THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 2 - Jean Jaurès

Présentée et soutenue par

Lyanne Silena Del Carmen AHUMADA EBRATT

Le 26 novembre 2020

L'influence translinguistique de la L2 sur la L1 : le cas des extensions sémantiques chez des bilingues tardifs espagnol-français en situation d'immersion

Ecole doctorale : **CLESCO - Comportement, Langage, Education, Socialisation, Cognition**

Spécialité : **Sciences du langage**

Unité de recherche :
Unité de Recherche Interdisciplinaire Octogone-Lordat

Thèse dirigée par
Barbara KÖPKE

Jury

Mme Monika SCHMID, Rapporteure
Mme María MATESANZ DEL BARRIO, Rapporteure
Mme Cyrille GRANGET, Examinatrice
M. Xavier APARICIO, Examinateur
Mme Barbara KöPKE, Directrice de thèse

à Rémi

Para mi Familia

Remerciements

Mes premiers remerciements s'adressent à Monika Schmid et María Matesanz del Barrio pour avoir accepté d'être rapporteures de ce travail de recherche. Un grand Merci également à Cyrille Granget et à Xavier Aparicio d'avoir accepté de faire partie de mon Jury de thèse et d'évaluer mon travail. Vous me faites tous un grand honneur de pouvoir partager mon travail de recherche avec vous.

Je tiens à remercier Barbara Köpke pour ces 6 dernières années d'encadrement de thèse, j'ai eu beaucoup de chance d'avoir un encadrement bienveillant, actif et toujours enrichissant qui m'a permis d'évoluer progressivement en tant que chercheuse. Merci pour ton temps, d'être toujours à l'écoute, de m'encourager, de m'éclairer avec des réflexions et des pistes de recherche. Merci pour ta confiance et de m'avoir donné l'entière liberté de mener ma recherche à ma façon.

Toutes ma reconnaissance à l'Université de Toulouse 2 et l'école Doctorale CLESCO pour m'avoir encadrée pour cette recherche dans des conditions que je n'aurais jamais pu avoir dans mon pays d'origine. Merci à l'Universidad del Atlantico de Barranquilla, Colombie de m'avoir permis d'implémenter mes expériences au sein de notre belle Alma Mater.

Un énorme merci aux participants de l'étude en Colombie et en France d'avoir participé à cette recherche, sans votre participation ce travail n'aurait pas eu lieu.

Aux stagiaires qui ont contribué grandement à la construction des tâches expérimentales: Isabelle, Alex, Victoria, muchas gracias. Aux responsables administratives pour votre travail au sein de notre laboratoire: Evelyne, Marion et Stéphanie.

Derrière ce travail de recherche se trouvent des chercheurs exceptionnels qui m'ont beaucoup apporté aux différentes étapes de la constitution de ma thèse. Je remercie Saïd Jmel pour ton temps et intérêt en ma recherche, pour les nombreuses rencontres qui m'ont permis d'avancer avec mes analyses statistiques. A Lorraine Baqué pour sa disponibilité et ses conseils précieux sur les écrans pour analyser mes données et me convertir en R, tout comme Dr Solier, Dr Garnier, Dra Orihuela et je ne vous remercierai jamais assez.

Je tiens à remercier les docteurs qui ont terminé cette aventure avant moi pour en commencer d'autres, et qui m'ont toujours soutenue au long de mon parcours doctoral, Clara, Laury, Olivier, Franchesca, Adhjar, Émilie. Je vous souhaite beaucoup de succès et que vos vies aient 'toujours des résultats significatifs' !

Une pensée pour Lucille, ma voisine de bureau, je te remercie pour ta patience infinie au tout début de nos réunions « stats » et pour ton soutien du début à la fin, tu es une excellente formatrice ! Une pensée affectueuse pour Kléo et Roge pour la fin de

vos thèses. Un grand merci à mes nombreuses correctrices de thèse et amies: Nour, Mélanie, Anne-Flore, et d'autres que je dois en oublier. Je ne serais pas là aujourd'hui sans le soutien moral de vous toutes, j'ai beaucoup de chance d'avoir pu entamer une relation d'amitié sincère et bienveillante avec vous toutes. J'espère pouvoir être à vos côtés à la fin de votre parcours doctoral et pour fêter vos succès au-delà de la thèse.

Merci à Enhao, Stéphanie, Katia, Marie et à tous les doctorants d'Octogone-Lordat, je vous souhaite beaucoup de succès ! Aux nouvelles venues, je vous souhaite beaucoup de courage.

À mes amies fidèles, Fiorella, Nicoleta, Marlen, Valérie merci pour votre écoute tout au long de ma thèse, et pour toujours être là à mes côtés pour le bon et pour le mauvais, je vous aime.

Une pensée à ceux qui ne sont plus là, À mon Père qui a toujours rêvé de venir en France. À Jorman qui est parti trop tôt mais qui restera toujours dans ma mémoire. À Marie Ange pour son soutien infini, son amour et ses encouragements, une pensée pour la famille Monzo, c'est grâce à Marie Ange que j'ai pu commencer mes études en France, sa lumière restera toujours en moi.

Merci à toute la famille Bégout de m'avoir accueillie comme l'une des vôtres, d'avoir cru en moi et m'avoir encouragée en permanence. Un Grand Merci à Christine pour sa bienveillance inconditionnelle et toutes ses petites attentions.

À Rémi, mon stagiaire - assistant doctorant - compagnon... Merci pour tes beaux tableaux Excel et pour ces nuits blanches dédiées à la thèse.

Merci pour ton soutien infini, pour ta patience, ta tendresse, merci de toujours me faire rire et toutes tes petites attentions, d'avoir toujours cru en moi, de m'avoir grondée quand c'était nécessaire et fait profiter de la vie en "oubliant" la thèse. Je te dédie cette thèse à toi, aujourd'hui un nouveau chapitre de notre vie commence (avec nos compagnons à quatre pattes). À Phil, à toutes ces nuits blanches à mes côtés et à toutes ces fois où tu as marché sur mon clavier pour m'aider à finir ma thèse. Merci de me m'avoir réveillée avec tes ronronnements tous les matins depuis ton arrivée. À Lexy pour sa tendresse, merci d'avoir acceptée de partager Rémi avec moi.

Por último pero no menos importante, gracias a mi familia en Barranquilla, que siempre estuvo « loin des yeux mais près du cœur », muy cerca a pesar de la distancia. Miles de gracias a mi mamá Ledys Ebratt por su amor incondicional, por haberme apoyado en todas mis decisiones, por todas sus lecciones de vida, éste logro es tuyo por todas las tareas que me ayudaste a hacer desde niña hasta la universidad, eres el mejor ejemplo de vida que puedo seguir, te amo. Gracias a mis hermanos Adalberto y Dayson, quienes siempre me apoyaron incondicionalmente y con quienes extraño pasar tiempo. A mi ilustre tío Manuel Ebratt, gracias por su apoyo moral y por sus valiosas reflexiones filosóficas y políticas del país Galo.

TABLE OF CONTENTS

LIST OF TABLES

TABLES OF FIGURES

APPENDIX

INTRODUCTION

At some point we are all concerned by bilingualism. Indeed, the use of more than one language on a daily basis constitutes a regular—rather than exceptional—condition in the worldwide society we live in. For example, we all have been confronted with L2 learning to respond to educational policies or to professional requirements in order to fully integrate a multilingual context. Among other examples, we find linguistic experiences such as spending a semester or year abroad, living temporarily in another country, or spending vacations. These changes of the linguistic environment, in which a foreign language is constantly used, will shape the L2 user's linguistic knowledge very quickly. In this context, the availability of the foreign language plays an important role for language learning and contact. We recognize bilingualism as a natural phenomenon that concerns half of the global population (Grosjean, 2010). We consider that an individual is *bilingual* as soon as he/she uses both languages in different contexts, at different levels, in different situations, independently of language proficiency.

This research focuses on the case of the semantic extensions in L1 by late bilinguals with Spanish (L1) French (L2). Here the term late bilinguals is used to refer to L2 users who had fully acquired their L1 until adolescence or/and early adulthood in an L1 dominant environment and subsequently learned the L2 throughout formal or informal instruction in an L1 or L2 dominant context.

The phenomenon of semantic extension induces an extended transfer of the meaning of a word from one language to a word in another, resulting in a subtle misleading association across languages (Jarvis, 2009). An example, illustrated by Grosjean & Py (1991, p. 58) involves the meaning of the verb '*entendre*' in French, which is extended in L1 (Spanish) in the structure; '*no entiendo el ruido del tren*' (' *I don't understand the noise of the train*'), resulting in a semantic extension, distinct from the canonical production of *'no oigo el ruido del tren' ('I don't hear the noise of the train')*.

Semantic extensions can manifest themselves from L1 to L2 and inversely, this type of crosslinguistic transfer is of great interest because it might be situated through two types of transfer: lemmatic or lexemic. According to Jarvis (2009), semantic extensions are supposed to involve the lemma level (i.e., the knowledge and the semantic and syntactic associations of a word with other words across languages), instead of the lexeme level (i.e., concerning phonological and orthographic information of words).This distinction opposes formal transfer from semantic transfer that we consider to be two sides of the same phenomenon.

Our interest is focused on two psycholinguistic factors playing a role during lexical acces: the neighborhood density of words and the morphological family since of words. These factors may enhance formal or semantic activation levels difference across languages, on one hand, at the lexeme level may observed through neighborhood density of words (e.g., Grainger & Segui 1990), and at the lemma level through morphological family size (Mulder et al., 2013). They are assumed to exert an influence on the way linguistic material is processed in L1 and L2. We are particularly interested in the links

that might exist between these two factors in terms of formal competition levels (e.g., Costa et al., 2007). One of other contributions that we are willing to provide is to try to determine the role of form and semantic links in the production, during word recognition processes, and during comprehension.

In Chapter 1 we describe the theoretical framework involving models of the bilingual mental lexicon and discuss lemmatic and lexemic transfer. Chapter 2 focuses on studies illustrating languages as a dynamic system based on CLI studies. In this Chapter we centered our interest particularly on how to interpret CLI and the specific case of semantic extensions among these. Chapter 3 outlines our research questions, describes the hypotheses proposed, provides a detailed description of the sample population studied, and the linguistic material used as well as the experimental task implemented. Finally, Chapter 4 and 5 correspond to our main results and interpretations to conclude with a general discussion explaining the main findings obtained and their implication for further research.

THEORETICAL FRAMEWORK

1. **The Bilingual Lexicon**

Research on the bilingual lexicon is part of an interdisciplinary field that brings together linguistic, psycholinguistic and neurolinguistic approaches to explore multiples issues to explain the functioning of the bilingual mind. To illustrate our purpose, we will start this Chapter with some definitions and examples of the different approaches with respect to the bilingual lexicon.

First of all, linguistic approaches to bilingualism focus on the description and documentation of 'observable' linguistic phenomena, such as code-switching, defined as the use of two or more languages or variants in the same conversation (Myers-Scotton & Jake, 2005). Studies on code-switching were the first to draw attention to lexical interaction between the two language systems in the bilingual mind. However, codeswitching is also interpreted to fill social and pragmatic functions in bilinguals (Gumperz, & Hernández-Chávez, 1970). In other words, it constitutes a linguistic resource used to indicate the membership in a community or just to emphasize a particular part of discourse for the interlocutor in a particular language. In line with the mental lexicon perspective, Poplack (2005) interprets codeswitching as an indicator of competence in early bilinguals, since this phenomenon is the result of the grammatical overlap between systems that respects complex grammatical constraints in both (or more) languages.

On the other hand, psycholinguistics focus their attention on the levels of representation and the cognitive processes in the brain which preside language processing (Nespoulous, 2004). Regarding bilingualism, SLA (Second Language Acquisition) research emphasized aspects such as the learning environment and age of acquisition (see Lynch, 2017 for a recent review). With respect to the lexicon, special attention has been given to the mechanism involved in learning vocabulary such as *the concreteness of the words*, opposing abstract and concrete words. Concrete words represent tangible ideas or concepts (*eg. table*) that benefit from high imagery resources in the mental lexicon (Paivio & Desrochers, 1980). In contrast, abstract words represent intangibles ideas or concepts that benefit less from imagery resources (e.g., *trust*). Indeed, the word type has an effect on the mental representation, particularly, during retrieval: concrete words are remembered better and are easier to learn than abstract words (Heredia & Brown, 2006; Van Hell & De Groot, 1998). In the same line, cognate words that share the same form and meaning between languages — e.g., table (English)-table (French) — are particularly easy to learn because of their formal and phonological similarities. Indeed word type would have an effect on learning, affecting speed/response time during processing and lexical access, and resistance to language forgetting (Ellis, 2005). Further analysis with regards to this subject is proposed in section 1.1.1.3. The psycholinguistics of bilingualism at the lexical level will be deepened in section 1.1 with the presentation of models and hierarchical representations of procedures that illustrate how languages are organized in the bilingual mind.

Neurolinguistics approaches, finally, study the relation between language and the brain in order to define the neural basis associated with language knowledge and use (e.g in healthy and pathological conditions) (Bambini, 2012). Disciplines such as cognitive

psychology, neuropsychology and cognitive neuroscience work interdisciplinarity along with Neurolinguistics. It also focuses on the storage architecture/organization of languages in the bilingual brain, in particular, on the way the two languages are processed (Meuter, 2009). ERPs (Event Related brain Potentials) and fMRI (Functional Magnetic Resonance Imaging) constitute major methods of research because they allow us to study how the brain behaves during linguistic performance. Thus, it is possible to identify the corresponding underlying cognitive process while a linguistic task is performed. As an example, we can mention the N400, which is a negative ERP component which arises around 400 milliseconds after a stimulus is presented. This measure is sensitive to semantic processing (violation, words in context or isolated words) lexical aspects such as word frequency (Van Petten & Kutas, 1990) and even phonological mismatches (Phillips et al., 2006). Indeed, ERPs provide some evidence that is undetectable with linguistic or psycholinguistic approaches alone.

