey 1962/2000; Baetens Beardsmore 1982; Hamers et Blanc 1989) préfèrent travailler avec des typologies ou les dimensions d'un phénomène, qui déterminent les limites d'un domaine particulier

¹⁴ La traduction est de moi, comme celles qui suivent. La citation originale est “having or using two languages especially as spoken with the fluency characteristic of a native speaker; a person using two languages especially habitually and with control like that of a native speaker”

¹⁵ “a person who speaks two languages fluently”.

¹⁶ “persons who possess at least one of the language skills [speaking, writing, listening or reading] even to a minimal degree in their second language”

¹⁷ “We shall therefore consider bilingualism as the alternate use of two or more languages by the same individual”

de recherche et qui se concentrent sur un type particulier de phénomène étudié, plutôt que de se contenter de définitions. La question des typologies est l'objet de la section suivante.

1.2. Typologies

1.2.1. La typologie de Hugo Baetens Beardsmore

Une des typologies les plus largement reconnues est celle proposée par Hugo Baetens Beardsmore. La typologie de Baetens Beardsmore est une classification très riche. Il rassemble un grand nombre de différents types de bilinguisme qui sont basés sur des facteurs tels que le temps d'acquisition du langage (précoce ou tardive, le bilinguisme consécutif ou successif), l'environnement de l'acquisition du langage (bilinguisme naturel ou primaire et secondaire), les compétences de la personne (gradation de ambilingualisme, équilibrisme, le bilinguisme « non-fluent » à semilinguisme, réceptif et productif), organisation mentale (subordonné, composé et coordonné), l'effet de L2 sur L1 (additif et soustractif), l'utilisation ou la mise en œuvre du bilinguisme (ascendant et récessif), etc. (Pour plus de détails, voir Baetens Beardsmore 1982).

Cependant, cette typologie a quelques faiblesses. Tout d'abord, il n'y a pas de structure interne. Le classement représente une sorte de liste où différents types de bilinguisme sont présentés plus ou moins au hasard. Certains types de bilinguisme sont des sous-catégories des autres, et en même temps, ils sont liés à l'une des catégories suivantes ; le lecteur est livré à lui-même pour faire ces connexions. Il n'y a pas de division des critères selon lesquels l'auteur distingue différents types de bilinguisme. Cette forme de représentation rend la compréhension de la typologie plus difficile, surtout en tenant compte du nombre des types de bilinguisme représentés. Deuxièmement, certaines explications sont dépassées par des études actuelles et la situation désormais mondialisée de contacts linguistiques.

1.2.2. La typologie de Hamers et Blanc

Dès le début, ces auteurs établissent une distinction entre bilingualité et le bilinguisme. « La bilingualité est l'état psychologique d'une personne qui a accès à plus d'un code linguistique comme moyen de communication sociale »¹⁸ (Hamers et Blanc 1989: 6). Le bilinguisme, quant à lui, est présenté comme un concept plus

¹⁸ La traduction faite par l'auteur de la thèse. La citation originale est "Bilinguality is the psychological state of an individual who has access to more than one linguistic code as a means of social communication"

général qui inclut celle de bilinguisme, « mais se réfère également à l'état d'une communauté linguistique dans laquelle deux langues sont en contact, avec pour résultat que les deux codes peuvent être utilisés dans la même interaction et qu'un certain nombre d'individus sont bilingues »¹⁹ (ibid). La bilinguisme et le bilinguisme tels qu'ils sont définis par Hamers et Blanc correspondent au bilinguisme individuel et sociétal, respectivement. Cependant, les auteurs vont plus loin en fournissant des classifications de chaque phénomène.

1.2.3. Conclusion pour typologies

Les typologies mentionnées ci-dessus décrivent le phénomène du bilinguisme d'une manière tout à fait explicite. Alors qu'il est plutôt difficile de déterminer les contours du bilinguisme par une simple définition, les typologies parviennent à expliquer les différents aspects et facettes de ce phénomène. Ils montrent qu'il n'y a jamais une situation univoque ; chaque cas de bilinguisme peut avoir ses caractéristiques particulières en raison de circonstances survenues dans une communauté ou une expérience bilingue personnelle. Bien sûr, ces deux approches ne sont ni exclusives ni unique.

Cependant, cette thèse n'a pas l'intention d'examiner toutes les définitions et typologies de bilinguisme ; au contraire, elle vise à donner une vue d'ensemble du phénomène du bilinguisme avec ses facettes multidimensionnelles et problématiques en même temps. Les grandes lignes du bilinguisme présentées dans ce chapitre mettent l'accent sur l'importance du relativisme ; à la fois le bilinguisme individuel et sociétal ou bilinguisme et le bilinguisme représentent un continuum. Chaque personne bilingue vit une expérience différente et fait face à un environnement unique où elle se familiarise avec la deuxième, troisième, etc. langue. Et chaque individu trouve sa place dans le continuum de bilinguisme/bilinguisme. Les arguments fournis dans ce chapitre permettent donc d'estimer que les apprenants de la langue seconde, qui sont les sujets de la recherche doctorale présente, sont bel et bilingues. Cette question sera examinée dans la deuxième partie, chapitre 4, qui donne plus de détails sur les participants dans l'étude.

¹⁹ La traduction faite par l'auteur de la thèse. La citation originale est “but refers equally to the state of a linguistic community in which two languages are in contact with the result that two codes can be used in the same interaction and that a number of individuals are bilingual”

Ce chapitre étudie également la question de la terminologie dans le domaine du bilinguisme, en examinant les termes en usage comme « langue maternelle », « première langue », « langue seconde », « langue dominante »²⁰, etc.

Le chapitre 2, quant à lui, examine le phénomène de l'alternance codique à la fois dans ses formes orales et écrites. Section 2.1. donne un aperçu historique de l'alternance codique parlé ainsi que le cadre théorique, qui comprend des modèles grammaticaux, psycholinguistique et sociales de l'alternance codique parlé, et le reste du chapitre explore d'autres facettes du phénomène.

Chapitre 2. Alternance codique

2.1.1. Études fondamentales

2.1.1.1. Aperçu historique de l'alternance codique

Le phénomène de l'alternance codique est connu depuis le début du XXe siècle. En 1913, le linguiste français Jules Ronjat publie ses observations de son fils bilingue - *Le développement du langage observé chez un enfant bilingue*. Ce sont les premières données enregistrées reconnaissables sur le bilinguisme. Cependant, l'essor des études sur l'alternance codique, en linguistique, est assez récent. Au début du XXe siècle, ce sont les psychologues qui ont montré le plus grand intérêt dans le bilinguisme, tentant de décrire le lien entre le bilinguisme et les compétences mentales et les défauts des locuteurs. Ensuite, une place importante dans la recherche sur l'alternance codique est occupée par Uriel Weinreich et son œuvre *Languages in Contact* (1953). Il s'y intéresse aux effets des contacts linguistiques sur les langues. Il décrit trois types de bilinguisme : coordonné, composé et subordonné. « La description par Weinreich des codes de commutation suggère que les individus bilingues possèdent deux variétés linguistiques distinctes, qu'ils emploient (idéalement) à des occasions distinctes »²¹ (Nilep 2006: 5). L'utilisation fréquente de l'alternance codique fut considérée comme un défaut, à comparer aux productions du « bilingue idéal ».

Dans les premiers travaux, l'alternance codique fut caractérisée comme aléatoire, mais maintenant beaucoup de linguistes en viennent à la conclusion qu'il s'agit d'un phénomène régi par des règles ; cependant, il existe encore des divergences sur la nature des contraintes. Dans la seconde moitié du XXe siècle, les

²⁰ "mother tongue", "first language", "second language", "dominant language", etc.

²¹ La traduction faite par l'auteur de la thèse. La citation originale est "Weinreich's description of switching codes suggested that bilingual individuals possess two separate linguistic varieties, which (ideally) they employ on separate occasions"

perspectives sur l'alternance codique changent. Une des œuvres influentes sur l'alternance codique parut en 1972 par Blom et Gumperz. Ils décrivent l'utilisation du bokmål et du ranamål à Hemnesberget, montrant l'importance des facteurs sociaux sur l'utilisation, et plus spécialement la commutation entre ces deux dialectes norvégiens. Blom et Gumperz introduisirent deux types de l'alternance codique ; *métaphorique*, lorsque les locuteurs changent de langue pour produire des effets communicatifs sur l'auditeur, et l'alternance codique *situationnelle*, quand il y a un changement d'une langue en raison d'un changement dans le contexte social.

En 1982 Gumperz développa un autre concept, celui de « l'alternance codique conversationnelle » qui comprend plusieurs fonctions conversationnelles telles que la citation, la spécification du destinataire, l'interjection, la répétition, la qualification du message, et la personnification vs. l'objectivation. Sa recherche inspira beaucoup de linguistes qui ont exploré et développé ses idées par la suite.

Une autre contribution importante à l'étude de l'alternance codique en termes de motivations sociales fut celle de Myers-Scotton, qui a construit sur l'alternance codique situationnelle et métaphorique de Gumperz son propre modèle. En 1983, elle a publié le livre *The negotiation of identities in conversation: a theory of markedness and code choice*, où elle propose le Markedness Model. Selon Myers-Scotton, les locuteurs choisissent les langues qu'ils utilisent pour indiquer leurs rôles sociaux, qu'elle appelle un ensemble de droits et d'obligations. Depuis 1970, il y a eu de nombreuses études démontrant diverses approches, modèles et contraintes de l'alternance codique. La section 2.1.2. fournira une présentation plus complète des principales théories de l'alternance codique.

2.1.1.2. Terminologie et définition

Ce chapitre examine également la question de la terminologie en explorant la différence dans l'orthographe du terme « code-switching »²², son origine et sa distinction de « borrowing », « interference » and « code-mixing ».

2.1.2. Cadre théorique.

Cette section esquisse un cadre théorique avec les aspects grammaticaux, psycholinguistiques et sociaux de l'alternance codique, illustrant différentes contraintes grammaticales : le modèle du Matrix Language Frame élaboré par Myers-Scotton, le principe psycholinguistique de « triggering » introduit par

²² Ici le terme anglais "code-switching" est employé à la place de son équivalent français "alternance codique" pour éviter la confusion et être explicite concernant le contenu de la section.