Considering altogether, the bilingual lexicon has multiples layers and distinct approaches work together to the enlightenment of its structure and functioning. Interference processes (e.g., codeswitching) provide valuable information about interactivity within the lexicon. In order to better understand its complexity, in this section we will first review some models explaining the relations within the bilingual lexicon, and then we will focus on studies of lexical access reviewing methods and findings. A subsection is dedicated to the definition of lemma and lexeme levels that framed this research as well as the transfer types corresponding to each level. Finally, we will focus on crosslinguistic influence at the two formal levels (phonological and lexical) and conclude with the proposition of a perspective focusing on the role of the orthographic/formal and the morphological level in order to explain how crosslinguistic words overlap.

1.1. How Are Languages Organized In The Bilingual Mind?

Mental models or hierarchical mental models allow understanding and making a plausible representation of how cognitive processes unfold in bilingual memory. Such representations of two or more linguistic systems interact in some cases at different levels, for instance, at the conceptual and lexical levels (representing subsequently conceptual links). They can also depict specific linguistics units such as phonemes, words, letters, or formal features. The functioning of a model may follow a 'bottom up' structure, which means that the process of activation begins at the lowest (formal) level represented to arrive at the highest (conceptual) level of representation; in contrast, 'top down' structure proceeds from the highest to the lowest level of representation (Green, 1986). Mental models have allowed researchers so far to validate or question hypotheses explaining bilingual processing and allowing us at the same time to improve and complete previous models in order to have a better representation of the actual processes in the bilingual mind.

This section aims to overview the principal mental models, beginning with the basis that allowed the construction of the current hierarchical bilingual mental models. Our presentation is meant to constitute a general introduction to our study, with this in mind, some concepts are simplified to their main ideas and other models, or concepts are deepened because of their relevance to the current research. Moreover, this section attempts to explore the fundamental theories put forward to explain bilingual memory, and bilingualism modeling. The core of this review focuses on salient aspects that had motivated bilingual modeling, with in the first-place concepts and translation processes as the center of the models, which than turned to word form characteristics playing a role in access to and organization of lexical units.

The main debate in these models is whether linguistic memory is language-specific, shared between languages or both. Furthermore, hierarchical models or stages to structure or represent the underlying processes are presented, according to each theory that depicts differently how languages interact between concepts and lexical representations.

1.1.1. Shared, separate or dual coding theory?

The departure point of research on bilingual memory was the question whether bilingual memory is shared or separated across the languages.

1.1.1.1. Shared Memory Hypothesis

The shared memory hypothesis (or interdependence hypothesis) argues that in the bilingual mind, the two languages are stored in one memory, thus a single meaning would be attributed to two labels or languages, and a tagging mechanism would allow bilinguals

to know to which language belongs a word during retrieval. Following this hypothesis, languages would be stored in memory in the form of language-free concepts (Caramazza & Brones, 1980), in other words, for a Spanish-French bilingual the underlying concept of '*gato'* and '*chat'* would be the same corresponding to a feline that can be tamed as a pet (cat). Additionally, a labeling process provides further information about the language of the word form. Following the Shared Memory Hypothesis bilinguals would store words in terms of semantic features only (López & Young, 1974: 981)

Supporting this theory, López & Young (1974) affirmed that a positive transfer effect of familiarization is observed in Spanish-English bilinguals during a free-recall task[1] Participants were asked to recall lists of words in two conditions: between language (e.g., '*yellow -amarillo'*) and within language (e.g., '*yellow-green'*). Results suggested that both languages shared the same underlying meaning because of the between-language familiarization effect suggesting that common semantic features are activated, thus, in favor of the Shared Hypothesis. This effect is more prominent in the dominant language (English) of the participants of this study.

Following the same principles, Glanzer & Duarte, (1971) attested for repetition effects testing language distance in a free recall task involving Spanish-English bilinguals. Language distance effects were observed in within language repetition condition — when a word is presented and repeated in the same language —, and in between language repetition condition — when the last word presented was followed by its translation equivalent of the other language —. When comparing the proportion of correct recalled words of bilinguals, results pointed out that bilinguals recalled better in the between language condition when distance between the words was shorter. Additionally, the more repetitions exhibited a word, the better the words were remembered (in comparisons to a condition presented without recall). The authors concluded that the two languages show similar ascending patterns, that is, the proportions of recalled words in between language and within language condition, which argues in favor of the Shared Memory Hypothesis (Heredia, 2008).

1.1.1.2. Separate Memory Hypothesis

In opposition to the Shared Memory Hypothesis, the Separate Memory Hypothesis proposes that instead of one, there are two separate independent memory system, one per language. In consequence, this hypothesis claims that one language cannot be accessed through the other, except during translation, which would be the only type of interaction between the two separated systems (see the review of Heredia, 2008). As the previous hypothesis, the Separate Memory Hypothesis concerns mostly semantics. However, this does not mean that representations exclude completely other types of

[1] A free recall task consists of a series of words that are to be recalled by the participant. Depending on the objectives of the particular study, the characteristics of the list of words vary, for example the position of letters, phonemes, or the language.

components such as syntax or phonetics (Glanzer & Duarte, 1971; López & Young, 1974). The following experiments take into account RT (response time) to observe how bilinguals and monolinguals process information in different conditions. This is commonly measured through a Priming Lexical Decision task, in which the participant must decide whether a word exists or not. During this task, primes are associated with targets words, to analyze priming effects, a facilitation effect (when time invested to respond is shorten) in which activation of the target word is facilitated by previous presentation of a prime word when both words are related formally or semantically. The aim of this task is to observe facilitation effects in relation to the stimuli presented, for instance if the target word is preceded by a prime that belongs to the same sematic category, the response is facilitated compared to unrelated words (Heredia, 2008). Interestingly, researchers observe priming effects in two conditions of word pair presentation: between languages or within language. Results showing that there is facilitation for the between language condition are interpreted in line with the Shared Memory Hypothesis because one system would influence the other through facilitation.

In favor of the Separate Memory Hypothesis, Scarborough et al., (1984) suggest that bilinguals are capable of differentiating linguistics codes and that response times are faster when identifying stimuli in the dominant language, at the same time bilinguals block interaction between systems. This suggests that languages might also have separate and independent stores. Furthermore, facilitation effect is observed only for within language conditions, for in between languages condition no facilitation or priming effects were observed, suggesting that each linguistic code is independent.

In the same line, Kolers & Gonzalez (1980) suggest that processing different concepts results in identical retrieval in bilinguals during a synonym repetition task paradigm [2] in between language condition and within language condition. Their conclusion is that concept associations are coded in more than one way and that these associations are coded differently in each language.

As can be observed, there exist data supporting both theories, these contradictory conclusions are related to the multiple experimental designs and theoretical frameworks as well as the different population studied. Considering that both theories are not conclusive, an additional theory may be considered: the dual coding theory.

1.1.1.3. Dual Coding Theory

The Dual Coding Theory (Paivio, 1991; Paivio et al., 1988; Paivio & Desrochers, 1980) proposes two independent but interconnected bilingual memory stores, that is, two verbal systems (specialized in linguistic processing and production), one in L1 and the other in

[2] A synonym repetition task paradigm consists in a free recall task in which translation equivalents are presented in between language conditions, e.g., *church-église*, or in a within language condition in which synonyms of a same language are presented, e.g., *church-chapel* (Paivio et al., 1988, p.163).

L2. The main difference regarding previous theories is that both systems would be linked through translation connections *(*e.g.*, 'boy-niño')* between systems that are stronger and more accessible than bilingual pair associates (e.g., '*girl-niño*'). Besides, this model proposes an image system that participates in the processing of visual information including non-verbal information and events. This image system is connected to the verbal system in L1 and in L2 via the translation connections suggesting that image and verbal systems can influence each other (Heredia, 2008).

In opposition to the Shared and the Separate Memory Hypothesis, the strong point of the Dual Coding Theory is that it argues for interconnectivity in which facilitation during retrieval depends on the double access to the two systems in verbal and nonverbal mode during encoding process, rather than a single code (verbal or nonverbal). A major point of this theory is that it explains the concreteness effect since concrete words benefit from the imagery system while abstracts words don't.

In favor of the Dual Coding Theory, Altarriba & Bauer's (2004) replication study confirms that concrete words are recalled better than abstract words, as concrete words benefit from tangible, verbal and non-verbal information. Additionally, Glanzer & Duarte (1971) attest for crosslinguistic effect in between language repetitions suggesting interconnections between the two systems whose semantic representations are independently activated.

Based on the theories considered here, the debate of Shared vs. Separate and Dual Coding Hypothesis remains inconclusive. At this purpose, Heredia (2008) proposes that task demands may influence the obtained results as well as the different interpretative frameworks adopted. In other words, the same results may be consistent with either theory. Heredia indicates that tasks demand in tasks such as free recall or other tasks evaluating semantic and conceptual knowledge favor the shared memory theory. Conversely, tasks such as the lexical decision task, word fragment completion or lexical naming results (assessing bilingual word forms) produce consistently language-specific evidence in line with the Separate Memory Models.

Given these points, as suggested by Durgunoğlu & Roediger (1987), both theories seem to be accurate for explaining bilingual storage at different levels. Following their interpretation, the semantic level is largely shared and accessible from the two languages and the formal level would be supported by two different and separate lexicons (Heredia, 2008).

1.1.1.4. Conclusion

There are two aspects that draw some limits to each theory. For the Shared Memory Hypothesis, the so called 'tagging mechanism' is too simplistic as it does not concern others linguistic levels, such as phonetic representations. As for the Separate Memory Hypothesis, which suggests that translation is a bridge for accessing two separate systems, the Shared Memory Hypothesis ignores that all linguistic levels interact and

converge during learning a second language, as demonstrated for example by changes in the conceptual representations of motion in more than one language (Brown & Gullberg, 2008; Hohenstein et al., 2004). It can be assumed that when learning new linguistic information, it is easier to compile it instead of creating new nexus for every new information, a shared network would allow more interaction between languages and reinforce previously learned information. In the 70s, (Goggin & Wickens, 1971) proposed the *proactive interference paradigm* which argued that already existing information interferes proactively with the learning of new material (Heredia, 2008). Keeping in mind this perspective, semantics would be an interesting subject of study because of the polysemy of words.

Another fact supporting theories based on the interaction of the different languages is reported by studies on languages in contact. It has been shown that languages interact while the bilingual reinforces his/her linguistic competence in the new language, for example, during retrieval of L2 words, tips of the tongue (TOT as reported by Ecke & Hall, 2013; Poulisse, 2000) that can occur as *codeswitching* or *unintentional language switching*, *borrowing* a word from one language into another (Grosjean, 2010; Myers-Scotton & Jake, 2005; Romaine, 1989) and *blendings of new words or loanword*s (Ringbom, 1987). These types of language transfer will be detailed in section (1.2). As we have seen in the present section, changes of the linguistic competence of the bilingual are generally speaking in line with the shared memory hypothesis, considering that language facilitation or transfer in-between languages occurs in L1 to L2 direction and in L2 to L1 direction (Chapter 2). In the next section, one of the most influent positions regarding bilingualism is going to be detailed to arrive progressively at more specific modeling of the organization of the bilingual lexicon.

1.1.2. ***Weinreich's*** Perspective

In the field of bilingualism research, Weinreich (1953/1968) was a pioneer who focused on types of bilingualism. He proposed three different ways of constructing linguistic representations following a perspective linguistically centered in the study of language contact. He described interference as *deviation* from standard norms at the phonetic, grammatical and lexical level as the result of the knowledge of more than one language, which implies rearrangement of linguistic units between languages (Weinreich, 1979). The organization of knowledge he proposes is centered on translation pairs as shown in **Figure 1** and relates to the Saussurean terms of *signifier* and *signified* [3],replaced here by form representation and conceptual representations by de Groot (2013, p. 173) distinguishing:

[3] Of special interest is the attention Weinreich paid to the sign. The sign is a Saussurian notion referring to the combination of a unit of expression (the so-called signifier) and a unit of content (also known as the signified)»(Woutersen et al., 1994: 448)

1. *Coordinative,* in which the L1 word and the L2 translation is supposed to have separate formal and conceptual representations.
2. *Compound* in which the L1 words and the L2 translations would share the same conceptual representation and have separate representation of word forms.
3. *Subordinative*, for which the L2 form representation does not map with the equivalent conceptual representation. In other words, the access to the L2 is indirect and available only via the L1. Thus, the L2 form representations are accessed through the L1 form representation including also its conceptual representation.

Weinreich proposed that a person or a group is not necessary categorized as compound or coordinative, instead some words could be considered compound and others coordinative, for instance (Weinreich, 1979: 10). Later interpretations assumed the three types of bilingualism (described above) as dependent of the context of language acquisition or learning (Ervins & Osgood, 1954).

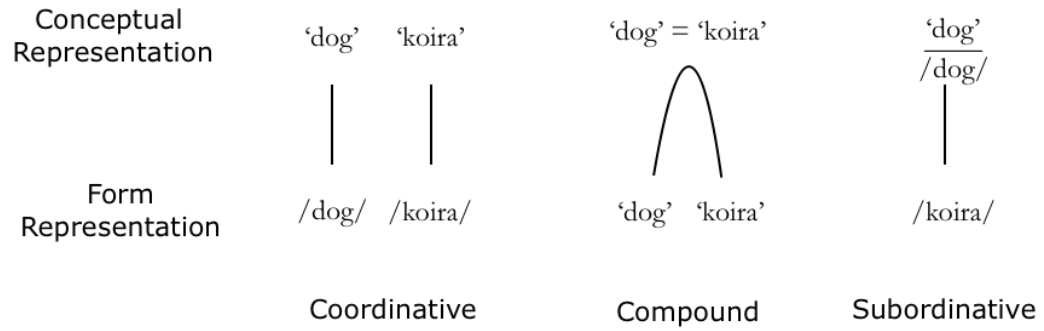

Figure 1 Organization of the bilingual memory as proposed by Weinreich (1953/1968) in (Groot, 2013, p. 173)

Some authors have extrapolated that compound bilinguals correspond to early bilinguals and to a shared bilingual memory; in contrast, the coordinative bilingual would correspond to late bilinguals and the separated memory hypothesis (De Groot, 2013). Furthermore, improvement of linguistic competence has been proposed to be a transition from subordinative to coordinative bilingualism. Weinreich's perspective constitutes a first step for a deeper understanding of bilingual processes in future models or proposals in the field. Weinreich further suggests that this transition between subordinative to coordinative bilingualism during language learning constitute an issue of interest of psycholinguistics (Weinreich, 1979, p.11).

Nevertheless, this idea was not taken up until 1994 when Woutersen et al. (1994) included this transition in a psycholinguistic approach associating the lemma and the lexeme level following the model of (Levelt, 1993). Briefly, the Levelt's model of monolingual speech production describes the processing from the pre-verbal message until the parsed speech. This model is divided into different stages of message generation at micro-planning and macro-planning levels through a processing system called *the conceptualizer*, then language encoding at the grammatical and phonological levels through *the formulator* and then language *articulation*. The center of these processes is articulated by

lemmas and *lexemes*. They constitute the mental lexicon and include conceptual and formal information of words allowing word activation in networks of connected nodes.

Woutersen et al. (1994) adapted the lexico-semantic part of the model to illustrate the bilingual mental lexicon with reference to Weinreich's three types of bilingualism as depicted in **Figure 2**.

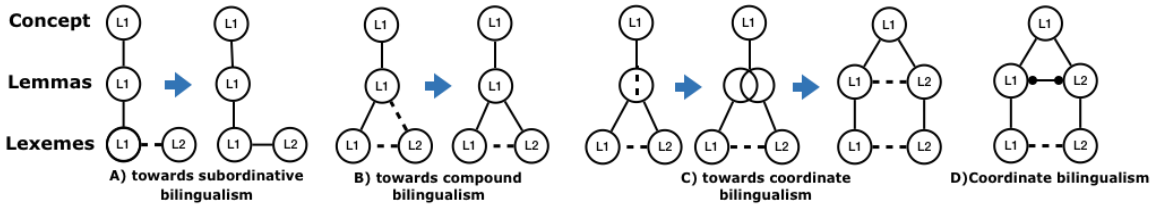

Figure 2 An adaptation of the bilingual modeling proposed by Woutersen et al., (1994, p. 464–466)

The subordinative bilingualism is represented in **Figure 2**A in which links between the L1 and the L2 are reinforcing progressively at the lexeme level depicted at the beginning with broken lines, and then with a solid line. In **Figure 2**B in compound bilingualism a new link is constructed between the L1 lemmas and the L2 lexemes, the reinforcement of which will give rise to coordinative bilingualism. In **Figure 2**C in which the lemma is finally split into two independent units as a result of two interconnected languages at the lemma and the lexeme level. Finally, coordinative bilingualism is represented with an inhibitory link between the L1 and the L2 lemmas whose function is preventing activation from the conceptual level. As can be seen, following the model of Woutersen et al., (1994), concepts are not language-specific which enables a richer perspective through a connectionist framework of bilingualism that considers each type of bilingualism as a developmental stage during the construction of links between lemmas and lexemes between two languages (Köpke, 2009).

1.1.3. Word Association Model and Concept Mediation Model

The Word Association Model and the Concept Mediation Model constituted the fundamental basis for the next models that were proposed later in the literature. The common point of these models is that they are based on two levels of representation, the conceptual level tagged as 'concepts' and the lexical level that is divided in two language-specific lexicons (Heredia, 2008). Besides, both models organized the bilingual lexicon as a whole, suggesting that all structures correspond to the same type, and do not only focus on translation pairs as did Weinreich (1979). The conceptual store is shared between both systems but the access to those common concepts differs depending on the model.

The Word Association Model (Potter et al., 1984) assumes that lexical levels between languages interact in translation processes, however the access to the conceptual

level is only possible through the L1. This means that the conceptual level is linked bidirectionally exclusively with the L1 as depicted in **Figure 3**. Thus, to access conceptual information from the L2, the activation of the translation equivalent in L1 is necessary. On the other hand, in the Concept Mediation Model (Potter et al., 1984) both lexicons interact and are independently able to access to the conceptual level without passing through the filter of translation as the Word Association Model does. In other words, the L2 is linked directly to the conceptual representation that is supposed to be neutral or independent of the language (this model is restricted regarding futures models that account for concepts that are language specific cf. Modified Hierarchical Model).

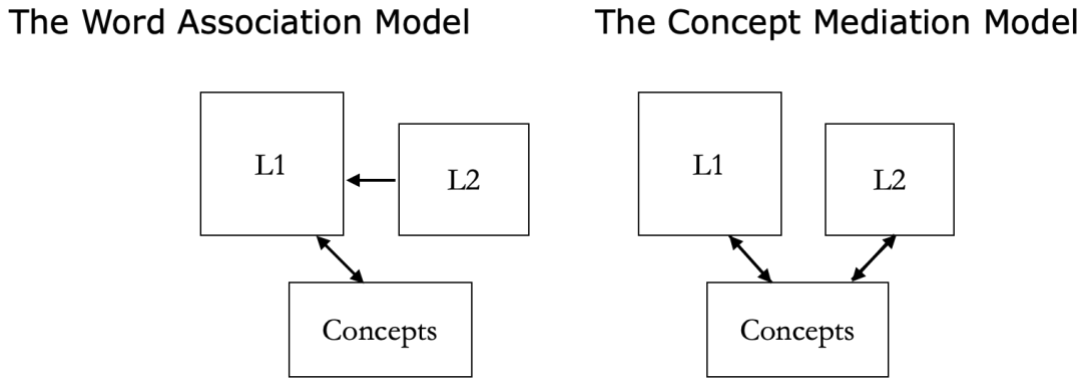

Figure 3 The Word Association Model and the Concept Mediation Model (Potter et al.,1984) in Heredia (2008, p. 53)

The Word Association Model and the Concept Mediation Model can be associated to different stages of second language learning, as discussed earlier.–Later research (de Groot & Hoeks, 1995) suggests that L2 learners would rely on words association processes in early learning while concept mediation processes would correspond to more fluent or advanced stages of second language acquisition. Considering altogether, the beginning of second language learning is supposed to start by associating L2 words with the L1 (Word Association Process). Both, Weinreich (1979) and Potter et al., (1984) had in common the hypothesis that there is a stage in L2 learning in which association processes are made. This stage is the subordinative type proposed by Weinreich. When L2 competence improves, stronger connections are made involving the construction of compound structures, which are direct links between L2 lexical representations and conceptual representations (Concept Mediation Process). In other words, the Concept Mediation Processes would correspond to a high proficiency level in second language acquisition. In the same line, at a superior stage, L1 to L2 correspondence increases, and conceptual representations become progressively shared between languages. This would correspond to the passage from Word Association Processes to Concept Mediation (as in Potter et al. 1984), and from Weinreich's subordinative to compound lexicon organization (1953, following de Groot, 2013, p. 173–174).

The models proposed by Potter et al. account for the assumption that vocabulary size differs between languages arguing that L1 vocabulary is larger than the one of L2. Note that this is also the reason why the squares representing languages are bigger for L1 than L2 in the models represented in **Figure 3**.

Now, the question arises at which stage these links emerge and what are the factors that trigger the connections between formal and conceptual representations? It has been suggested that word use frequency facilitates direct links between word forms and conceptual representations. Bilingual memory structures having direct access between form representations and conceptual representations are frequent L2 words. On this matter, de Groot affirms during her review of these models and based on her own research that:

> *"At a given stage of L2 fluency a bilingual has developed more direct connections between the form representations of frequently used L2 words and the corresponding conceptual representations than between the form representation of infrequently used L2 words and their conceptual representation"* (de Groot 2008, p. 175).

She highlights the importance of taking into account frequency that certainly plays an important role in processing language implying that increase in L2 fluency is related to changes of the linkage patterns of word form representations and conceptual representations. Furthermore, word frequency is one of the factors that triggers direct linkage or access between form representations and conceptual representations.

In sum, regarding The Word Association Model and the Concept Mediation Model (Potter et al.,1984), it is important to mention that language learning seems to equal constant changes; second language improvement would be reflected in the way bilinguals link formal linguistic information into concepts or meanings. Additionally, word frequency, language use and the context involved in learning (simultaneously or independently) plays a major role in constructing the giant puzzle that is bilingualism. Once the learning of second language processes begins, the simple fact of using a new word in a new language on a daily basis would affect the mental lexicon to the point of rearranging/reorganizing lexical connections that might never be the same as before, just as a bilingual cannot avoid to understand or be able to forget completely a language that was already learned.

1.1.4. The Revised Hierarchical Model (RHM)

Some years later, Kroll & Stewart (1994) proposed the Revised Hierarchical Model (RHM), inspired by the previous models and combining them. Brysbaert & Duyck (2010) suggest that the RHM enhanced the preceding models by proposing the inclusion of shared and separate representations in their modeling of bilingual language processing. This suggests that bilinguals can store information in two different stores, but shared representations may also exist. Moreover, the RHM proposes a selective access suggesting

that bilinguals are able to inhibit or activate their languages, in order to prevent language interference.

Activation refers to the availability of a representation in the lexicon. For example, before a speaker names an object in a specific language, identification of the object '*activates*' its related conceptual representations. However, since multiple conceptual representations may be activated simultaneously, a concept selection process is necessary, leading to lexical selection and finally to word articulation (La Heij, 2005). More specifically, when activation starts at the semantic level (e.g., *'feline')* then, it is extended to the lexical level in which a lexical node is selected, hence, the most available lexical element is chosen among others (e.g., *'cat, panther, animal'*). Once the lexical node is selected (e.g., *cat)* the activation flow spreads to the next level, i.e., the phonological level (e.g., /k/; /ae/; /t/).

Furthermore, the RHM added two unidirectional links at the lexical level between the L1 and the L2 and vice versa (as observed **Figure 4**) instead of one bidirectional links as in previous models (see **Figure 3**). The bidirectional links of The Word Association Model and the Concept Mediation Model are depicted differently in the RHM, they vary in strength and represent translation routes or asymmetries between L1 to L2 processing (de Groot, 2013). It should be noted that the RHM is obtained from experiments involving translations from the L1 to the L2 and vice versa (Heredia, 2008).

Following this model, semantic information would be better accessed from L1 arguing that category interference [4]occurred when translating from L1 to L2 (referred to as forward translation) suggesting that L1 to L2 translations are conceptually mediated, so, translation takes longer than L2 to L1 translation. In contrast, L2 to L1 translation trace is stronger because of direct connections between L2 form representations and L1 form representations (de Groot, 2013). For example, for a native French speaker it would be easier to translate '*mouse*' into '*souris*' (L2 → L1) than translate '*souris*' to English (L1 → L2) because L2 words are mapped onto the L1 word equivalents, but not every L1 word is mapped on the L2. This is supposed to be due to a lack of L1 to L2 translation practice. This kind of lexical link is represented by a broken bidirectional line as shown in **Figure 4** (Heredia, 2008).

[4] Interference observed between a list of semantically related words, for example in picture naming: the image of a *shirt* being easier accessed when preceded by a semantic related stimulus such as *jacket* or *belt* (clothing category).

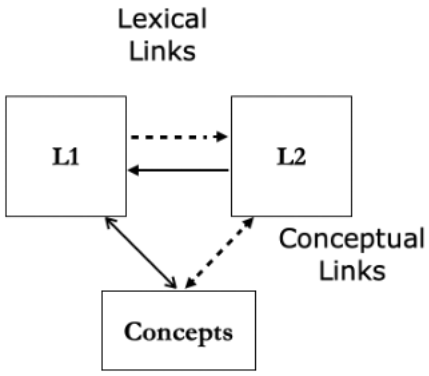

Figure 4 The revised hierarchical model (Adapted from Kroll & Stewart, 1994, p. 158) in Heredia (2008, p. 55)

Regarding the conceptual links between the L1 and L2 lexicon and the conceptual store, in Figure 4 the L1 lexicon is depicted by a solid line, representing strong conceptual links, while the link with the L2 lexicon is depicted by a broken line showing weak conceptual links, and is independent of the level of L2 proficiency of the bilinguals. This difference corresponds to the idea that the L1 is the native language and that L1 meanings are (and remain) more familiar than L2 meanings.

Hence, the RHM suggests that translations are faster and induce less category interference from L2 to L1, consequently, lexical associations play a more important role in translation than concept mediation (de Groot, 2013). In other words, L2 to L1 translation does not involve the conceptual store. Studies in favor of this model report priming effects or facilitation only in L2 to L1 direction (Keatley et al., 1994).