Michael Clyne, le « we-code » et « they-code » de Gumperz, la distinction entre l'alternance codique situationnelle et métaphorique faite par Gumperz et son collègue Jan-Petter Blom, le modèle de « Markedness » proposé par Carol Myers-Scotton.

2.2. L'alternance codique dans l'écriture

Alors que nous avons tendance à associer l'alternance codique à la seule langue parlée, la section 2.2. se focalise sur l'alternance codique dans l'écriture et démontre les domaines dans lesquels ces mêmes concepts ont été mis en œuvre par la recherche précédente, tels que la production des enfants bilingues et les apprenants d'une langue seconde, la littérature et la communication électronique.

Lorsque l'intérêt pour l'alternance codique écrite surgit, la plupart des chercheurs ont essayé d'appliquer des méthodes développées pour l'alternance codique parlée à leurs corpus écrits, en somme pour voir si les tendances que l'on observe dans l'alternance codique écrite reflètent celles de l'orale. Toutefois, si certaines pratiques d'écriture telles que la correspondance, les scripts de jeu, les communications en ligne, etc. pourraient sans doute être facilement étudiées avec les mêmes outils que l'alternance codique orale, d'autres types d'écriture qui ne sont pas du type communicationnel, ou en tout cas d'un type de communication dans laquelle l'identité de l'interlocuteur est inconnue ou multiple - par exemple, dans la publicité, les affiches, les panneaux signalétiques, etc. - nécessitent une autre approche mettant en évidence la valeur de ces genres littéraires particuliers. Le cadre multimodal proposé par Mark Sebba (2012) est le seul modèle élaboré spécifiquement pour les textes multilingues en tenant compte de leur diversité.

Quand on parle de l'alternance codique écrite, on a tendance à oublier le côté multidisciplinaire de ce phénomène qui intéresse non seulement la sociolinguistique, mais aussi les sciences de l'éducation et la psychologie. Ce chapitre met en lumière les différentes perspectives de l'alternance codique écrite, mettant l'accent sur les cinq thèmes suivants : la lecture/écriture des enfants bilingues, l'écriture en L2, l'alternance codique dans la littérature et l'écriture électronique, et enfin, le chapitre explique la seule approche existante pour l'analyse des textes multilingues.

Chapitre 3. Notes

Chapitre 3 traite de la notion de notes. Le processus de prise de note est examiné comme une activité cognitive.

3.1. *La prise de note*

La prise de notes occupe une place importante dans la vie des étudiants. Voilà pourquoi elle est d'un si grand intérêt pour les éducateurs. Chaque pays a son propre système éducatif avec ses propres caractéristiques. En tenant compte de ces différences, les chercheurs de différents milieux universitaires et de différents systèmes éducatifs ont pu montrer l'influence qu'ils peuvent avoir sur des habitudes et des pratiques de prise de notes. Le processus de la prise de notes peut être défini entre autres facteurs comme une activité cognitive, ce qui implique l'enregistrement sous une forme écrite, provenant d'une source orale ou écrite, des informations que les preneurs de notes estiment importantes (Piolat 2006: 7). Le mot « cognitive » joue un rôle important dans cette définition. L'effort cognitif utilisé dans la prise de note ne doit pas être sous-estimé. Cette activité ne consiste pas seulement à écrire ce qui est dit ou écrit. « La prise de notes est une activité complexe qui exige la compréhension et la sélection des informations et qui inclut des processus de production écrites »²³ (Piolat et al. 2005: 291). Les preneurs de notes doivent sélectionner des informations tout en écoutant, car ils ne disposent pas de temps à perdre avant et pendant la notation des informations.

3.2. *Critique génétique*

Après avoir regardé la prise de notes documentaires (à partir d'un texte lu ou discours écouté), au début de ce chapitre, l'accent se déplace vers un autre type de notes personnelles et qui est un deuxième « compagnon » des étudiants, à savoir les brouillons. Une proposition intéressante a été faite en 1972 par Jean Bellemin-Noël, qui a inventé le terme « avant-texte » :

l'ensemble constitué par les brouillons, les manuscrits, les épreuves, les « variantes », vu sous l'angle de ce qui précède matériellement un ouvrage quand celui-ci est traité comme un texte, et qui peut faire système avec lui (Bellemin-Noël cité dans Grésillon 1994: 108).

L'approche qui étudie de telles séries de documents est connue sous le nom de la critique génétique. Elle n'a rien de commun avec la biologie et de la génétique, au contraire, elle traite de la genèse des œuvres littéraires. L'objet de la critique génétique est un texte comme étant en « mouvement » (Grésillon 2008: 23). Jean-Louis Lebrave (1983) explique la dualité dans les positions que les lecteurs peuvent adopter à l'égard des avant-textes. Pour la majorité des gens, faire face à une copie

²³ La traduction est de moi. La citation originale est "Note-taking is a complex activity that requires comprehension and selection of information and written production processes"

signifie de la lire. Cependant, quand on commence à réaliser cette activité habituelle, on comprend qu'elle est tout à fait différente de la lecture d'un texte ; c'est au cours de cette pseudo-lecture que l'on génère des informations génétiques quasi-spontanément conduisant à des hypothèses sur la manière dont le texte a été produit. Dans le même temps, nous sommes tous des producteurs de brouillons ; cette position d'écrivain nous fait sentir une appartenance énigmatique aux processus génétiques. Ainsi, Lebrave a proposé un modèle d'analyse de la production de textes basé sur ces deux aspects : être à la fois lecteur et écrivain. Fondamentalement, la lecture est une tentative d'organiser un avant-texte dans une double perspective. Tout d'abord, on lit un texte contenant des corrections et des substitutions de l'auteur qu'on a l'intention de mettre dans un ordre chronologique ; ce processus peut être compliqué par un grand nombre de corrections, comme par exemple les avant-textes des auteurs tels que Heine, Proust ou Flaubert, qui sont composés presque exclusivement de ratures. Deuxièmement, on essaye d'établir un ordre évaluatif dans un avant-texte ; allant au-delà de la conventionalité culturelle de la perfection du texte, le système évaluatif et en même temps interprétatif pourrait être une tentative d'expliquer à nous-mêmes pourquoi il y a tant de corrections dans les avant-textes. Ayant procédé du point de vue de la perspective d'un écrivain à celle d'un lecteur, Lebrave met en évidence la différence entre production orale et écrite ; à savoir qu'il est toujours possible de revenir et de rectifier, de supprimer ou d'enrichir, etc. le texte qui est déjà écrit, tandis que dans un discours oral, on peut seulement insérer un commentaire méta-linguistique (« non », « Je suis désolé, je me suis mal exprimé », « Je vais me corriger », etc.). Le fait qui détermine cette différence et l'écriture en général, est que « l'écrit produit quelque chose qui se trouve, non pas dans le temps, mais dans l'espace, et qui est un objet matériel manipulable à deux dimensions » (Lebrave 1983 : 18). Les deux dimensions, que Lebrave mentionne, sont un niveau textuel et graphique d'analyse. Lebrave explique que le niveau textuel fournit une sorte de matériel pour la lecture dans son sens général ; ce niveau correspond à la description des analyses des avant-textes de Grésillon, qui peuvent comprendre la langue (paraphraser, représentation du temps, de la syntaxe, le discours, etc.), la narration, les thèmes, la psychanalyse (l'inconscient du texte) et la contribution de l'auteur vs la contribution de l'époque historique (Grésillon 1994). En ce qui concerne le niveau graphique, il approche des avant-textes des arts visuels avec l'orientation sur l'activité d'écriture de son auteur et non pas sur le jugement d'un brouillon par des lecteurs. L'écriture est considérée comme une activité visuelle, car elle est guidée par le regard, et aussi comme une activité spatiale parce qu'un auteur organise un texte, autrement dit un tracé graphique, dans un espace (Olive et al., 2010: 35) .

3.3. Une étude avec un double corpus

Un dossier génétique assez original a été recueilli et examiné par Annie Piolat et Marie-Laure Barbier (2007). Elles ont réussi à regrouper deux types d'avant-texte : la prise de notes documentaires et des brouillons rédactionnels, qui précèdent le processus d'écriture. Avec cette étude, Piolat et Barbier font remarquer un nouveau champ d'écriture personnelle, qui est riche dans sa diversité des matériaux et qui est encore peu étudié. La direction ouverte par ces auteurs pourrait être très prolifique pour d'autres sciences telles que les sciences de l'éducation, la psychologie cognitive, linguistique, psycholinguistique, etc.

3.4. La prise de note des interprètes

Admettant l'importance de la prise de notes des interprètes dans un domaine général de la prise de notes, la section 3.4. fournit une brève description de cette question.

Après avoir examiné la notion de notes, ses formes et ses contextes ce chapitre finalise la première partie de ce travail et apporte une contribution essentielle à la compréhension complète des brouillons des étudiants, qui constituent le corpus de cette thèse.

Partie II

A la suite de l'état de l'art, qui a été représenté ci-dessus, la partie II se concentrera sur l'étude originale menée dans le cadre de cette thèse de doctorat. Elle examinera l'écriture bilingue, notamment l'alternance codique dans l'écriture personnelle académique. Comme le montre la partie I, le phénomène de l'alternance codique a été bien étudié depuis le début du XXe siècle. Cependant, l'intersection de l'écriture bilingue et la critique génétique n'a jamais fait jamais l'objet de l'intérêt qu'elle mérite. Est-ce simplement un domaine de recherche, qui attend son temps ? J'adhère à cette idée. Les multiples raisons du choix du thème de cette étude, qui seront présentées au chapitre 4, vont avancer des arguments en faveur de son intérêt et son importance.

Le premier chapitre de la partie II explore les raisons du choix du sujet et la méthodologie employée, aussi bien dans les premiers stades de la collecte des

données, les décisions prises quant à la transcription du corpus et la présentation des résultats.

Chapitre 4. Questions méthodologiques

4.1. Le choix du sujet

Comme nous venons de le voir, la majorité de la recherche jusqu'à présent a été axée sur l'alternance codique dans les communications orales et écrites, ce qui implique au moins deux participants : un locuteur/écrivain et un auditeur/lecteur. Très peu de travaux ont porté sur ce qu'on appelle l'écriture personnelle (Montes-Alcalá 2000). La spécificité de cette forme linguistique est qu'il n'y a qu'une seule personne impliquée, parce que l'écrivain joue le rôle du lecteur ; l'écriture personnelle ressemble à un monologue avec soi-même.

C'est un terrain potentiellement riche pour la recherche. Ce type d'écriture est inhérent à chaque individu lettré. C'est une sorte d'un espace personnel où les auteurs se sentent libres de s'exprimer sans aucune instruction ou jugement, qu'il soit moral ou formel. Cet espace présente un grand intérêt pour les linguistes. L'écriture personnelle peut jouer un rôle d'intermédiaire entre le fonctionnement du cerveau et la production finale, ce qui se reflète dans des exemples tels que les autocorrections qui sont présentes et acceptables dans l'écriture personnelle et tous les types d'ébauche qu'on fait en tant qu'esquisse d'une œuvre finale ou d'un plus grand projet. Ainsi, l'écriture personnelle des individus bilingues peut apporter une contribution précieuse à la psycholinguistique, fournissant des données tangibles et naturalistes pour l'analyse de la production bilingue. En outre, les individus forment des communautés et des sociétés. Dans le même temps, ces structures sociales exercent une influence sur les individus, influence qui est incorporée dans les normes sociales et morales, la culture, la politique, etc. Par conséquent, l'individu reflète la communauté et/ou la société, à laquelle il appartient, dans la même manière que l'écriture personnelle reflète l'individu. En conséquence, les questions sociolinguistiques sont nombreuses : Combien l'écriture personnelle est-elle représentative de la communauté/société ? Quelles sont les variables sociales qui affectent l'écriture personnelle ? Est-ce que l'alternance codique dans l'écriture personnelle bilingue varie avec différentes paires de langues ? Aussi paradoxal que cela puisse paraître au premier abord, l'écriture personnelle, comme un sous-type de l'écriture en général, peut être analysée du point de vue sociolinguistique. « [...] [L'] écriture représente une forme particulière de la performance linguistique qui est inévitablement influencée par les conditions sociolinguistiques de sa production, de sorte que les textes de toute nature constituent également un type de source

linguistique qui peut être d'intérêt en soi, vue de la perspective sociolinguistique »²⁴ (Schneider 2013: 170).

L'écriture personnelle bilingue représente tout à fait une source novatrice pour les études bilingues. On doit définir l'écriture personnelle pour permettre son application à la suite de la recherche en matière de bilinguisme. Le présent travail se référera, dorénavant, à l'écriture personnelle comme à des éléments écrits ou des séquences d'éléments écrits produits par un auteur et qui, initialement, ne sont pas destinés à des lecteurs externes.

4.2. Corpus

La nouveauté du travail sur l'alternance codique dans l'écriture personnelle implique la difficulté d'un corpus inexistant. En ce qui concerne tous les types possibles de l'écriture personnelle qui auraient pu être recueillis, les brouillons d'examen des étudiants ont été considérés comme les meilleures données disponibles pour atteindre les objectifs de ce travail.

Les brouillons d'examen ont été recueillis dans trois départements d'anglais respectifs : Université Lumière Lyon 2 (France), Université d'État de Volgograd (Russie) et Université St. Kliment Ohridski de Sofia (Bulgarie). Les étudiants français étaient soit en 3ème ou en 4ème année de leurs études. Les participants russes étaient des étudiants soit en 4ème ou en 5ème année. En ce qui concerne les données bulgares, elles diffèrent en deux aspects : d'abord, seulement 30% des participants étaient en 4ème année d'études, tandis que 70% étaient en leur 1ère année; d'autre part, seulement 44,5% des données sont constituées de brouillons, tandis que les 55,5% restants sont des copies au propre. En dépit du fait que la disparité des données bulgares va compliquer la comparaison entre les trois groupes, ces données méritent d'être étudiées toutes seules.

Étant un adepte, comme nous l'avons vu, du point de vue que le bilinguisme est un concept relatif, qui représente un continuum (Mackey 1962/2000; Baetens Beardsmore 1982; Hamers et Blanc 1989) – une notion qui fut examinée dans la Partie I – je considère les apprenants de la langue seconde, en leur 3ème, 4ème et 5ème année d'études dans le département d'anglais de leurs universités respectives, comme des bilingues aux fins de la présente étude.

²⁴La traduction est de moi. La citation originale est "[...] [W]riting represents a special form of linguistic performance that is unavoidably shaped by sociolinguistic conditions of its production, so texts of whatever kind also constitute a type of linguistic source that may be of interest in its own right from a sociolinguistic perspective"

Toutes les copies ont été recueillies après les examens, qui étaient basés sur un stimulus langagier. On entend par là, d'abord, que la langue de la réponse d'examen était imposée : les français et les étudiants bulgares devaient répondre en anglais, tandis que les étudiants russes avaient certaines tâches à rédiger en anglais et certaines en russe. Deuxièmement, les feuilles contenant les questions d'examen et les textes sources étaient en la langue en laquelle la réponse devait être donnée. Plus de détails sur les examens sont fournis dans la description des données pour chaque groupe linguistique. Le montant total des données collectées est 632 de copies (572 brouillons et 60 copies au propre). Sur ce total, seuls ont été gardés pour l'analyse à proprement parler les brouillons avec la présence de deux ou plusieurs langues, soit 167 brouillons et 12 copies au propre.

Toutes les données ont été collectées soit en format original, manuscrit, soit ont été photographiées si le document original ne pouvait pas être déplacé. Toutes les copies ont fait l'objet d'une lecture minutieuse. Cependant, celles qui étaient entièrement monolingues : anglais, français ou russe, ont été rejetées, et seulement celles dans lesquels plus d'une langue était présente, transcrites. La transcription est un outil d'analyse, ce qui permet un transfert d'une écriture manuscrite vers un espace numérique de travail, c'est-à-dire sur un écran et dans un programme de traitement de texte. Le système de transcription adopté dans cette recherche a cherché à reproduire, le plus fidèlement possible, les copies originales conservant des éléments grapho-spatiaux de l'auteur, les passages raturés, la grammaire et l'orthographe d'origine, etc. Si l'alternance codique apparaissait fréquemment dans la copie, elle était transcrite entièrement. S'il n'y avait que quelques occurrences d'alternance codique, seule l'alternance codique dans son contexte était transcrite.

4.3. Méthodologie

Outre ce travail sur une forme de corpus jusqu'alors inexistant, une autre nouveauté de la présente thèse tire son origine d'une autre difficulté, qui est l'inexistence d'une méthodologie communément acceptés pour l'analyse de ce type de données.

Pour comprendre la raison de l'alternance codique dans les brouillons, une approche multi-perspective est essentielle. La perspective pragmatique montre les fonctions pragmatiques de l'alternance codique ; cependant, celles-ci sont étroitement liées avec la perspective linguistique, qui permet l'examen du schéma de l'alternance codique, qui reflète souvent les fonctions pragmatiques. En outre, une perspective visuelle fournit des informations complémentaires, qui peuvent soit clarifier l'utilisation des fonctions pragmatiques ou indiquer une direction inaperçue.

Dans cette thèse, la perspective visuelle sera utilisée pour démontrer si l'alternance codique est visuellement marquée.

La méthodologie de la perspective pragmatique est basée sur les travaux de Montes-Alcalá (2000), où elle explora les fonctions socio-pragmatiques de l'alternance codique écrite espagnol-anglais. Après avoir adapté cette liste au corpus de l'écriture personnelle, sept fonctions ont été sélectionnées : la citation directe, la clarification et/ou l'élaboration, le commentaire explicatif parenthétique, l'alternance codique déclenchée, l'alternance codique stylistique, le besoin lexical et l'alternance codique sans but particulier. Les sept fonctions mentionnées ci-dessus ne couvrent pas tous les usages pragmatiques de l'alternance codique dans le corpus recueilli ici. Ainsi, sept autres fonctions pragmatiques ont été déterminées et ajoutées pour embrasser les caractéristiques spécifiques de l'écriture dans les brouillons de l'examen. Ces fonctions reflètent les caractéristiques structurales et sémantiques de la rédaction des brouillons, et sont l'encadrement, la structuration, l'alternance codique terminologique, l'économie du langage, la répétition, la liste de vocabulaire et la vérification de l'orthographe.

La troisième perspective est la perspective linguistique, qui examinera l'alternance codique au niveau textuel, ce qui n'est pas forcément chose aisée dans des textes déstructurés comme le sont les brouillons d'examens, mais qui présente néanmoins un intérêt certain, nous le verrons. L'alternance codique peut se produire dans un mot, dans une phrase et entre les phrases voisines.

Les trois chapitres suivants présentent la description du corpus suivant les perspectives visuelles, pragmatiques et linguistiques de l'alternance codique dans les données françaises, russes et bulgares.

Chapitre 5. Les données françaises

Les résultats démontrent que les données françaises sont riches en alternance codique. Elles présentent 13 fonctions pragmatiques parmi lesquelles l'alternance codique sans but particulier, l'alternance codique déclenchée, l'alternance codique stylistique, la clarification et/ou l'élaboration, la structuration et l'encadrement sont les plus fréquentes. Tous les trois types d'alternance codique : intra-/inter-phrastique et à l'intérieur d'un mot, sont illustrés dans les brouillons français, avec une répartition presque égale des types intra-/inter-phrastiques. Un examen plus approfondi des 6 fonctions pragmatiques les plus fréquemment employées à travers la perspective structurelle apporte un soutien à la distribution égale de l'alternance codique intra-/inter-phrastique. Les fonctions pragmatiques de structuration,

l'alternance codique sans but particulier et l'alternance codique déclenchée admettent l'alternance codique intra-/inter-phrastique ; l'alternance codique stylistique et la clarification et/ou l'élaboration suggèrent essentiellement l'alternance codique intraphrastique ; alors que l'encadrement présuppose essentiellement l'alternance codique inter-phrastique. La distribution de ces six fonctions laisse entendre également les motifs de l'alternance codique qui illustrent le fait que l'alternance codique est facilement utilisée par les étudiants français ; les participants changent de langue pour éviter l'interruption dans le processus de l'expression ; d'ailleurs ils appliquent l'alternance codique pour donner des informations plus précises ; et enfin l'alternance codique est utilisée pour structurer ces informations et fournir des marqueurs pour une utilisation plus fonctionnelle du brouillon.