In contrast, other studies (de Groot & Poot, 1997; Heij et al., 1996) have found some counterevidence to the RHM showing the opposite pattern, that is, shorter time in forward translation (L1 → L2) in bilinguals non-fluent in L2 (de Groot, 2013). These results are against the hypothesis that backward translation does not involve the conceptual store.

Another problematic issue of the RHM is that the L1 is supposed to be always dominant over the L2, independently of the level of proficiency of the bilinguals. As proposed by Pavlenko (2000), the rapidity of processing depends on the frequency of use of that language. Considering this, the more dominant language in terms of activation is not necessarily the L1 or the language acquired first. On the contrary, for the language used the most frequently, lexical access is facilitated, and it can be considered as the dominant language. For example, if a French (L2) Spanish (L1) bilingual uses more frequently his L2 than the L1, the L1 will no longer have a major incidence on lexical retrieval compared to the L2.

As can be seen here, the dynamics of the lexicon as suggested by Pavlenko (2000) is evidenced by the fact that the L2 influence the L1, which corresponds to an important developmental stage during second language acquisition. Hence, language status (in terms

of dominance) is not static, it changes over time and depends on individuals, as well as the language involved.

1.1.5. The Distributed Conceptual Feature Model (DCFM)

The Distributed Conceptual Feature Model of de Groot (1992) deepens the difference between the lexical and the conceptual levels. The lexical level concerns word forms, which are composed of the phonological and orthographic representations of the words in L1 and L2. The conceptual level concerns the meanings of the words, which are language independent and are constructed by a collection of features that are activated by word-forms.

The main contribution of this model is the distinction between concrete vs. abstract and cognates vs. non-cognate words and their representation in the mental lexicon. Cognate translation is associated with facilitatory effects, that is, conceptual overlap of cognates (e.g., *'hospital* 'and *'hospital'* in **Figure 5**C) would affect bilingual performance, as they are translated more quickly than non-cognates (Kroll & Stewart, 1994).

The model focuses on the number of conceptual features that are shared across languages depending on the type of word and postulates connections between the lexical and the conceptual nodes corresponding to a word. This conception of the mental lexicon suggests that concrete words (detailed in **Figure 5**a) share all conceptual features across languages (e.g., '*casa*' and its L2 translation '*house')*. Since concrete words involve perceptual referents that are mostly shared across languages and, consequently, have similar or identical conceptual features (e.g., here: inanimate object defined as a building for human habitation) and hence a close translation equivalent across languages (Kroll & De Groot, 1997).

The predominant idea is that concrete words are thought to share many features across languages, while abstract words are more language specific. It also suggests that cross-linguistic priming effects can be explained by the number of overlapping features between the prime and the target. This is the case in the example in **Figure 5**a in which both lexical and conceptual levels overlap in L1 and L2. In contrast, for abstract words such as in **Figure 5**b '*amor*' and '*love'* only a few conceptual features are shared. In fact, in English, the verb *love* can be used with animate and inanimate objects (e.g., '*I love this song*' vs '*I love my husband*'), while in Spanish the verb '*gustar*' ('*like')* is restricted to inanimated objects (e.g., '*me gusta ésta canción')* and *love* is used exclusively for animate objects (e.g., '*amo a mis esposo')*[5].

[5] For further information see the illustration proposed by (Heredia & Brown, 2006, p. 240) in **Figure 14**.

This example can also be interpreted in terms of the density of the conceptual features of animate categories. Sholl et al. (1995) showed with picture naming and translation tasks that animates are more rapidly translated than inanimate concepts. In contrast, non-cognates would be similar to abstract words.

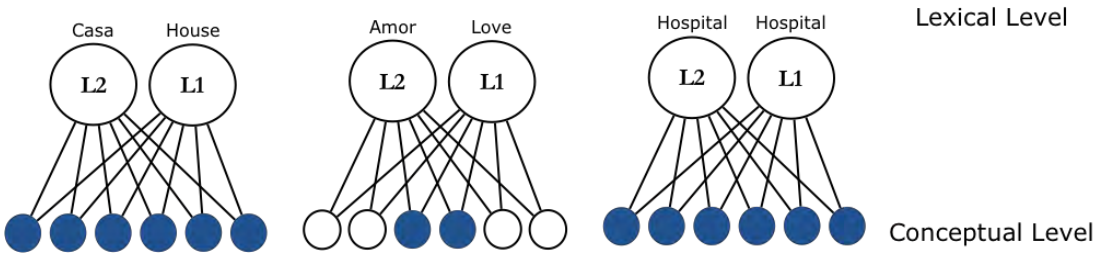

Figure 5 The Distributed Feature Model of bilingual memory for concrete (5a), abstract words (5b) and cognates (5c) (Adapted from de Groot, 1992 by Heredia (2008, p. 58)

Differences in response patters for these two types of words are observed in various research paradigms including word translation, between language priming and words associations between languages (see De Groot, 2013).

One of the elements that support this model is the concreteness effect (discussed in section 1.1.1.3) showing that concrete words are recognized and translated faster than abstract words (Van Hell & De Groot, 1998). Concrete words would benefit from a denser imagery system than abstract words, likewise they benefit from multiples codes (Paivio et al., 1988). It seems that frequency (Schwanenflugel et al., 1992) and the grammatical status of words (verbs are generally more abstract) may also influence the retrieval and access of abstract words (Van Hell & De Groot, 1998). Another interesting aspect of this model is the idea of 'distribution of features' related to a more connectionist view which contrasts with previous highly modular models (e.g., RHM, Word Association Model or the Concept Mediation Model). Hence, the DCFM enables us to represent the mapping of lexical and conceptual representations in a smoother way, taking into consideration the dynamics of bilingualism.

While the DCFM succeeded in going further than previous models with a new typology of words and their respective conceptual representations, it shows limitations for handling semantic specificities such as the polysemy of words e.g., the word '*carte'* in French can mean *'menu', 'credit card', 'student/identity card', 'postcard', 'map', 'card'* and so on.

In sum, the Distributed Conceptual Feature Model was the first model of the bilingual lexicon taking into account the specificities of concrete vs. abstracts words and cognates vs. non-cognates in the bilingual mind. These two types of words might differ in the amount of shared meaning: they can be completely different or can share a certain degree of meaning. Nevertheless, convergence and divergence at semantic levels vary a lot between languages and the model does not sufficiently account for this.

1.1.6. The Shared Distributed Asymmetrical Model (SDAM)

As seen in the Distributed Conceptual Feature Model, links may exist between conceptual and lexical representations distributed differently depending on the nature of the words. The Shared Distributed Asymmetrical Model (Dong et al., 2005) develops the principle of distribution further and adds the representation of different strengths of links between the L1 and the L2. Compared to the Word Association and the Concept Mediation Models, the Shared Distributed Asymmetrical Model continues and illustrates more deeply the developmental changes from elementary to advanced second language learners.

One of the contributions of this model is the proposition of a dynamic approach that concerns not only L2 vocabulary learning but also the inclusion of a new distinction at the conceptual level which allows the author to explain bilingual performance at different levels of second language acquisition (Pavlenko, 2009). The innovative proposition of Dong et al. (2005) relies on the distinction of L1 and L2 elements that are language and culture specific in their conceptual components, e.g., the color *red* is more salient to the concept of *bride* in Chinese than it is in English because brides wear red instead of white (p. 233). The second contribution of this model is the introduction of conceptual convergence between the L1 and the L2 during second language acquisition adding to the model shared conceptual representations (defined as common elements) besides language specific representations.

The Shared Distributed Asymmetrical Model (Dong et al., 2005) claims as the previous model that translation pairs do not share meaning completely. Instead, the representation of word meaning is spread in a set of elementary conceptual units, hence, translation pairs have a subset of features that are shared between languages. Additionally, each word is associated with several language specific conceptual elements as described in **Figure 6**. To illustrate this, following the current model, for the translation pair '*love'* and '*amor',* the conceptual representation in Spanish would involve that *amor* is a feeling, but its equivalent in English (*love*) can also express a preference, such as *'I love chocolate ice cream',* so the concepts of '*love'* and *'amor'* have different constraints depending on the language. Additionally, in Spanish the concept of *amor* is mainly restricted to the use with animates, e.g., '*amo a mi mama'* (*I love my mother*), but this is not the case in English, e.g., '*I love pancakes'*. On the other hand, the shared subset units might be that in the context of '*I love somebody'* the verb may be used in both languages, e.g., '*amo a alguien'*.

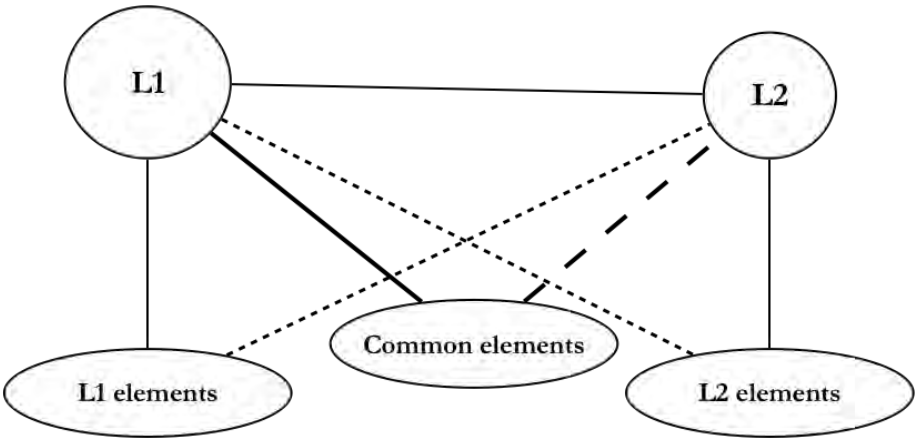

Figure 6 The Shared Asymmetrical Model (Adapted from Dong et al., 2005) in (Pavlenko, 2009, p. 146)

As in the RHM, L2 form representations and conceptual representations are supposed to be weaker than for L1 in the first stages of L2 learning, but as improvements in the L2 are made, the connections strengthen progressively (Groot, 2013). It is worth mentioning that one of the strongest points of this model is to take into account of the evolution of second language learning in the modeling of the mental lexicon.

The SDAM proposes that during the beginning of L2 learning, L2 form representations are connected to conceptual units that are shared with the L1 system and to L2 specific elements as depicted in broken lines in **Figure 6**. In this stage, it is a natural process for the second language learners to compare the grammatical systems of his languages, one first step towards the construction of linguistic awareness that characterizes bilinguals (see Bialystok, 1988).

As the level of L2 proficiency is increasing, connections will gradually strengthen between L2 form representations and L2 specific concepts. Meanwhile, previous connections between common elements and L1 specific elements weaken (de Groot, 2013).

The SDAM is supported by the study of Dong et al. (2005). Six groups of subjects with different levels of L2 proficiency in English (two groups of 1st year English majors and two groups of 3rd year English majors) as well as two monolinguals' groups (Chinese monolinguals knowing little English and English monolinguals knowing little Chinese). Additionally, different language materials are presented (Chinese or English) to the different groups of subjects on a ranking task that consists of rating sets of words with respect to semantic closeness between a *head word*[6] and seven other words that were related to different degrees. The main conclusion of this study was that bilinguals' ratings were closer to monolinguals' according to their proficiency levels in the L2: advanced L2 English learners' ratings were closer to English monolinguals' ratings than less advanced

[6] One of the head word used in this experiment was FRUIT associated to '*lamp', 'apple', 'melon', 'nut', 'lower', 'tomato'* (Dong et al., 2005) p., 229.

English learners' ratings. Most interestingly, the ratings of proficient English learners were dissimilar to monolingual Chinese ratings. This suggests that L2 learners can gradually become more native like. Moreover, the difference in the ratings between beginners and advanced learners supports the hypothesis stating that during L2 learning, connections between the L1 form and the L2 specific conceptual units are made (de Groot, 2013).

This SDAM makes the assumption of an asymmetrical representation of concepts between languages, implying that there are concepts that are not equivalent in both languages. This is explained through a second study of English-Chinese bilinguals based on a lexical decision task in order to analyze different semantic relations between words that varied in conceptual associative strength. The methodological design was structured in 6 classes of primes and targets (e.g., sail-SHIP, taste-FOOD, whisper-SPEAK) (Dong et al., 2005, p. 225), and 2 control stimuli (e.g., unrelated words). Additionally, different language conditions were used, e.g., between-language and within-language. Results suggested that responses of L1 conditions (Chinese) were processed faster than L2 conditions (English). This supports the hypothesis of stronger links between the L1 words and L1 concepts, which is also in line with the Distributed Feature Model (de Groot, 1992). Regarding the types of words semantically related at different levels, priming effects were shown only in the between-language condition suggesting evidence for a shared conceptual system.

The SDAM is one of the first to propose an approach taking into consideration proficiency of languages and crosslinguistic differences in vocabulary learning. Moreover, it recognizes that L2 learners can progressively attain native-like standards once strengthening of connections between the system and the conceptual units is obtained. Nevertheless, this model has also been criticized regarding the nature of the structure of the conceptual representations, as defined later, and distinguished from semantic representations(Pavlenko, 2009, p. 146).

1.1.7. The Modified Hierarchical Model (MHM)

Since semantics is the main focus of interest in this research, the Modified Hierarchical Model or MHM's (Pavlenko, 2009) framework is of specific interest. The MHM follows the Shared Asymmetrical Model (Dong et al., 2005) and the Distributed Feature Model (de Groot, 1992) and agrees that crosslinguistic conceptual representations are not completely shared but partially shared in linguistic categories. Additionally, it holds that lexical concepts are language-specific, a conception that enables a dynamic perspective in which concepts are defined following the specific social context and depending on the language concerned, and where transfer is conceived in terms of activation and inhibition. An illustration of this assumption is supported by the fact that some words cannot be fully translated from one language into the other. Actually, some concepts are not necessarily lexicalized in all languages as for example English *privacy* and *frustration* that have no equivalent in Russian (de Groot, 2013). A similar example is the

expression of the Spanish concept of '*estrenar*' which means '*wearing something for the first time*' that is inexistant in English.