En dépit d'une grande variété de fonctions et des types de l'alternance codique, cette alternance passe inaperçue pour le lecteur. Un peu plus d'un tiers des participants marque visuellement l'alternance codique (45 sur 107); alors que seulement un seul brouillon contient toutes les occurrences de l'alternance codique marquées, sans autre marquage visuel dans le brouillon ; 43 brouillons démontrent que l'alternance codique est partiellement ou entièrement marquée en parallèle avec d'autres contenus du brouillon. Ce ratio signifie que le marquage visuel est utilisé pour mettre l'accent sur des points particuliers de la teneur de la réponse d'examen plutôt que d'attirer l'attention du lecteur sur un changement de langue.

D'ailleurs, une petite enquête, menée dans le cadre de cette thèse, ne montre aucune relation entre l'alternance de codes dans les brouillons et les notes attribuées aux étudiants après la correction de leur copie au propre. La conclusion, que cette description des données françaises suggère, est que les étudiants utilisent l'alternance codique à la fois comme un outil méthodologique et une alternative pour l'auto-expression. Ils utilisent toutes les ressources disponibles, et l'alternance codique est l'une d'entre elles, à préparer le brouillon d'examen de manière efficace avec une structure interne et externe convenable et un contenu explicite.

Chapitre 6. Les données russes

Les données russes comprennent 93 brouillons, qui ont été recueillis après les examens finaux dans le département d'anglais de l'Université d'État de Volgograd (Russie) au cours de la période d'été 2014–2015. La description des données russes montre que malgré le nombre limité de brouillons collectés, les données russes contiennent une variété importante de fonctions pragmatiques (leur nombre s'élève jusqu'à 12). Leur distribution est progressive ; l'occurrence des fonctions est distribuée de 9 à 1. Cependant, du point de vue linguistique, la préférence est

attribuée à l'alternance codique intra-phrastique. L'alternance de codes dans les brouillons russes est plutôt d'un type d'insertion ; généralement un mot ou un groupe de mots est inséré dans un énoncé avec une langue matrice. Cette inférence est soutenue par le fait qu'il n'y a aucune occurrence de l'alternance codique entre la proposition principale et subordonnée, et que la majorité des alternances codiques est de type intra-phrastique.

La moitié des fonctions pragmatiques, dont la fréquence varie entre 9 et 6, est composée de l'économie de la langue, la clarification et/ou l'élaboration, l'alternance codique sans but particulier, la réitération, la liste de vocabulaire et le besoin lexicale. Seule la fonction de l'alternance codique sans but particulier comprend l'alternance intra-/inter-phrastique, tandis que les cinq autres fonctions impliquent la plupart du temps seulement l'alternance intra-phrastique. Ce fait reflète une quantité supérieure d'alternance codique intra-phrastique. Dans le même temps, ces cinq fonctions pragmatiques démontrent que l'alternance de codes constitue un outil pour la composition efficace d'un brouillon ; elle sert à gagner du temps et de l'espace, et d'augmenter la précision de l'information.

La perspective visuelle montre également que l'alternance codique n'est pas marquée en tant que telle, mais constitue une partie du contenu visuellement marqué. Les étudiants russes accentuent les éléments importants de leurs réponses à l'examen soit pour se rappeler les principaux aspects de la réponse ou de se rappeler certaines questions douteuses, qui devraient être revues avant/pendant la composition de la copie au propre. Les brouillons avec une alternance codique visuellement marquée constituent près de la moitié de tous les brouillons bilingues (12 sur 28). La répartition de toute ou une partie de l'alternance codique visuellement marquée (5/7) ainsi que d'autres contenus des brouillons étant visuellement marqués, suggère que l'alternance codique représente un outil dans la composition d'un brouillon d'examen ; à savoir un outil économique appliqué pour les trois dimensions : le temps, l'espace et de qualité.

Chapitre 7. Les données bulgares

Les données bulgares sont les moins homogènes de toutes. La moitié des données est composée des brouillons et une autre de copies au propre. Toutes les données ont été recueillies pendant l'année scolaire 2014–2015. Les résultats montrent que les données bulgares sont assez hétérogènes en raison de la variété des types de copies et l'année d'inscription à l'université. Alors que la majorité des copies au propre a été rédigée par les étudiants en 4e année, les brouillons étaient collectés, soit après un examen d'entrée soit un examen de linguistique des étudiants

en 1ère année ; le niveau de L2 à la fois des candidats et des étudiants en 1ère année peut affecter l'utilisation pragmatique de l'alternance codique. Pour cette raison, la description des données bulgares doit être considérée comme une étude préliminaire, qui donne un premier aperçu de l'alternance codique bulgare–anglais dans l'écriture personnelle académique ; cette étude devrait être développée et réalisée sur un corpus plus grand et plus homogène, qui prendra en compte les limites et les contraintes des données actuelles, pour présenter des conclusions définitives. Cependant, certaines observations méritent d'être mentionnées.

La répartition des fonctions pragmatiques et des types intra-/inter-phrastique de l'alternance codique est déséquilibrée. La fonction pragmatique la plus fréquemment utilisée est l'encadrement avec 43 des 53 occurrences, qui est suivie par l'alternance codique sans but particulier avec seulement 7 occurrences. Dans la mesure où la fonction pragmatique de l'encadrement suggère surtout l'utilisation de l'alternance codique inter-phrastique, le rapport de l'alternance intra-phrastique et inter-phrastique est biaisé en faveur de ce dernier type (4:38, et 2 brouillons contenaient les deux types). D'ailleurs, très peu d'étudiants marquent visuellement l'alternance de codes, mais si elle est marquée, l'accent est mis sur l'énoncé et non sur l'alternance codique en soi. Ces résultats permettent de déduire que les étudiants bulgares, qui participèrent à cette étude, utilisent l'anglais et le bulgare dans des buts différents : tandis que la langue anglaise est utilisée pour élaborer la réponse d'examen, la langue bulgare est utilisée pour fournir l'information externe, qui est sans rapport avec le contenu de la réponse d'examen. Chaque langue a sa propre utilisation délibérée.

Partie III

La partie descriptive est suivie par les discussions analytiques dans la partie III.

Chapitre 8. Comparaison

Le chapitre 8 établit une comparaison interne des résultats des sous-corpus et revoit la possibilité de la comparaison externe.

8.1. Comparaison interne

Section 8.1 effectue la comparaison des données françaises, russes et bulgares sur la base des perspectives visuelles, pragmatiques et linguistiques et propose une comparaison fondée sur le principe de la relation entre la langue et le contenu proposé par Sebba (2012). Les résultats sont considérés à la fois individuellement et en relation avec d'autres résultats.

L'étude montre que les participants utilisent l'alternance codique dans leur écriture personnelle académique quel que soit le groupe linguistique auquel ils appartiennent. Chacun d'entre eux emploient l'alternance codique comme un outil pour rendre leur travail plus efficace ; l'alternance de codes est l'une des options ainsi que L1, l'organisation structurelle du texte, marquage visuel, etc., qui sont disponibles pour les étudiants. Les deux types de l'alternance codique, à savoir intra- et inter-phrastique, sont présents dans le corpus. Cependant, la distribution de variables est fort différente dans les trois sous-corpus, laissant supposer que les participants français, russes et bulgares utilisent l'alternance codique pour des raisons différentes. Toutes les trois perspectives (visuelle, pragmatique et linguistique) ainsi que les relations de la langue du contenu, produisirent les résultats hétérogènes. Alors que tous les participants, comme mentionné ci-dessus, utilisent l'alternance codique comme un outil d'efficacité, il faut se rappeler que l'efficacité est multi-facettes et ici chaque groupe linguistique met l'accent sur un aspect particulier de l'efficacité – ce qui est le plus approprié pour leur rédaction des brouillons d'examen.

8.2. Comparaison externe

Section 8.2 présente une comparaison générale de l'écriture personnelle académique avec la communication orale et la littérature sur la base des fonctions pragmatiques utilisées. Les fonctions pragmatiques ne sont pas comparées directement en raison de leur distribution considérablement variée à la fois dans le discours oral en raison de différents contextes sociaux (Gumperz 1982; Myers-Scotton 1999; Gardner-Chloros 2009; Walters 2005) et dans les textes littéraires à cause des multiples genres littéraires et l'origine de l'œuvre littéraire (Jonsson 2005; Jonsson 2010; Montes-Alcalá 2012). Cette variation est également observée dans le corpus de la thèse actuelle, probablement à cause de la différence dans le degré de la maîtrise de la langue et le contexte culturel des étudiants. La comparaison générale des fonctions pragmatiques dans cette section conclut que les brouillons d'examen représentent une forme d'écriture oralisée.

Chapitre 9. Discussion

Chapitre 9 entreprend une discussion dialectique et indique à la fois les contributions scientifiques et pratiques de la présente recherche.

Conclusion

L'état de l'art présenté dans ce travail a dévoilé une délimitation théorique de l'écriture personnelle bilingue. Après avoir exploré les notions générales du bilinguisme, l'alternance de codes et les notes, on a vu qu'au carrefour entre les trois se trouvent une forme de production bilingue écrite naturalistique, qui reste largement sous-étudiée pour le moment.

Visant à mener à la fois la recherche qualitative et quantitative sur l'alternance codique dans l'écriture personnelle, cette thèse fournit une description multimodale complète de l'alternance codique dans l'écriture personnelle académique. La nouveauté de cette recherche réside à la fois dans son corpus et sa méthodologie. C'est la première étude à recueillir un corpus de brouillons universitaires bilingues. Après la description des données basée sur les perspectives visuelles, pragmatiques et linguistiques, les résultats ont été examinés dans leur interrelation les uns aux autres en soulignant le caractère composite de l'alternance codique. L'alternance de codes représente un outil, que les étudiants utilisent habilement dans leur écriture personnelle académique. Les raisons d'utilisation diffèrent non seulement entre les individus mais aussi entre les sociétés. En ce qui concerne les limitations et les contraintes du présent corpus, des recherches supplémentaires sont nécessaires pour attester ou de contester les résultats. Néanmoins, l'incohérence des résultats met en évidence l'influence de facteurs externes sur l'écriture personnelle. Par conséquent, cette étude ouvre deux perspectives sociolinguistiques. Tout d'abord, d'explorer l'écriture personnelle en tant que témoignage sociolinguistique ; de déterminer les facteurs sociaux qui influent sur l'alternance codique et qui la rendent différente dans l'ensemble des pays/cultures/communautés. Deuxièmement, d'explorer l'écriture personnelle pour avoir un aperçu de la société reçu du côté de ce que j'appellerais volontiers le côté "personnel" caché de la société.

Ce travail a placé l'alternance codique dans l'écriture personnelle académique sur la même ligne que l'alternance codique orale, écrite et littéraire, à la fois en insistant sur ses caractéristiques sociales et apportant la reconnaissance de l'alternance codique dans l'écriture personnelle qu'elle mérite. En outre, cette étude fit valoir la nécessité d'une approche interdisciplinaire pour fournir une

compréhension globale du phénomène de l'alternance codique et de son rôle sur l'arène linguistique.

Appendices

Appendices will comprise two parts. Appendix 1 will demonstrate the primary sources accompanied with the transcription. Appendix 2 will present the marks of the 38 sociolinguistic examination papers with regard to the presence and absence of the code-switching evidence in respective student drafts. Appendix 2 makes reference to Chapter 5.4. Grading.

Appendix 1.

Appendix 1 is intended to be representative of the work which precedes the analysis demonstrated in Part II and III of the present thesis. Four samples of each language group will be presented bellow. The entire corpus (both original student papers and their transcriptions) is available at

https://drive.google.com/drive/folders/0B_C6pnkGJ6jHQmVneThYa0JSMkK?usp=sharing

Each sample contains a scan or a photo of the original draft/fair copy and its transcription. The latter is followed by the table assembling the key information which is necessary to proceed with analysis (such as Language content relationship, Pragmatic function, Linguistic type and Comments). An inscription in pencil indicating a number (Sample X) is added to the original papers by the author of the present work for an easier management of the corpus.

The code-switched passages are italicized in the transcription according to the Matrix and Embedded language of the sample/paragraph. If some lexical items are illegible, they are transcribed with three question marks (???)

1.1. French samples.

Draft 1.

This article entitled "Coop between rhetoric & achievement defined this era" was written by Philip Johnston & published in The Daily Telegraph a British (right-wing?) newspaper on May the 10th of 2007. The title of the article sets the tone of the article & the source or biased.

On 10th July 2007 Tony Blair has been ~~designated~~ is Prime Minister since 1997 & leader of the Labour Party since 1994. ^{only to be} had he in losing the general elect in favour of Gordon Brown, his former Chancellor of the Exchequer.

The author develops the idea that Tony Blair's premiership was many hopes & expectations 10 years ago after the Conservative years of Thatcher & John Major, but that it has ~~retained~~ ~~gone~~ ~~remained~~ or not that positive.

His major political success has been to brought peace in Northern Ireland with the Good Friday Agreement in 1998.

④ eco Gordon Brown.

in ~~the~~ social policies (crime, universities, welfare) }
 democratic & foreign policies (Iraq, Kosovo) }
 democratic policies (budget ...)

(unresolvable left & right)
 Attractiveness of the structure reform
 → need for change to meet challenges of a globalized world.

II) Expectations of New Labour & Third Way.
 → Failure of Conservative & LR policies & gov.
 → Failure of old left, old Labour (James Callaghan)

III) Success of Blair's premiership
 → maintenance success of Gordon Brown → unemployment, already growth...
 → peace in Northern Ireland, secure fuel...
 → **Failure of Blair's premiership** → public spending
 → social policies → foreign policies (Iraq 2003, strong state like Fukuda's when in 1982)
 → democratic policies → democratic policies

→ independence of Bank of England
 → gov in eco
 → number

Free economy ⇒ Thatcherism

Strong state

by 8th month of 2007

expectations

success

failures

→ independence of Bank of England
 → gov in eco
 → number

Free economy ⇒ Thatcherism

Strong state

main ideas of the text:
 The legacy of Blair is contrasted.

Social:
 → crime (ASBOs) → ↑ pop in prison.
 → welfare → failure
 → universities (tuition fees) → debt of students.

Economic:
 → neo-liberalism (control money supply) → good results.
 → Bank of England independent.
 → Private-public partnership → ↑ spending
 Democratic & Foreign policy:
 → Devolve Northern Ireland, Scotland, Wales → not conflict with IRA.
 → War in Iraq & Kosovo.
 → Freedom Act (European → loss of Sovereignty of Parliament)
 → Human Rights Act 1998 → ↑ transparency.
 → House of Lords Act 1999.

2000

Issues:
 → identifier (source, author, date, bias ...)
 → context
 → main ideas
 → phenomena → Thus raises the q...
 → plan (concrete)

To what extent Philip Johnston's criticism of Blair's premiership...
 one could say like Philip Johnston that Blair's era has been defined by a gap between rhetoric & achievement?

Context of 1997 → expect d'1 programme.
 hopes & expectations.

14-15
crisis
sample 2

bildin
 ↳ record
 ↳ gap between rhetoric & achievement defined this era
 ↳ promises
 ↳ results of policies
 ↳ 10th May 2007 → end of the premiership of Tony Blair Prime Minister (97-2007)
 ↳ leader of the Labour party (94-2007)
 ↳ Gordon Brown = Chancellor of the Exchequer
 ↳ Devolved to Northern Ireland, Scotland, Wales ⊕ democratic renewal
 ↳ ↳ great success, major achievement of Tony Blair, his "legacy"
 ↳ Peace in Northern Ireland but Nationalists → influence
 ↳ Good Friday Agreement in 1998 → end "civil war" with IRA
 ↳ Independence of Bank of England
 ↳ Ministerialism by Gordon Brown who reaffirmed Labour's economic programme
 ↳ ⇒ good eco results ≠ from the past
 ↳ Introduction of tuition fees →
 ↳ Schools pb
 ↳ Centralisation & privatisation (like Thatcherism)
 ↳ Private - public projects (schools, hospitals)
 ↳ idea that the private sector is efficient than the public sector
 ↳ Game ↑ no of prison pop ↑ ↳ crime offences → target young people
 ↳ with ASBOs
 ↳ House of Lords Bill in 1999
 ↳ Mayor of London
 ↳ ⇒ Theme of the text = a criticism of Blair's record - How has gov
 ↳ achievement been achieved / rhetoric
 ↳ discrepancy between the 2
 ↳ Blair promised a Third Way, ≠ old left & NR
 ↳ ↳ from old Labour policies

Transcription Draft 1.

Sample 1.

1.

<p><u>Intro:</u></p> <p>-> identificat° (source, author, date, bias...)</p> <p>-> context</p> <p>-> main ideas</p> <p>-> <i>problematiq</i> . Thus raises the q°...</p> <p>-> plan (<i>annonce</i>).</p>

Translation:

problematiq = *problematic* (shortening from "problématique")

annonce = *announcement*

Language content relationship	Disjoint
Pragmatic function	1. Free 2. Parenthetical comment`
Linguistic type	Intra/inter-sentential c-s
Comments	Visual marking

Draft 2.

- | | 97-2001 | 2001-2005 | 2005-2007 |
|---|--|-----------|-----------|
| | Great achievement = Thanks to Blair. | | |
| → | monetary policy under Blair. | | |
| → | New deal for unemploy ⁺ = mean-test for families | | |
| → | his 1 st term was quite good in terms of eco growth and low inflat ^o / a steep fall in unemploy ⁺ of b/w 2% - 3%. | | |
| → | freeze on petrol and diesel duties | | |

What he didn't achieve during his terms in office:

"an avoidance of promised reform ... (p. 15):

- Sth index → 2014 → by referendum
- educat^o: Darling Report.

ASBO: Anti-social behaviour order = crime offences

- high repressive system
- the job Dev introduced = job for the New Lab.

To what extent is Blair.

Blair = 94-2007 = leader of the Lab P^{ty}

97-2007 = PM of the UK

97 GE = a landslide victory for the Lab.

- a few days before his resignation
- 97: Indpdee of the B of Eng.
- 97: a lot of promises / once he came into power = a lot of reforms / continuity of his predecessors' social policies
- HL reform Act 99
- A clear rupture w/ the old Lab
- Good Friday Agreement 98 → a peace settlement in NI.
- The journalist highly criticized Blair's policies / his promises.
- a strong acceptance of neo liberal policies of Thatcherism.
- G. Brown as the new Chancellor of the Exchequer under Blair.
- The Constitutional Reform: After Smith's death, Blair maintained to follow.
- Eco / Social policies under Blair
 - Justice:
 - ↑ pop in jail: old Lab "soft" on crime

Education:
tuition fees.

NI: initial not even good: end of his term

A special relationship w/ US

- Strongly criticized by the public: war in Iraq and Kosovo.

new / and / old / lab

The main ideological f. betw the new and old Lab
 → rewriting of Clause IV in the Lab Constitution, which stated
 that the Lab PTs were committed to nationalising the industries
 of the UK, notably before 1979 GE.
 → ex: of changes made by Blair
 He thought that the PTs were aware to win a GE under
 its current guise. He transformed the PTs from a left
 wing, tax rising, nationalising PTs into 1 which was
 more attractive to across the broad spectrum of the UK's
 social class. Lab shifted from left wing to a centre
 right up in a very short space of time.
 The main way to make the PTs more attractive
 than the Tories was to use marketing and focus
 on techniques to find out what public wanted.

This speech was made by _____, leader of _____ from _____. We would therefore expect it to be rather biased and unlikely to criticize past policy. However as we shall see, he is critical of some of the past policies of his own party. This speech was given on _____, little over _____ after his _____. He begins his speech by highlighting the _____ ... the policies of both the previous policies that the New Lab intended to follow, detailing his key economic, social and constitutional reforms policies.