The MHM proposes a separation between the semantic and the conceptual representations. Pavlenko, (2009) also distinguishes two types of transfer, arising at the conceptual or the semantic level (this issue is detailed in section 1.1.7.1).

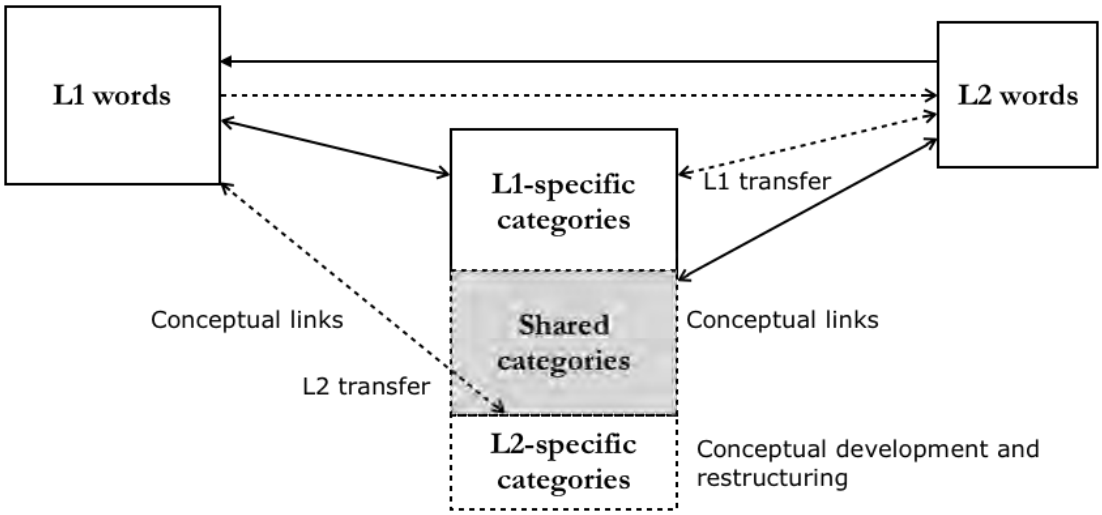

Figure 7 The Modified Hierarchical Model (Pavlenko, 2009, p. 147)

In contrast to the models detailed previously, in the MHM, the conceptual store is depicted in a block divided in three parts that represent words that are fully shared or partially overlap between languages and fully language-specific categories in the L1 or the L2, as represented in **Figure 7**.

This model is interesting since it describes L1 and L2 transfer and introduces the notions of conceptual development and restructuring which are related to L2 language proficiency.

1.1.7.1. Conceptual Transfer

Shared representations in both languages may trigger unusual or non-canonical productions in the L1 or in the L2, corresponding for example to the use of the L2 in accordance with L1 linguistic categories or the other way around. This is what Pavlenko calls conceptual transfer: a conceptual content that is reassigned from the L1 to the L2, producing atypical meaning associations between languages. For example Spanish-English bilinguals may replace the meaning in Spanish of '*casualidad*' ('*coincidence*' in English) by the meaning of *'casualty'*, producing as result atypical meaning associations with '*death*' or '*injury*' instead of expressing the meaning of '*casualidad*' in Spanish (Heredia & Brown, 2012, p. 241).

The main purpose of Pavlenko is to propose a distinction between semantic and conceptual representations that would also have an influence on the definition she

develops for conceptual transfer. Semantic representations involve implicit knowledge about the mapping between lexical units and concepts or between others lexical units. Following this way, different concepts are expressed through a single lexical representation (e.g., polysemy of words) affecting the linking processes between lexical and conceptual representations (see section 1.2.1.3 about Lemmatic Transfer). Hence, lexical and conceptual links are constructed including collocation, and synonymy /antonymy associations (Pavlenko, 2009, p. 148).

On the one hand, semantic transfer occurs during the mapping of lexical forms onto concepts, for example an L2 English learner could use the expression '*he turns me crazy*" instead of '*he drives me crazy*' because of his/her mapping on Spanish structure '*me vuelve loco,*' which is a temporary state of being crazy because of someone else. On the other hand, conceptual representations concern the structure itself of the category of words, including multimodal conceptual representations that exist in a specific language (e.g., the polysemy of words). Here, the conceptual representations are considered to be dynamic and context dependent. Hence, conceptual transfer may be related to category knowledge that is unsatisfactorily transferred from one language to another. For example, an English Speaker of L2 Russian who uses the word '*Chashka*' (to refer mistakenly to a paper drinking container following the English concept of '*cup*') resulting in a conceptual transfer because it involves multimodal conceptual representations (Pavlenko, 2009, p. 149).

The MHM proposes three types of conceptual equivalence that influence L2 vocabulary learning and conceptual transfer, as described in **Figure 8**, indicating how the conceptual links facilitate L1 vocabulary learning and the type of L1 transfer observed in L2 learners.

1) In the first case, there is conceptual equivalence between L2 and L1 lexical forms; here learning involves positive transfer since no conceptual restructuring is necessary.
2) Second, the partial non-equivalence produces conceptual overlaps that may render learning difficult through overgeneralization of the concepts between languages and exhibit negative transfer, likewise, positive transfer is achieved once common overlaps are identified and restructuring processes are made.
3) Third, the conceptual non-equivalence helps to create new categories since they do not exist in the other language, however this is supposed to be more helpful for learning new objects or concrete concepts than abstract concepts.

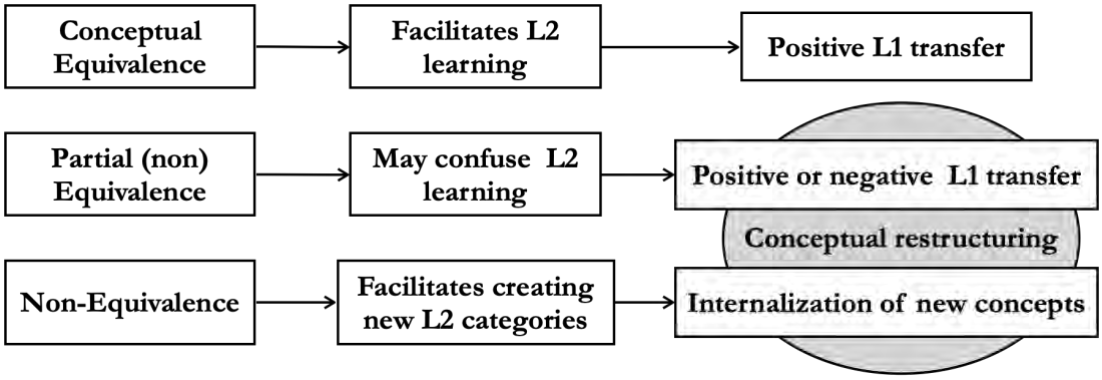

Figure 8 Representation of the 3 types of conceptual equivalence (adapted from Pavlenko, 2009)

It seems that conceptual restructuring constitutes a passage from partial or non-equivalence concepts and non-equivalence to concept learning. In the next section, this definition is going to be developed.

1.1.7.2. Conceptual Restructuring

Conceptual restructuring is defined as a gradual restructuring process of the existing conceptual representations in order to develop new ones that get closer to the L2 concepts (Pavlenko, 2009). In other words:

> "*it implements the idea that a main goal of L2 vocabulary learning is conceptual restructuring such that L2 conceptual representations will ultimately be as native-like language as possible*" (de Groot, 2013, p. 181).

As mentioned before, during L2 learning processes conceptual transfer occurs, that is, L1 concepts that are not completely cross linguistically shared with the L2 can be mistakenly assigned to the L2 (see **Figure 8**).

In order to explain this conflict of conceptual non-correspondence between languages, Pavlenko assumes that the use of L2 words in bilinguals differs from native language use, illustrated through what she calls a *Bilingual Semantic Accent* (2009), i.e., a kind of accent that is not phonological but semantic in nature and manifests itself through the use of L2 words in nonnative-like manner. This non-native likeness is explained by the fact that linguistic information is influenced by the L1, or because it deviates from the context in which a native speaker would produce that particular word. As an illustration, in French the word *'pâté'* can be used in different contexts, such as: the one we eat (*'un pâté de canard'*), the one made by children on the beach to construct a sandcastle (*'le pâté de sable'*), or a group of houses (*'le pâté de maison'*). For sure, the use of this word may be confusing for L2 French learners, because it depends on specific contexts.

Another process that involves conceptual restructuring is a phenomenon called *Meaning Nuances*, in which a single concept is composed by L1 and L2 concepts, merging specific meaning nuances of each language. In this case, both languages are '*accentuated*'.

In other cases, merged concepts can be created, that is containing both L1 and L2 specific meaning nuances. *Bilingual Semantic Accent* involves an L2 uncanonical production involving the L1, instead *Meaning Nuances* involves a merged production including the L2 and the L1 conceptual features.

Taking into account the MHM and its hypothesis about the stages in which the conceptual restructuring occurs, about how it is followed by successful L2 learning, and the place attributed to the L1 during this process, the following questions emerge:

- *What are the factors that determine that some L1 concepts remain native-like while others are unified (L1+L2) into a single concept?*
- *Assuming the restructuring process that superposing two languages can produce during learning, are previously acquired languages an obstruction to foreign language learning at the conceptual level?*
- *During the improvement of L2 learning, is the L1 affected at the semantic/ conceptual level?*

As can be seen, from the moment a speaker starts using two languages simultaneously, as a result of bilingualism, languages evolve together and complete each other at all linguistic levels, reflecting subtle changes. Different studies stand for restructuring processes, at some point they even suggest that bilinguals differ from monolinguals not only regarding their L2, which is explained above, but also regarding the use of their L1 (Cook, 2003). This is shown through different phenomena such as crosslinguistic word naming patterns, crosslinguistic grammatical gender assignation and crosslinguistic categorization tasks. A brief discussion of these studies is proposed below (for deeper analyses see the literature review by de Groot 2013, p. 181–189).

Regarding crosslinguistic word naming patterns, proficient bilingual's L2 naming patterns differ from native naming patterns even after years of second languages acquisition and immersion (Malt & Sloman, 2003). Such findings are obtained after analyzing meticulously a recipient object naming task submitted to highly proficient bilinguals, whose performance differed from monolinguals, suggesting that bilinguals' L2 concepts differ from the corresponding concepts of native speakers of the L2 even after a long period of L2 immersion. More interestingly, it seems that this also is the case for the other languages the bilingual uses. Ammel et al. (2005, 2009) report that not only is the L2 performance of bilinguals deviant from L2 monolinguals' but also from L1 naming patterns reported for monolinguals. Malt & Sloman (2003) claim that bilinguals' concept category boundaries overlap between their languages. These boundaries differ from native speakers of either language (L1 and L2). However, the more experienced bilinguals are in their L2, the closer the concepts correspond to those of native speakers.

Nevertheless, restructuring might not be limited to naming patterns but arise also with respect to grammatical cross-linguistic differences. The following study is an example of how the grammatical gender in L1 affects the L2 and how speakers of English (as a language without grammatical gender) cope with the assignment of grammatical gender in the L2. Malt & Sloman (2003) hypothesized that 1) gender assignment is not arbitrary, and will be related to properties of animated concepts (feminine or masculine) and 2) that

gender assignment in L2 naming may be influenced by gender in the L1. In order to test these hypotheses a first experiment involved English L2 speakers in their gender assignments while naming animals and objects in their L1, Spanish or German. Results showed that English speakers successfully assigned gender in both languages for animals in spite of the fact that English is a genderless language. Researchers explained that the perceived masculine or feminine attributes associated to the properties of referents influence gender assignment of English speakers. E.g., '*el búfalo*', buffalo in Spanish or '*der Büffel*' in German might be associated to masculine attributes such as a big, strong and imposing. The experiment concludes that gender assignment may not be entirely arbitrary but influenced by their referents' properties.

Boroditsky, Schmidt, & Phillips (2003) studied two groups of native speakers (Spanish and German) who were naïve in the English language. Briefly, the task consisted in learning object names in English that were associated to a fictional gender, for example a *frog* called *Harriet*, after going through a distracting activity, participants were asked to recall the fictional gender of the objects that were based on half feminine and masculine in German or in Spanish. The objective of this procedure was to compare the gender assignment made by the participants with different L1 (German or Spanish). Results suggest that the hypothetical grammatical gender system of English was remembered better when the gender was congruent with the L1 of the participant, suggesting that grammatical gender assignment in L2 (in this case a new language) is influenced by the corresponding gender in the L1.