To what extent is / are ...

This commentary will be divided into 3 parts:

1st we'll explain

2nd we'll examine

3rd we'll argue

Transit: entre les parties:

We'll now analyze the key element of the 3rd way according to Blair, and then attempt to determine whether New Lab has really managed to provide ...

CCU:

In this speech he _____ successfully managed to present the 3rd way. However w/ the exception of devolution there was much continuity b/w the policies of the previous Conservative gov and those of New Lab. Both parties followed similar economic policies which departed from social democratic consensus of the post-war period in order to promote market freedom ...

Act of Scotland



Devolutⁿ: Scot / NI

11/09/98 = 74,3% yes
25,7% no

NI = it was in the interest of the NIth for
The Greater London Authority Act = 99.

Pb: The change in Newlab:

To what extend is Blair's Newlab policies is being down in history?

I - The rupture w/ the Old lab:

II - The new lab

III - The Devolution in Scot and NI:

Last term of Thatcher was a complete disaster and Major

I -

- The Good Friday agreement
- The Blair and Brown = duo
- making new eco policies...
- His 3rd way: as a changing world = renewal of a complete eco policy

II - A continuity w/ the Thatcherite policies:

- social: Blair took some of Major's social policies.

III - Some promises which weren't respected:

- Devolutⁿ in Scot
- NI's case:

CRM1 Sample 2

Transcription Draft 2.

Sample 2.

1.

Transit° : <i>entre les parties</i>

Translation:

Transit° (shortening from "transition"): *between the sections*

Language content relationship	Disjoint
Pragmatic function	1. Clarification and elaboration
Linguistic type	Intra-sentential c-s
Comments	Visual marking

III How culture is a ^{class} distinction today

- 1) history guide (demonstrations, strength... of the M-C & W-C)
- 2) some institutions no longer accessible for the upper class only
- 3) middle class growing & the majority. => talented.
↳ knowledge of rising.

Conclu:

- Going to the opera ... => culture. if access to culture => "smarter"
- Access to culture not always easily accessible => eg. Uni. fees rising.

Sample 3

Transcription Draft .

Sample 3.

1.

Pblmtq: To what extent did the aristocrats^{upper class} have less and less influence on politics and social classes in Britain, which makes a great difference today.

Translation:

Pblmtq = *problematic* (shortening from "problématique")

Language content relationship	Disjoint
Pragmatic function	Structuring
Linguistic type	Intra-sentential c-s
Comments	Visual marking

Draft 4.

PLAW

explain

(a)

Ⓘ Idea of class consciousness / awareness

a) emergence of classness

b) class distinction

} merge

(ZAMORA Claire)

Ⓜ each class = a specific rel to art

a) upper

b) middle

c) working

} class

} west

Ⓜ link between globalization of arts & disappearance of notion of class

a) disappearance of notion

b) ~~cause~~ cause: globalizat° of arts

} blu

T1 End 19th C notion of class-ness appeared
 & reinforced 3 blue classes. ~~FF~~ Events
 in blue position of middle class ~~is~~
 imitating upper class in vain but closely
 related to m. class.

T2 ~~university~~ a century ago culture
 was a way of differentiating
 the 3 classes: the upper class ~~was~~
 being it to down it whereas, it was
 the ~~upper~~ middle class was ~~the~~ source
 of it & the m. class ~~was~~ ~~the~~ ~~source~~
 However, now culture is a way
 of uniting society

Conclusion we can say that
 M.B deals with the idea that ev of
 notion of class has to be linked with the ev of
 of the one of culture.

Culture first a way of opposing 3 different
 class (not the same rel to work).
 The climax of opp being the emergence of class
 -ness (end 19th C). However, now
 that culture has been globalized, notion of
 class appeared to be no longer useful.

Class, History, Culture

(1)

Notes
L 2 §

⑤ a century ago
L 3 last §

"But a century ago"

"today"

"old hierarchy" => 3 classes: upper, middle & working class

"the class of our parents - our grandparents"
L 19th C.

} idea of class absent
} idea "it disappeared end 19th C."
L "losing its center"
= no longer used

idea of a change "as if it class have been"

"we" = British people } he is including himself
-> generalizing phenomenon

what characterize people will = culture

"identifying themselves by their cultural choices"

concrete ex. given to illustrate his words/pt of view

metaphor of a storm "stormed by forces of a brood culture"
L power of culture

"society" ⑤ "individuals" = globalization - seeing ourselves
as a "whole" ⑤ categorization of 19th C.

"labels & designations of the ancient regime"

disappearance of aristocrats & servants.

still used in 19th C.

L being ironical?
L until WWI.

not a ref to king & everything but

3 § Rel class/culture

"the upper class remains - culture"

ex given to illustrate

"a hundred years ago" // "Today"

II a)

not a radical change "even though some of that remains"

L "it has slipped out of"

(1911) > end 19th / beg 20th = class consciousness

idea of belonging to a class (idea, type, well)
of E. Thompson, The Making of the working class

3 established classes - working
- middle classes
- upper

"most important" clearly identifying to one of 3 classes

~~"Ruling class"~~ Role of \pm class on arts

↳ middle class "duplicating" "cultivating"

dominating arts at that time

"big red cheeks" ???

"need money" ref to index value > earnings

who owned it? upper classes > the richest > expensive

↳ who dictated arts = aesthetics

ex of Lady de Grey

vs working class "one and a quarter million domestic servants" = "see"

↳ vs house "designator of the ancient regime"

⇒ 3 classes = 3 \pm ref to arts (vs) everyday liberalization

legitimacy of upper classes "mad way little"

"feels released into the leisure..."

Boundaries?

48 Middle class Relations between 3 classes

Middle class copying upper class

- share of the servants

- "aped the auctors in..."

- each class = main class characteristics
class consciousness > division

middle class as reading class "made" "supervised" = actions

"does" "drives" the origins of

II b)

② laziness of upper class

enumeration = link culture/middle class

"minor but deadly distinctions" idea of boundaries
not pass to go from middle class to upper class
to save money this to money (ed)

I b)

"The in-between class"

"where the class differed

from the aristos"
class"

"fear & loathing of the working

strong boundaries x upper & middle class

x middle & working class = more tangible boundaries

"these men" & reactions in front of
working class = feeling of sup & fear

"a comfort to the establish"

I b

idea that syst in place will be still
= feared to be "away"

categorization ② globalization

① class consciousness in 1900s.

② link @ a different use of class

③ link @ globalizatio of arts & disappearing
of the notion of class.

Transcription Draft 4.

Sample 4.

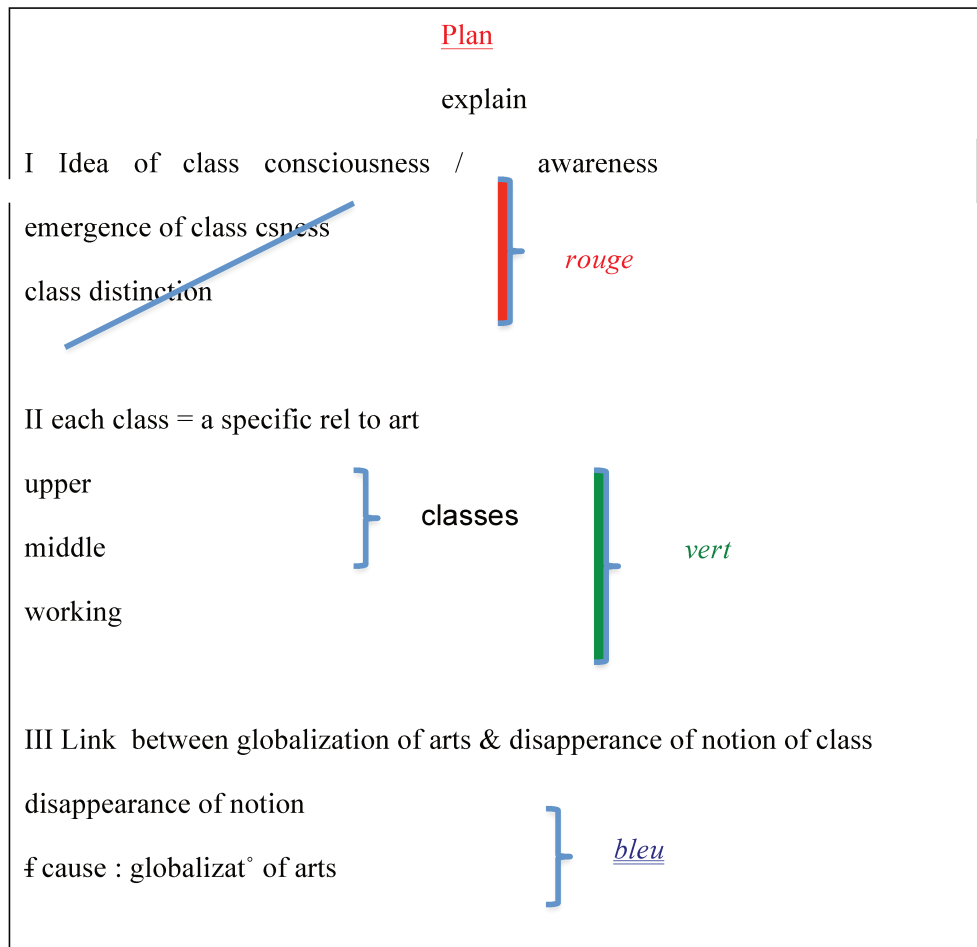
1.

PBQ culture = from a tool to ≠ classes to unifying tool ?

Translation:

PBQ = *problematic* (shortening from "problématique")

2.



Translation:

rouge = red

vert = green

bleu = blue

Language content relationship	Disjoint
Pragmatic function	1. Structuring 2. Framing 3. Framing 4. Framing
Linguistic type	Inter/intra-sentential c-s
Comments	Visual marking

1.2. Russian samples.

Draft TAL 1.

revitalization

The ~~title~~ ^{topic} of the article is — contact . social

The main ~~idea~~ ^{subject} of the article is the new service which was ^{thanks to which} by Yandex ^{because} of which users profiles of the ^{social network} site kontakte.ru can be found through a search engine.

Much attention is given to the ^{of} ~~fact~~ ^{question} of confidentiality of information ~~of users~~ ^{social networks users}.

A brief information is given about Facebook, which settings make it difficult to hide their personal information. ~~from other users~~ ^{these} users can delete their profiles but information remains on a site. Russian analog kontakte.ru also demands truthful data about users and it's difficult to delete an account and change a name.

However, users were not informed that the search engine ^{could} find them. Yandex considers providing ~~the~~ ^{the} most reliable and relevant information as the main objective and ^{the} new service will ~~promote~~ ^{help} to achieve this aim. Now, search engine ^{engineyrobauca} ~~can~~ ^{find} only — information, ^{but} users can address to the support service and hide some data if necessary. At the end of the article the fact that soon deleted accounts will not displayed in ^{complex} search engines is stressed. In my opinion this is very complicated and ~~relevant~~ ^{relevant} ~~for~~ ^{actual} problem. Users should thoroughly study all details and then register.



Transcription Draft TAL 1.

TAL Sample 1.

1.

индексиорванная indexed (marked)

Now search engine can find only _____ information, net users can address toⁱⁿ the support service and hide some data if necessary.

Translation:

индексированная = indexed

Language content relationship	Equivalent
Pragmatic function	Lexical need
Linguistic type	Intra-sentential c-s

Transcription Draft TAL 2.

TAL Sample 2.

1.

<p>Темат. ориентация</p> <ul style="list-style-type: none"> - монотематические (e.g: перевод новое прогноза погоды) - политематические

Translation:

Thematic (shortening from "тематическая") orientation

- monothematic (e.g: translation of a ~~new~~-weather forecast)
- polithematic

2.

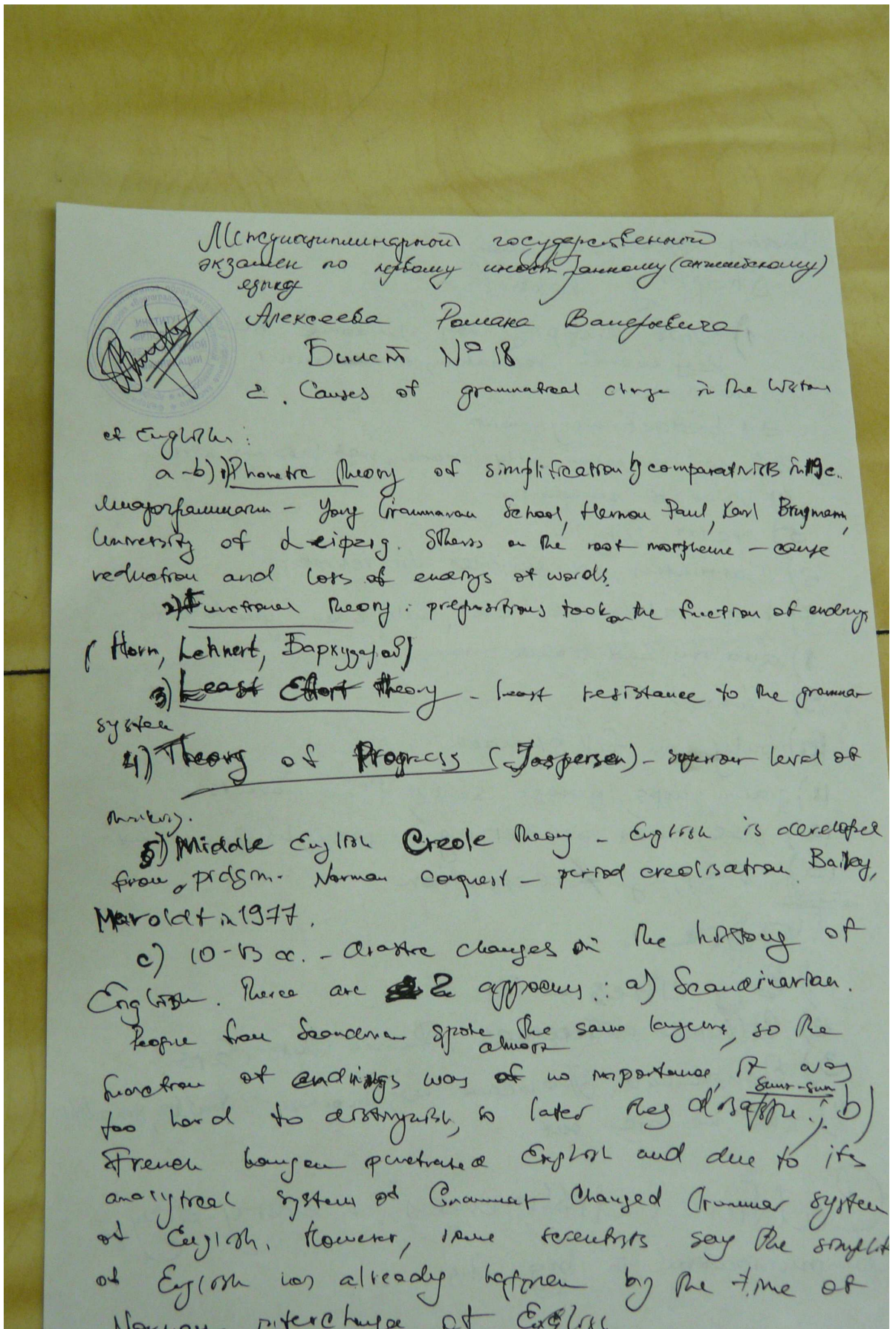
<p>Автоматизир перевод - перевод текстов на ЭВМ с пом. комп.терхнологий. От машинного он отличается тем, что компьютер лишь позволяет ему произвести готовый текст - либо или - работа со словарями, глоссариями, <i>transl. memory</i>.</p>

Translation:

Computer-aided (shortening from "автоматизированный") translation - translation of texts on the computer with the help (shortening from "помощью") of computer (shortening from "компьютерных") software. It differs from the machine translation in a way that a computer only allows to produce a finished text - ~~or~~ work with dictionaries, glossaries, *transl. (shortening from "translation") memory*.

Language content relationship	Disjoint
Pragmatic function	1. Language economy 2. Terminology switch
Linguistic type	Intra-sentential c-s

Draft LINGUISTICS 1.



Transcription Draft LINGUISTICS 1.

LINGUISTS Sample 1.

1.

a-b) 1) Phonetic Theory of simplification of comparatives in 19c.

Младограмматик - Yong Grammarian School, Hernou Paul, Karl Brugmenn,
University of Leipzig.

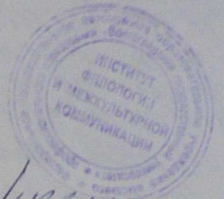
Translation:

Младограмматик = *Neogrammarian*

Language content relationship	Disjoint
Pragmatic function	Terminological switch
Linguistic type	Intra-sentential c-s

смактис — законности разности ~~и~~
 закон ~~и~~ в конкретном к абстрактному.
 "Бителог (по виду)" — труднее. "Иногда (по виду)" —
 переделка. Больше — вайтлогит
грамматик —
терфелогит.
 - ваз. мифологическая часть и мифологическая
 - во всех м. сд. мифологическая часть мифа
 - первая часть мифологическая и мифологическая часть мифа
 мифологическая часть мифа.
Синтаксис. — в каждой к. разная тема. Темы.
 в любом м. к. м. сд. мифологическая часть мифа,
 мифологическая часть мифа.

Myzoreema S.A



What is web-surfing?
Web surfing is a process of jumping from 1 page to another. This way of spending some free time has become popular since internet got not satisfactory speed.

Myzoreema

- web surfing may become uncontrollable and turn into a waste of time.
- reasons: person has completely nothing to be busy with.
- person fails to manage his time.

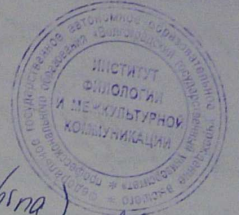
As for me:

- I can't help web-surfing. It happens, when I am bored with my tasks.
 - I search for something interesting just to other sites by links. read news in online news p. and may find interesting topic and start searching the same with videos on video-hosting sites such as You-tube. ~~watch videos on some topic and when it ends - check with from what service can offer.~~
 - I can last pretty long and I look at the time and start doing my tasks in fear.
 - W-s costs me a plenty of time and I often fault myself for doing this.
 - I personally spent very many surfing: 2-3 hours on reading news or online-shopping and watching videos.
 - That's how I relax sometimes. It helps to get out emotionally.
 - agree, that w-s steal my time that I can devote to my study and personal life, but I try to minimize time spending online.
- As I'm experienced web-user and a student of P&AL I know best search techniques and strategies that always apply to in searching for info and doing HW and projects and other.
- According to Internet Digests. com - in 2014 there were 960 mln websites.
- There can be 1000s of web pages connected, connected with the topic quite interested in, but in practice only few

Музыкалинский А.А.
Михновский.

key words:

- communication, language, business context, (writing) English, Chinese
- заграницей
- написание
- Анкетирование
- key words.



Музыкалинский

The article ~~deals~~ ^{simplification of} with ~~English~~ ^{English} variation of English ~~of by Chinese people~~ ^{in a globalized business context.} especially Chinese people.

The author points out that ~~non-native~~ ^{non-native} speakers adopt language to their own comfort, ignoring grammatical and style correctness. He goes on saying that it is important to take into account cultural background while interpreting meaning in a business context. ~~and stressed~~ ^{and stressed} the fact that English absorbs the characteristics and features of different cultures of ~~the world~~ ^{the world}.

~~In this paper the author observes four main compared business emails of Hong Kong and China agents, working in an exporting company and~~ ^{In this study, the author analyzed} the ~~characteristics of~~ ^{influence of non-native speakers on} English and ~~its various features~~ ^{its various features} ~~and~~ ^{and} ~~reflected in~~ ^{reflected in} digital ~~resources~~ ^{resources} comparing through corpus-based analysis of emails by China and Hong Kong employees and founders of ~~exporting companies.~~ ^{exporting companies.}

~~At the end of the article, the author summarizes~~ ^{At the end of the article, the author summarizes} that English tends to be more informal when used by Chinese writers.

Transcription Draft PhAL 1.