Regarding grammatical number and object concepts and how these aspects are also influenced by bilingual categorization abilities, a study by Athanasopoulos & Kasai (2008) compared advanced and intermediate L1 Japanese learners of English (L2) in a trial matching task in which participants categorized three objects that alternately corresponded in shape or color to the target stimuli. The results were compared with those of Japanese and English monolinguals and bilinguals because both systems differ in the way individuation is constructed, specifically for mass nouns in English and for inanimate nouns in Japanese. Individuation serves to grammatically mark substances or objects, for example in English, objects (count nouns) are individuated (e.g., *two apples*) and substances (mass nouns as *water*) are nonindividuated unless the speaker uses quantifiers (e.g., *two glasses of water*). In contrast, in Japanese, it is impossible to refer to common nouns that have inanimate referents using a grammatical number as in English; to do so, the use of numeral classifiers is necessary (e.g., '*san ko no ringo'*, 'three pieces of apple' (Athanasopoulos & Kasai, 2008, p. 106). In other words, in Japanese quantifiers are used for common nouns (*apple*) as in English for mass nouns (*water*). Athanasopoulos & Kasai (2008) focused on shape patterns because these constitute a mark of *individuation* in order to mark plurals, so the shape feature would represent a more prominent aspect for English speakers than for Japanese speakers. The results suggest that Japanese with advanced L2 English show the same shape patterns as English monolinguals while less fluent L2 English learners followed a shapeless pattern of categorization as do Japanese monolinguals. The authors suggest that the growth in the L2 English competence makes

shape a prominent feature in the learners' conceptual object structure, in other words, a cognitive restructuring process has taken place in advanced Japanese-English bilinguals prioritizing shape categorization.

These three studies provide examples showing at which levels the restructuring process is observable in second language acquisition, as well as the semantic nuances in the L1 and the L2 at different stages of conceptual construction.

To summarize, the strong points of the Modified Hierarchical Model of Pavlenko (2009) are the idea of conceptual transfer and the restructuring processes that shapes differently the bilingual lexicon depending on the level of proficiency in L2 and language specificities. This allows recognizing the dynamics and the complexity of the mental lexicon since Pavlenko proposes that L2 and L1 transfer processes is a fundamental process of language learning and improvement. It is worth mentioning that transfer phenomena are bidirectional (not only from L1 to L2 but also from L2 to L1). Thus, L2 knowledge might also change L1 usage (Levy et al., 2007). This is the case when there is a semantic accent in both languages.

1.1.8. Summary

The main aim of this section was to explore hierarchically the models that contributed to the understanding of the bilingual lexicon as well as presenting some of their limits. This section is limited to the most representative models according to the current research and objectives, but it is important to recognize the existence of other models that are not included in this review.

The models detailed above differ in the way representation of the lexical and the conceptual levels are built up and linked together. As can be seen in **Figure 9**, at the beginning the Word Association and the Concept Mediation Model established limited interaction at the lexical levels through translation processes. Later, the Revised Hierarchical Model added modulation of the strength of the links between the conceptual and the lexical level, links with the L1 being stronger than links with the L2. Subsequent models such as The Distributed and Conceptual feature Model take into consideration the word type effects in the bilingual mental lexicon proposing a distribution of conceptual features as well as crosslinguistic facilitator effects between languages. In the same line, the Shared Asymmetrical Distributed Model goes further with a more dynamic representation of the bilingual processes. Finally, the Modified Hierarchical Model included conceptual transfer and restructuring aspects that are in line with the present research.

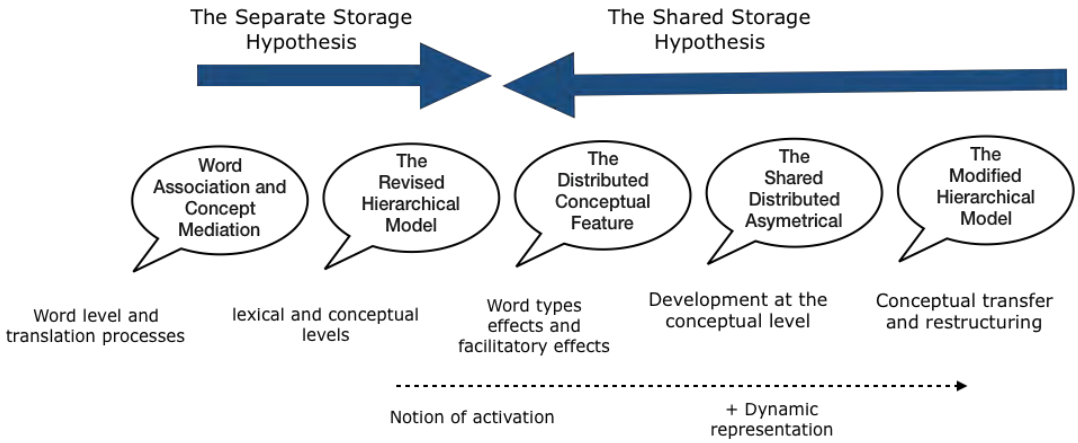

Figure 9 Summary of the models presented and their representation of the mental lexicon

One of the points that is taken into account in the present review is the nature of the experimental tasks involved: in early research, experimental tasks were centered on translation processes that are underrepresented in the actual processes involved in bilingual language use. For this purpose, Kroll & Groot (1997) advance that tasks such as translation or picture naming are out of the context of bilingual performance, the authors stress the need to adopt experimental tasks that are in accordance with bilingual performance in representative contexts of second language use for further research.

A second aspect that is considered here is the place of the lemma level (Levelt et al, 1999) which is problematic since it is not sufficiently developed in bilingual modeling (e.g., Woutersen et al., 1994). This may be explained by the challenge related to the analysis of syntactic and semantic processes in controlled experimental tasks. One of the contributions of the present research is to take into consideration the lemma and the lexeme levels, which have been underestimated in the past, through the study of the intersection between word form and semantics in lexical transfer.

1.2. CROSSLINGUISTIC INFLUENCE (CLI)

Crosslinguistic influence (CLI) involves the way linguistic knowledge of an already learned/acquired language affects or influences the learning/acquisition of another language in production, perception and comprehension (Jarvis & Pavlenko, 2010) with respect to all linguistic levels: phonological, lexical, semantic, syntactic and morphological. This interaction between systems has been traditionally investigated from the L1 to the L2 mostly in morphological, syntactical and lexical levels; other levels such as semantics have been gradually studied over the years (Gathercole & Moawad, 2010). In contrast, reverse CLI from L2 to L1 has been less reported in the literature despite the recognition of the fact that L1 competence is not a stable and fixed system as presumed, so the native language is also subject to change even in adulthood (Major, 1992). For example, in phonetics, research suggests that the human perceptual system is at some degree flexible throughout life course, thus, modification can be linked to the way the input changes. We will refer to CLI as the influence of one language on the other in bilingual contexts (e.g., phonological and lexical). In contrast, transfer will be used here to refer to crosslinguistic influence in more specific levels (e.g., lemmatic transfer, lexemic transfer). From our positioning, the terms CLI and transfer are interchangeable describing the same phenomenon (e.g., Jarvis & Pavlenko, 2010).

The aim of this section is to review studies concerned with reverse CLI, that is, from the L2 to the L1. We will focus our review on the phonological and the lexical level which are the most represented in our data collection.

1.2.1. CLI at the Lexical Level

Lexical transfer refers to the influence of word knowledge or language use of one language in another language (Jarvis & Pavlenko, 2010). CLI at the lexical level is of great interest in SLA since it allows researchers to explore the question of interconnectivity between lexicons (Cenoz et al., 2007) in the which typological relation between languages play a role, as well as the crosslinguistic interconnections such as language proficiency.

Following Ringbom (1987, p. 37), the lexical level involves 6 types of knowledge or dimensions that define holistically the knowledge of a word in a language :

1) Accessibility of the word in the mental lexicon;
2) Knowledge related to spelling and pronunciation of a word;
3) Knowledge related to the word's grammatical class and syntactic constraints;
4) Knowledge related to the meaning of the word (semantics);
5) Knowledge related to conventional multiword combination of a word or collocations;
6) Knowledge related to the associations with other words and concepts.

Lexical knowledge of a language includes also the acquisition of some linguistic competence that allows the speaker to use successfully the number of words learned in a particular language in the right context and with the right interlocutor. That is, the knowledge of words is not limited to linguistic knowledge but also includes a set of sociolinguistic aspects, both types of knowledge (tacit knowledge and the ability to use that knowledge) conforming the definition of the communicative competence of a speaker (Hymes, 1972).

During the process of second language learning, as a new word acquired in the new language, it also implies the construction of multiple mental interconnections as discussed in section 1.1.2 related to lemma and lexeme levels not only with other L2 words, but with the L1 as well.

Jarvis & Pavlenko (2010) propose an interesting illustration in English of how mental interconnections can be constructed following the list presented above:

1) when an English speaker hears the word wiggle it co-activates frequent words such as wriggle.
2) Speakers identify go, goes, went, gone as different forms of the same word.
3) The word thankful in a sentence will be followed by the preposition for.
4) Speakers recognize feline and cat as synonyms.
5) They realize that launch tends to follow words like ship and rocket.
6) While thinking of the word chair, associations such as couch, sofa, desk, table or even the department chair are activated.

In bilinguals, such interconnections can be limitless. The multiple connections that occur between languages are illustrated in cascade models (e.g., Caramazza, 1997; Dell, 1986) in which lexical activation is spread or co-activated to the next layers or levels of activation, that is the phonological or sub lexical representations, explaining interconnection of lexical nodes between languages.

We propose to make a distinction between lemmatic and lexemic transfer since it will allow us to evidence and analyze more deeply the nature of lexical CLI. In most studies, the distinction between lemma and lexeme transfer is not taken into consideration (since they are rarely included in bilingual modelling) with the exception of Jarvis, (2009). He proposes a review presenting the most prominent findings related to lemmatic and lexemic transfer that are detailed in the next section.

1.2.1.1. Lemmatic and Lexemic Levels

In this section, lexical transfer is categorized into two levels: lemmatic and lexemic. Our theoretical positioning is in line with Levelt's definition of the lemma (Levelt, 1993).

The lexeme level concerns the phonological and orthographical structure of words, i.e., information about how to pronounce and spell the inflectional form of words. Hence, this level includes the orthographic and phonological representations of words, e.g., for '*go'*, related phonological and orthographic representations such as '*goes', 'went', 'going'* are associated (Jarvis & Pavlenko, 2010). These variants are allomorphs of the lexeme *go*, representing here morphological declinations of different lexemes linked to the same lemma.

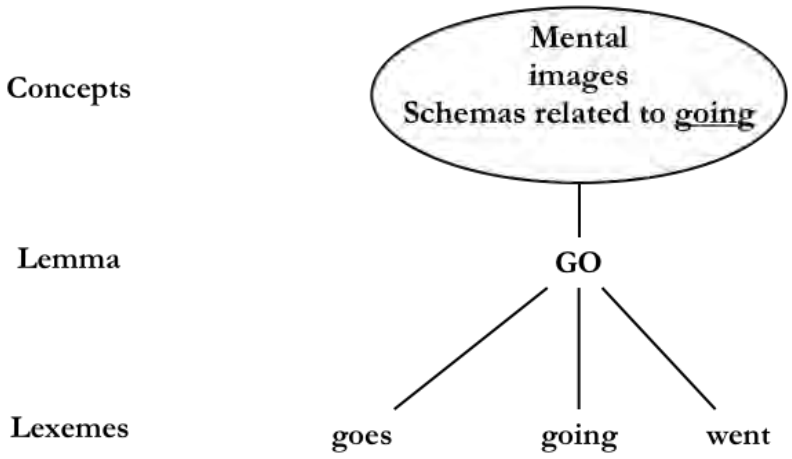

Figure 10 Three levels of Representation adapted from (Jarvis & Pavlenko, 2010, p. 83)

The lemma level concerns semantic and syntactic properties of lexical units, involving information about grammatical class, collocations and syntactic constraints related to it, including polysemy between lexical units.

Recall that lexemes can be linked in different ways during acquisition (Jarvis & Pavlenko, 2010), hence, the learner can link a lexeme in an L2 directly to the closest lexeme in the already acquired L1 (Kroll & Stewart, 1994), that is lexeme-lexeme link, or the learner can link the new lexeme in L2 directly to an L1 lemma that includes a group of lexemes and their respective representations, such as the previously cited example in **Figure 10** in which *'go'* is related to '*goes', 'went', 'going'*, i.e. a lexeme-lemma link (Jiang, 2002). Other possibilities of construction are also possible, such as a links from the L2 lemma to an L2 lexeme or L1 lexeme, or L2 lemma to an L1 lemma.

While formal features such as orthography and phonology intervene at the lexeme level, at the lemma level, each lexical unit involves semantic links with other units, consequently, during reading or processing of a word, other related complex words belonging to the same family are co-activated as well, e.g., '*birthday*' is closely related to '*birth'* and '*day'* (Mulder et al., 2013). The morphological family constitutes a set of words that are derived from a common stem and it has been shown that the number of words that are derivations from the same stem, i.e., the *Morphological Family Size*, affects semantic processing (De Jong et al., 2000). In other words, a morphological family is based on

morphological links between several words, that share not only phonological and orthographic features but also semantic links.

Jarvis & Pavlenko (2010) propose a three levels representation of lexical access, the first level corresponding to lexemes, the second to lemmas and the third level to concepts (see **Figure 10**). Hence, L1 and L2 knowledge form mental links between languages (e.g., Kroll & De Groot, 2005) and different linkages between this information is constructed around these three levels of lexical representation.

The conceptual level includes mental images or impressions stored as sensory information (auditory, olfactive, tactile and kinesthetic) and schemas. These elements are then organized in conceptual categories (Murphy, 2004) that play a role in individual conceptual network constructions that depend on personal knowledge and experiences, for example a *dog* can be categorized as an animal, a mammal, but also as a pet or a friend. As outlined with previous examples, different associations are possible, concepts can be associated to a) a particular language, this is the case when there is no conceptual equivalence between different languages e.g., '*dodo'* in Spanish, and b) words that are at the same time associated with other lemmas, e.g., in case of synonyms.