PhAL Sample 1.

1. (Photo 1_1)

Словарн статья (-*entry*) – основн структурн ед-ца словаря. [...].

Translation:

Diction. (shortened from dictionary) entry (*entry*) – a basic structural un-
y. (shortened from unity) in a dictionary. [...].

2. (Photo 1_1)

Лексикографич. полита (*gloss*). лексикографич сл-ря [...].

Translation:

Lexicor. (shortened from lexicographic) ??? (shortening from ???)
(*gloss*). in lexicographic dictionary [...].

3. (Photo 1_7)

Задание №1
- Название
- Аннотация
- *Key words*

Translation:

Task №1
- Title
- Annotation
- *Key words*

Language content relationship	Overlapping
Pragmatic functions	1. Reiteration 2. Reiteration 3. Free
Linguistic type	Intra/inter-sentential c-s
Comments	Visual marking

1.3. Bulgarian samples.

1.3.1. Fair copies

Fair copy 1.

Десислава Боранова, I п., 25647

15/.

I.

①.

Edgar A. Poe is seen as one of the first American fictional writers, who embodies typically southern beliefs. A proponent of aristocratic views on society, slavery ~~and~~ the one hand and an opponent of technological advances and urbanization, on the other, Poe's (writing style belongs to the late Romantic period.) Contribution to the emergence of new popular genres is an inspiration to modern writers like Stephen King, as well.

As an answer to the Age of Reason, the Romantic movement that Poe is part of does not concern itself with realistic or deistic ideas but rather with beauty, the arts; and however the writer adds to the list the extraordinary and fantastic.

Another element of interest to Poe is the human psyche and the mechanisms behind it and it is no wonder that the gothic is a part of his repertoire. The fear of being buried alive in a casket, or women

turning into vampires ~~is~~ are themes present in, for example, *The Fall of the House of Usher*. Such fantasy elements are picking up speed in today's works but it was actually Poe, who "gave birth" to the sci-fi and detective novels. The gothic style of writing with its gloomy setting usually in a castle ~~may not~~ be something he invented but rather developed so that other writers could be inspired as he was from his predecessors and contemporaries.

Transcription Fair copy 1.

Fair copy 1.

1.

<i>Десислава X– –, I гр., 12345</i>

Translation:

Desislava X– –, I group (shortening from "група"), 12345

Language content relationship	Disjoint
Pragmatic functions:	Framing
Linguistic type	Inter-sentential c-s

Fair copy 9.

Bunmoy Breab 25644 Group 3

Task 1

20/1

Edgar Poe's contribution to the development of new genres and the evolution of those used earlier on was paramount. Not only a writer of prose and poetry, but an esteemed literary critic as well, he pioneered genres which were new and original to the stage, but they had their roots in what ~~was~~ already ^{been} in use.

It's known, first of all, that Edgar Poe was part of the late Romanticism movement in America. It can be safely said that some of his works fall into that category one way or the other. There are Romantic characteristics in Poe's works no matter how original he made them sound. His characters demonstrated these characteristics the most for not all of them were as original as the others. In "The Fall of the House of Usher", Usher is the prime suspect of harboring Romanticism. He goes through traumatic and life-changing experiences and a sort of a quest. These are known characteristics of a Romantic setting although they do not complete it entirely. The setting itself, the mansion, the climate sound horrifying, but to such an extent that they may well fit into the

1

description of being supernatural. Even though the effect which is socketed into the work isn't really what the reader would expect from a partially Romantic work, ~~even though~~ ^{despite} the picture ~~is~~ ^{being} completed with the death of Usher these features are illustrated in the image of Usher throughout the tale. This connection to art is a good example; his music, his singing they border with the image of a Romantic musician. Usher escapes into a different reality through his music and emotions which are explained, in a way, and sound obscure and odd to an uncomfortable degree. This is actually a part of Poe's style, present in other places in this work and many others. In short, some of his detailed descriptions of what goes on in his works.

In short, sometimes, although an explanatory description of, let's say, emotions, is present, it's written and layed out in such a way that it sounds detailed and complete, but actually leaves the notion that words are all ~~but~~ powerless to truly explain it, that the emotions and horrors are beyond definition and understanding at this point.

Poe also helped pioneer new genres in fiction as well as poetry - he ~~was~~ contributed much to the development of sci-fi and crime fiction

as well as being the chief pioneer of the horror story.

His other works have been described as being fusions between late Romanticism and Gothicism, yet different in his own brilliant way, for example "The Pit and the Pendulum", "The Raven".

It's possible that this great variety of genres used by Poe actually happened, in part, due to his style and ideas on originality and what an original work should look and sound like in terms of long-lasting effect. To him this effect is the soul material for a good work; every word should contribute to a desired effect which should be long-lasting and never before experienced. Because he was so meticulous ~~and~~ ~~to~~ regarding these works [The essay of composition] and true to his word it's no wonder he experimented so much until he had his revolution outlined. With all the works mentioned so far, plus "The tell tale heart" among others, there was practically no real competition that could match Poe's compositions.

Transcription Fair copy 9.

Fair copy 9.

1.

<i>Виктор</i>	<i>X--</i>	12345	Group 3
---------------	------------	-------	---------

Translation:

Victor *X--* 12345 Group 3

Language content relationship	Disjoint
Pragmatic functions:	Framing
Linguistic type	Intra-sentential c-s

1.3.2. Drafts

Draft 5.

Sample 5
 789 11:42
 300-350
 1. Intro 64
 2. For 124
 3. Against 100 83
 4. Conclusion 60

rising need raise
 In addition
 Furthermore
 Moreover

121
 + 83
 + 64

 268
 178

 346

yearly increase
 In addition
 Furthermore
 Moreover

С. П. КО. УЧЕБНО-МЕТОДИЧЕСКИЙ
 ЦЕНТР
 ТЮМЕНСКОГО
 АГРОНОМИЧЕСКОГО

Transcription Draft 5.

Sample 5.

1.

<p>11:42</p> <p style="border: 1px solid blue; border-radius: 50%; padding: 5px; display: inline-block;">300 - 350</p>	<p><i>уседнял</i> <i>застоял</i></p> <p>rising rise raise</p>	<p><i>In addition</i> <i>Furthermore</i> <i>постижения</i></p>
<p>1. Intro 64</p> <p>2. For 121</p> <p>3. Against 100 83</p> <p>4. Conclusion 60</p>	<p style="text-align: right;">121 + 83 + 64 ----- 268 +78 ----- 346</p>	

Translation:

уседнял = settled,

застоял = stagnant,

постижения = achievements.

Language content relationship	Disjoint
Pragmatic functions:	1.Free switching 2. Free switching
Linguistic type	Inter-sentential c-s

Draft 6.

Тема

101

Sample 6

Thousand of selfies (photographs ~~is~~ one has taken of oneself) are uploaded everyday.
 Why have selfies become so popular with young people?

Plan:

Intro:

As the internet became a huge part of our lives in the past decade, so did social networks sharing our everyday life and connecting with other people, regardless if they are acquaintances, or strangers. And thus the selfie was created.

Points:

1. origin - nothing new, born out of necessity, but morphed into something else in a negative way.
2. what it is today and why is it ~~negative~~ easy to turn to something bad, when it was harmless at first. Negative points - ~~overexposure~~ easy access for predators, crippled morals.

1. As we ~~wanted~~ wanted to be more and more connected to the internet and the people on it, we searched for new ways of sharing our experiences. First there was blogging, then came the video sharing, also known as vlogging. And when the social networks became trendy - it was pictures. Of course when we visit a new and exciting place we want to share it via word of mouth, or even better - illustrating it with photographs. The trend started with sharing special moments, beautiful places where we had been, or wished to be. But as the popularity grew, so did the demand to share more and more. It was good fun to compete with your friends, and the need to have more photos and be more popular led to the self taking of pictures. After all you can't always have someone with you to take your picture when you are lounging, shopping, cooking etc.

↓
 What was all fun and games at the beginning, is now considered by many an obsession and a manner of bad taste. There is nothing wrong with taking a quick shot of yourself on top of a mountain for example,

But that easy acceptance of the selfie led to the trends "take a selfie in a bathroom", or "take a selfie ~~care~~ ^{carefully} dressed, or doing explicit things. The oversharing and nudity affect even the youngest of internet users, unless parental control is present. Many of the teenagers feel the need of attention, that their seniors are getting. Also that ^{no} gives them an inflated sense of self-assurance. It can lead to bullying and the so-called "fat-shaming" or "thin shaming". But probably the ~~most~~ ^{worst} consequence is that all of this oversharing gives the sexual predators a way into the lives of their victims. ~~It~~ Internet stalking, identity theft and others are also results of this sudden need for attention.

157
 167
 32/wordy

g

87

Transcription Draft 6.

Sample 6.

1.

<u>Чернова</u>

Translation:

Draft

Language content relationship	Disjoint
Pragmatic functions:	Framing
Linguistic type	Inter-sentential c-s
Comments	Visual marking

Appendix 2.

Table 1. The marks of 4th year French students.

M1 / №	Marks of the drafts WITH c-s	Marks of the drafts WITHOUT c-s
1/ -		6,5
2/ - sample 72	6,5	
3/ - sample 71	11,5	
4/ -		7
5/ - sample 70	9	
6/ - sample 69	14	
7/ - sample 68	14,5	
8/ -		11
9/ - sample 67	14,5	
10/ -		17,5
11/ -		7,5
12/ -		8,5
13/ -		12,5
14/ - sample 66	14	
15/ -		11
16/ -		12,5
17/ -		15
18/ -		13
19/ -		10
20/ -		10
21/ -		8
22/ - sample 65	9	
23/ - sample 64	13,5	

24/ -		12,5
25/ -		12,5
26/ - sample 63	12,5	
27/ -		16
28/ - sample 62	10	

Table 2. The marks of 3^d year French students.

L3 / №	Marks of the drafts WITH c-s	Marks of the drafts WITHOUT c-s
1/ - sample 61	4,7	
2/ -		8,2
3/ -		9,5
4/ -		11,6
5/ -		13,7
6/ -		11,2
7/ -		7,2
8/ -		13,5
9/ -		15,3
10/ -		17,5