Considering this, recall that a second language learner has multiple ways of constructing crosslinguistic associations with his/ her already learned languages. Nevertheless, when comparing conceptualization in different languages, there exist fine distinctions that are organized differently, following Ameel et al. (2009). These authors suggest that naming patterns are language-specific and not universal, for example in Chinese there are 19 objects that can be categorized as '*jar*' and 13 different objects as '*bottle',* which may be rarely the case of English. Therefore, it can be assumed that associations between the three levels can be structured differently, depending of the language. In order to illustrate the interactions between the lemma and lexeme levels in the bilingual lexicon, Ecke (2009) investigated TOTs (words on the Tip Of the Tongue) evidenced by retrieval failure of a word in one of the languages of the bilingual. In TOTs, the access to the lemma of a word at the syntactic level, and then is succeeded by the conceptual level, meanwhile, at the lexemic level, the phonological and orthographic representations, remains inaccessible. Indeed, TOTs constitute a window into lexical access in speech production. During retrieval, phonological encoding fails which entails the inability to produce the word, even though the semantic and syntactic information is available. In other words, the speaker knows the meaning of the word, the grammatical category of the word but the words do not come out of the mouth. The principal findings about TOTs is that during retrieval, the first letter or sound of a word is commonly the element most frequently accessed in all languages. However, some differences can be observed depending on the syllabic structure of the language (Brown, 1991).

Poulisse (2000) compared L1 and L2 slips of the tongue[7] and articulated her analysis of existing data according to the spreading activation mechanism following the model of Dell (1986) and Levelt (1993) to explore backwards and forward activation. L2 data support the idea that networks of linguistic units are constructed between semantically related lemmas that are co-activated simultaneously. Accordingly, the analyses of slips of the tongue of Dutch L1–English L2 learners involved blends between synonyms or semantically related words in 89% of the data analyzed (e.g., 'sind' = 'sort' and '*kind*' Poulisse, 2000, p. 140).

As we have seen, TOTs and slips of the tongue may arise both at the lemma and at the lexeme level. In the next two sections we summarize the typology of transfers proposed by Jarvis (2009) categorized depending on the level concerned during CLI, section 1.2.1.2 describes types of lexemic transfer and section 1.2.1.2 types of lemmatic transfer.

1.2.1.2. Lexemic Transfer

Lexemic errors provide valuable information about the underlying processes during lexical transfer. Three types of lexemic errors have been observed and studied in the literature by Ringbom (1987, 2001) and Jarvis (2009). These three types are false cognates, language switching and word blending.

1.2.1.2.1. False Cognates

False cognates, deceptive or false friends are words that share some formal features (orthographic and phonological) across languages that involve different referents and meaning in each language, although their form is identical or similar (Lalor & Kirsner, 2001), e.g., *coin* in English and *coin* in French (meaning *corner*).

For a better understanding of false cognates, we must take into account that there are different types of cognates. Jarvis(2009) proposes a typology of the different types that we present in **Table 1** a synthetized version.

Table 1 Typology of Cognates adapted from Jarvis, (2009, p. 107–108)

Genetic relationships		
1) Similar meaning & same form	Meaning:	Form:
	Sw. offer= victim; offer	Eng. & Sw. offer

[7] Slips of the tongue refers to erronous productions which difers from TOT in which an 'imposibity'to production is experienced.

2) Similar meaning & form	Sw. bon= leg; bone	Eng. Bone Sw. bon
3) Dissimilar meaning & same form	Sw. strand =beach Eng. strand (fil)	Eng. & Sw. strand
4) Dissimilar meaning & similar form	Sw. gris =pig Eng. grease	Sw. gris Eng. grease
Non genetic relationship		
5) Similar meaning & same form	Sw. student (in college) Eng. student	Eng. & Sw. student
6) Similar meaning & form	Sw. pensel (paintbrush) Eng. pencil	Sw. pensil Eng. pencil
7) Dissimilar meaning & same form	Sw. kind (cheek) Eng. kind	Eng. & Sw. kind
8) Dissimilar meaning & similar form	Sw. aktuell (current) Eng. actual	Sw. aktuell Eng. actual

Table 1 illustrates two categories of false cognates in Swedish (Sw.) and English (Eng.) whose distinctions are genetic (etymological relations) or non-genetic (non-etymological relations).

These 8 types of cognates explain the differences in the linking processes between lexeme to lexemes / lemmas in bilinguals, and how the lexemic errors illustrate the interaction between lexeme and the lemma level.

Two main types of cognates that are described here: interlingual word pairs (e.g., '*bon'* in Swedish and '*bone'* in English) that share genetic relationship, that is, the same etymological origins, considered as true cognates, and interlingual word pairs that do not share any etymological links, in other words, their form or meaning similarities are coincidence such as '*kind'* in English and Swedish, including international loanwords, e.g., '*radio'* in English and Swedish (Jarvis, 2009).

Cognates that share genetic relationship are further divided into four types as illustrated by Jarvis (2009, pp. 108–109) and reviewed in table 1.

1) In the first case, they share meaning at some levels, and have identical forms such as the Swedish word offer that shares parts of the meaning in English ('offer' but also victim in Swedish).
2) The second has similar meaning and form, such as the Swedish-English pair: 'bon'-bone ('bon' means bone or leg in Swedish).
3) The third etymological false cognates share only form and have different meanings, e.g., 'strand' (beach in Swedish) and 'strand' (English).
4) The fourth group do not share the same meaning but have similar word-form, such as in 'gris' (pig) in Swedish and 'grease' in English.

The second group is composed of false cognates that do not share genetic relationships even though they have similar word-form. These non-genetic cognates or false cognates are divided into four subgroups that might:

5) Share meaning partially and have identical form such as for '*student*' in Swedish and English whose semantic distinction is very subtle. They are also called identical false cognates.
6) Share form and meaning such as in '*pensel*' in Swedish which stand for paintbrush and pencil in English.
7) Share form and have dissimilar meaning inducing this way semantic transfer or misleading crosslinguistic associations. For example, '*kind*' in Swedish which means '*cheek*' associated to '*kind*' in English.
8) Have dissimilarities in form and meaning and inducing crosslinguistic associations. This is the case of the Swedish '*aktuell*' which means '*current*', associated to '*actual*' in English.

As discussed in section 1.1 the role of cognates and non-cognates and their representation in the bilingual mind has been of interest for the modeling of the bilingual lexicon. The main objectives of these studies were to observe differences between bilingual and monolingual during the processing of cognates and non-cognates, as manifested in differences in response times. Hence allowing to decode spreading activation processes as well as the role of non-selected languages while processing a selected language. Conclusions differ depending on authors and experimental procedures.

Early studies with bilinguals reported priming/facilitatory effects exclusively for cognate words (e.g., '*rich-rico*') that share the same form and meaning and no priming/facilitatory effects for non-cognates words (e.g. '*gabia-jaula')* that share only meaning (Scarborough et al., 1984). Later studies observed priming effects in both cases, for cognates and non-cognates (de Groot & Nas, 1991) suggesting that cognate translations (e.g., '*student-estudiante*') share conceptual representation, while noncognate translations (e.g., '*face-cara*') have separate representations. This led to the proposition of

the Distributed Feature Model (de Groot, 1992). Sánchez-Casas and García-Albea's (2005) review suggests that the level of processing of true cognates shows high priming effects, that is, the orthographic form and their meaning are activated while processing the target word in the other language. They suggest that cognates are stored jointly in memory, which may not be the case of non-cognates.

For false cognates, priming effects are observed only when interlingual word pairs are formally identical, however this effect is less pronounced than in cognates, in which priming effects increase due to formal and meaning overlaps (Gerard & Scarborough, 1989). However, the García-Albea et al. (1996) study did not show any priming effects for false cognates and concluded that form similarity itself cannot account for facilitatory effects for cognates. Yet, meaning by itself is not responsible for priming effects, as shown by the absence of priming effects for non-cognates. Final remarks point out that cognate effects are not exclusively related to cross-linguistic form and meaning similarities but also to morphological links that exist between languages (Lalor & Kirsner, 2001). This aspect is going into be treated in section 1.2.2.2.

An observation that underscores the hypothesis of morphological links between cognates is that during SLA, cognate words seem to be easier to learn than non-cognates words(de Groot, 2011).

In favor of meaning overlaps, Sánchez-Casas et al., (1992) showed that in a translation recognition task, recognition latencies were slowed down for non-cognate translations, which was explained by distinct meanings between words. However, for cognate translations, recognition was faster regardless of the degrees of meaning overlap between the words. Word form and meaning play distinct roles in early and later stages of the recognition processes: word form plays an important role in early processes, which is supported by facilitation effects in false cognates at very short priming settings. In contrast, meaning would play an important role in later recognition processes supported by facilitation at longer priming settings Sánchez-Casas et al., 1992).

Bilinguals show some sensitivity to transfer depending on the similarities that exist between their languages. For example Finnish speakers who learn English have shown less difficulties for false cognates, and fewer lexical transfer than Swedish learners of English (Ringbom, 2007), suggesting that the levels of activation play a major role in lexical transfer.

To conclude it should be noted that false cognates are of great interest because of the partial formal similarities and the divergent meaning they involve. This relation would function as an intermediate point between form and meanings overlaps.

1.2.1.2.2. Unintentional Language Switching

Unintentional language switching concerns the use of words in a different language which results from high activation levels of the non-selected language during production in another language.

Following the studies of Cenoz et al. (2001) and Ringdom (2007), activation threshold levels are not the only theoretical framework explaining transfer. Typologically related languages would play a role as well during lexical intrusion. This was concluded from a study of multilinguals (L1 English learners of Swedish with L2 German) for whom the lexical intrusions observed in Swedish arise from the L2 German, instead of the L1 English, suggesting that languages overlap, and that proximity may also influence the direction of transfer, independently of their language status (L1 or L2).

1.2.1.2.3. Word blending

Word blending (or blending) is a type of transfer which results in new words that are created from formal properties coming from different languages. In this type of transfer, the word form is modified creating a new lexeme that does not belong to either the recipient or the source language. In other words, there is a blending of two lexemes into one and sometimes the blending also involves a word stem from one language mixed with the inflectional morpheme from the other, e.g., "*All these wooden golves must be cleaned, Sw. golv* = 'floor"' (Ringbom, 1987, p. 154) cited by Jarvis (2009), in this example the lexeme 'golve' arises from Swedish 'gol' that means floor.

A similar phenomenon is called *loanword* by some authors (Grosjean, 2010; Romaine, 1989), referring to cases of morphological adaptations made on the L1 following the rules of the L1 as in "*codeswitché*" in the example of Grosjean (2010, p. 56) as in "*Ça m'étonnerait qu'on ait codeswitché"* (Grosjean, 2010, p. 56).

Given all these points, lexemic transfer can be manifested in the use of false cognates, word blending or loanwords. What is common to all these types of transfer is that they are explained through competition and activation threshold levels that would trigger the presence of intrusion such as cross-language similarities at some representational levels. At the lexeme level, factors such as word frequency, language mastery and the chronology of language learning play a determinant role.

1.2.1.3. Lemmatic Transfer

Lemmatic transfer, also called semantic transfer by Ringbom, encompasses collocational, morphological, and syntactic constraints or restrictions of words in which semantic mappings between lemmas are made, e.g., the restricted use of phrasal verbs in English in which a combination of verbs, prepositions, or/and adverbs creates a new definition. As discussed previously the lemma involves semantic associations and constraints (de Bot, 2004). Following Jarvis (2009) lemmatic transfer is composed by two major features: lemmas and conceptual associations (e.g., the links between two lemmas that are synonyms such as *homework* and *assignment*), the second feature correspond to the connections between the lemmas themselves (*e.g., 'goes, went, going'*) as seen in **Figure 10**.

Lemmatic transfer involves 'learned' crosslinguistic associations or mental links that join lexical units with concepts and lexical units between them (Jarvis, 2009). He

proposes that contrary to lexemic or formal transfer, lemmatic or semantic transfer is supposed to occur when the speaker has successfully acquired language-specific meanings in the source language (Jarvis & Pavlenko, 2010, p. 82).

Jarvis (2009) proposed a typology contributing to a better understanding of the crosslinguistic interactions at the lemma level in the bilingual lexicon. He defined 4 types of lemmatic transfer: semantics extensions, calques, collocational transfer and subcategorization transfer that we are going to discuss in what follows.

1.2.1.3.1. Semantic Extensions

Semantic extensions or loan-shifts concern the way the lemmas are linked to concepts. As an illustration, the polysemy of words works differently in each language, that is, some features can be cross-linguistically shared and others not. The studies of Meriläinen (2006) on written L2 production in English focused on particular errors involving mistaken attributions or associations of meaning in a language. She analyzed the written production of a Finnish speaker that reflects the use of L2 English lemma 'spin' calqued on the sense of L1 Finnish '*Kehrä ta'* that defines two different concepts: *'to spin'* and *'to pur'*. In this example, the speaker has extended the meaning of the English '*to spin*' into '*to pur*' such as in the polysemic lemma '*kehrä ta'* in Finnish (Meriläinen, 2006, p. 92) cited in Jarvis, 2009).

In the reverse direction, Pavlenko and Jarvis' (2002) study of oral narratives in L2 English of Russian speakers suggested that L2 meanings had influenced L1 lexical structures such as in the use of "kamera" (referring in Russian to video, TV or movie camera) instead of the standardized Russian word "plenka" and "photoapparat" respectively (Pavlenko & Jarvis, 2002, p. 201). Pavlenko observed the same kind of transfer in Russian-English bilinguals when they wanted to express concepts like *privacy* or *personal space* and *frustration* in Russian, since there is no equivalence for these concepts in Russian (2002, 2003).

These illustrations suggest that conceptual representation varies depending on the language and semantic specificities, advocating for the existence of erroneous associations between lemmas and concept levels during learning processes Jarvis (2009). In opposition to the explanation based on misleading learned associations, we may assume that semantic extensions could be attributable to linguistic resources that complete the semantic spheres between languages such as in the case of Russian which has no means to express concepts such as *privacy*. We can imagine that bilingualism may lead to the discovery of nuances of meaning across languages that the bilingual would probably integrate in his/her L1 competence, that is enriched with new semantic representations.

It is also interesting to observe transfer in written production as a highly specialized off-line process allowing the speaker to think along and modify during the entire writing process. This is contrary to oral production which is a direct and spontaneous on-line process in which corrections are only possible through accumulation of propositions. Written production is an indirect mode of communication which allows

for different stages of construction with the help of reformulation, corrections, and so on. Hence, the resistance of such semantic transfer in written production suggests that the use of these forms is rooted in a deep level of the bilingual lexicon.

Other studies concerned judgment tasks (Jiang, 2002, 2004) and observed that Chinese and Korean speakers judged the relation between words pairs in English better when their translation equivalents corresponded to a single word in Chinese (e.g., '*problem*' and '*question'* that are translated by *'wenti*') than when their translations matched with different words in L1 (e.g., '*interrupt'* and '*interfere'* for '*daduan'* and '*ganrao')*. This motivated an alternative interpretation, that is, learners seem to associate L2 lexemes with L1 lemmas, so the L1 and L2 lexemes became morphological variants of the same lemma. In the present case, '*problem'* and '*wenti'* were linked to the lemma '*wenti',* which is now included in another variant 'problem'.

Another interesting example involving Spanish-French bilinguals is the study by Grosjean & Py (1991) in which acceptability and judgment tasks were used to observe the restructuring of L1 competence and language use in Spanish immigrants in Neuchâtel, Switzerland. The authors followed 4 principles[8] and studied 5 types of structures (1991, p. 53-58) in which L2 (French) had influence the L1 (Spanish).

1) Prepositional phrases, e.g.*, 'fuimos de vacaciones* ***a*** *España /* ***en**** *España'*
2) Direct objects, e.g.*, 'El león quería morder* ***al*** *hombre/* ***el**** *hombre'*
3) Infinitive verb phrases, e.g.*, 'decidió de* llamar al médico/ decidió llamar al médico'*
4) Highlighting[9] e.g.*, 'es mañana* ***cuando**** *llega mi hermano /es mañana que llega mi hermano'*
5) Semantic borrowing: '*no entiendo* el sonido del tren/ no escucho el sonido del tren'*

These rules were either resistant or flexible to change according to the 4 principles. Thus, a rule in the L2 would be reenforced over the equivalent rule in the L1, e.g., when a rule in L1 has multiple exceptions, is ambiguous or hard to apply in different contexts, then this rule cannot be resistant to change. In contrast, if the corresponding rule in the other language is easy to generalize, involves no ambiguities or exceptions, this rule would remain unaffected by change. This is similar to the Reduced Redundancy Principle proposed by Seliger (1991). This principle defines a process in which bilinguals combine elements from both languages resulting in a fused rule, then, the bilingual speaker adopts the less restrictive rule among both languages, hence reducing redundancy. This process is supported by linguistic data received via direct or indirect feedback from which the speaker extracts positive or negative evidence about his/her production in a particular language. Bilinguals would benefit from negative or positive evidence in both languages,

[8] These principles are: the number of rules; the generality, simplicity and clarity of the rule; the function of the rule and the type of norm in the grammatical system of the target language.

[9] As the translation of « *mise en rélief* » in French.

for example when there is restricted exposure in the L1, the L2 resources intervene to palliate the lack of information, resulting in L2 rules governing L1.

To come back to the study of Grosjean & Py (1991), these authors found differences in the mean acceptability rates between Spanish-French bilinguals and Spanish monolinguals, suggesting that the L1 competence of bilinguals was restructured and highly influenced by the L2, which was supported by the fact that deviant L1 structures were rated as frequently attested in the linguistic environment of the Spanish-French bilinguals who had been immersed in L2 for a long period of time (8 - 25 years with a mean of 16 years).

The 5 types of structures presented previously were analyzed following principles such as the number, generality, simplicity, clarity, and the function of the rules in both languages to explain the degree of L1 restructuring and L2 influence in the L1. Unfortunately, this approach was not successful to analyze semantic borrowing (5) because this phenomenon could not be matched with grammatical rules and degrees or ambiguity criteria as was the case for the other structures studied here. Semantic borrowing (equivalent to our definition of semantic extensions) will not apply explicit resistance, either respect or violate a particular grammatical rule in the L1 as for the other 4 structures.

The fact that semantic borrowing or extension involves lexical elements rather than grammatical rules, suggests that the lexicon might be more flexible than grammar since its use is not constraint in terms of grammatical rules, but by context of language use. Semantic extension in Spanish-French bilinguals is the main research interest in the current investigations (see Ahumada-Ebratt et al., 2018). Moreover, the method used by Grosjean & Py (1991) is highly interesting and in accordance with the research questions that were adopted in the present study.

Assuming lexical CLI to be motivated by similarities at the formal level (Singleton, 1999), semantic extension may be triggered by false cognates, such as in the Spanish example provided by Grosjean & Py (1991, p. 58) of '*entender*' in "*no entiendo el ruido del tren*" ('*I dont understand the noise of the train*'), following the structure of French '*entendre*' ('*to understand*') instead of 'escuchar' ('*to listen*'). However, it seems that the underlying processes are more complex than just the activation of formal similarities. Phonological overlap of the word forms does not seem to be sufficient to induce semantic transfer. The Activation Threshold Hypothesis and language competition have also been proposed to explain this phenomenon (Paradis, 1993) at the semantic level. Semantic overlap will probably play a significant role since differences have been observed during the processing of true cognates (e.g., Sánchez-Casas & García-Albea, 2005) and false cognates (e.g., Sánchez-Casas et al., 1992). The typology of cognates presented in **Table 1** illustrates that formal and semantic relationships across languages are indeed very diverse affecting differently CLI.

Following Jarvis, semantic extension would be the result of learned interlingual associations of lemmas, rather than the result of activation levels or processing

interference. It seems that the learner's proficiency level is a predictor of the direction of semantic extension, thus, L1 to L2 direction is more current than L2 to L1 (Ringbom, 2001). Besides, crosslinguistic similarities do not seem to be a factor explaining successfully semantic extensions because these are reported also between typologically unrelated languages (Ringbom, 1987, 2001, 2007).

Andersen (1983) assumes that semantic transfer occurs without any perceived crosslinguistic similarity, suggesting that semantic transfer contrary to formal transfer is defined by the lack of awareness of semantic differences between two lexical units. For instance, in the production '*he bites himself in the language'*, instead of '*tongue'*), transfer is motivated by the Finnish word '*kieli'* standing for the meaning of both *'language'* and '*tongue',* contrary to English.

1.2.1.3.2. Loan Translations

Also called calques, are simple or compound words, complex constructions or collocational words that are transferred from one language to another as a translation (Jarvis, 2009). In these translations, the meaning is imported, however the lexemes used to express it are "native-like" (Andersen, 2011). This means that the translation itself does not make sense in the target language because it is a literal translation, calqued or inspired from the other language: e.g., the use of '*animaldoctor*' for veterinarian in English is calqued from Finnish *elä inlä äkäri* literally translated as *animal doctor* (Meriläinen, 2006, p. 91). Loan translations can also involve collocational words, which are often expressions that are language specific and not transferable. Grosjean (2010, p. 60) proposes the example of the expression used by Hispanics immersed in Florida *"tener un buen tiempo"* calqued from the English "*to have a good time"*. In Spanish, the canonical translation of this expression would be the word *"divertirse"*.

This kind of transfer results in non-native like production because of language-specific semantic constraints, so the elements calqued do not match the standard form in the other language. Malt et al. (2003) affirmed that words or categories in one langue can be labeled in a subset of classes which are inexistent in the other language, for example in English there are '*fingers'* and '*toes'* but in Spanish there is a single label '*dedos',* that evokes both. To refer to *'fingers*' Spanish uses '*dedos de las manos'* and to refer to '*toes'*, '*dedos de los pies'*. In contrast, in Spanish *'walls'* inside a house are called '*paredes'*, walls around a castle are called '*muros'* and walls around a city are *'murallas'*, while in English all these are referred to with a single label: *walls*.

Thus, the links that underlie this type of semantic transfer are lemma to lemma associations and involve syntactic specificities as is the case of semantic extensions (Jarvis, 2009). Hence, following Jarvis, Calques and Semantic Extensions are supposed to be the result of interlingual or crosslinguistic learned associations, which excludes the possibility a theoretical frame based on levels of activation.

1.2.1.3.3. Collocational Transfer

There are different terms used to define collocations, such as fixed expressions, word-combinations, idioms, phrases, or prefabricated patterns (Wang & Shaw, 2008). Their definition encompasses two or more words that co-occur in the use of a language. Collocational transfer maps words that co-occur together and is manifested in word combinations that do not match in the other language, e.g., '*do children'* in Finish instead of '*have children'* in English (Meriläinen, 2006). So, lemma-to-lemma associations are transferred from the L1 to the L2's associations or the other way around (Jarvis, 2003).

This kind of transfer has received a lot of attention in the literature on transfer in SLA (Odlin, 1989; Biskup, 1992, Wang & Shaw, 2008). It has been studied using translation tasks, words association tests, written production, however, only few studies take into account natural language production.

As for the previous category, collocational transfer involves also the semantic level, as mentioned previously, it is considered as a continuum of loan translations (Jarvis, 2009). It has also been observed between languages that are typologically distant and which are not supposed to involve levels of activation of lexemes or processing interference, because collocational transfer (and others types of lemmatic transfer) are explained as the results of learned interlingual identifications that affects how lemma are constructed in the L2 (Jarvis, 2009, p. 115), which means that the lexical transfer is triggered exclusively by formal features (e.g., word-form, phonemes).

1.2.1.3.4. Subcategorization Transfer

For some authors, the previous categories are gathered together in one single type of transfer (this categorization depends on the definition adopted of collocations). Here we distinguish a fourth type of lemmatic transfer, i.e., subcategorization transfer because it focuses mainly on syntactic constraints. This transfer involves for example the specific syntactic principles that are associated with a verb, and that may be mistakenly transferred between languages, e.g., in *She kissed with him* vs *She kissed him* (Jarvis, 2009) where a reflexive form is transferred from one language to the other. As the other transfer types, subcategorization transfers are bidirectional that is, produced from L1 to L2 or from L2 to L1. Since the syntactic information is related to the lemma level, it is considered to be a lexico-syntactic transfer as well as lexico-semantic. To illustrate this, Helms-Park (2001) studied the causative in English for Hindi-Urdu and Vietnamese speakers whose task performance differed depending on the way causation was expressed in their L1.

1.2.1.4. Discussion

Following the studies of Jarvis (2009), lexical transfer is explained through two processes: processing interference or learned interlingual associations (e.g., association between a lemma and lexemes). Lexemic transfer (e.g., false cognates) is explained as the result of online interference processes, which means that the transfer is due to competition during the activation of word forms in both languages. In contrast, lemmatic

transfer (e.g., semantic transfer) is considered as learned crosslinguistic associations that would reflect underlying strategies to compensate intentionally or unconsciously the learning difficulties encountered.

However, word-form associations at the semantic levels may trigger the production of semantic extensions, suggesting that these might also be the result of formal competition across languages, just as false friends may be the result of processing interference. Assuming that associations are generally stronger between related languages, some examples reported in the literature suggest that semantic extensions may also occur more frequently between related languages.

Assuming semantic extensions to be due to lemmatic transfer only (and related to learned interlingual association processes) is ignoring that interference during processing may also play a role triggered by formal features. As suggested by Jarvis, lexemic transfer increases with perceived similarities; semantic extension as well engage formal similarities between languages such as in the example *go up the stairs* in Spanish and French (*montar las escaleras* vs *monter les escaliers*) proposed by Grosjean et Py (1991). Therefore, we suggest that lexemic factors can be an explanatory framework for semantic extensions under the scope of competition between lexemes. Thus, semantic extension as the result of interference processing is as plausible as the learned crosslinguistic association process explanation proposed by Jarvis (2009).

1.2.2. Observing Lemma, Lexemes and Crosslinguistic overlaps

In the previous section, a 'traditional' perspective of the study of semantic extension has been proposed, that is the lemmatic approach attributing semantic extension to the links between two or more lemmas in different languages. This network between the lemmas of each language is supposed to be the result of the bilingual's learned interlingual associations. However, in the present study the role of lexemic relationships between languages is one of the centers of interest considering that lexemic information involves formal aspects of words (e.g., phonological and graphemic information) which is supposed to be explained under the scope of activation and competition levels. At this purpose, Jarvis (2009) considers that lexemic transfer can probably be extended to the lemma level (pp. 106) but it is supposed to be driven by purely lexemic factors (e.g., false cognates). Keeping in mind that semantic extensions are formed by interlingual pairs of words, which include lemmatic and lexemic information, it can be considered that both approaches are head and tails of the same coin of the nature of semantic extensions.

Given the limitations of the study of semantic extensions (e.g. Grosjean and Py, 1991), lexical ambiguity is difficult to identify in the analysis of deviated productions. Taking into account the relevance of the Reduced Redundancy Principle (Seliger, 1991), this study proposes a combined approach that includes crosslinguistic associations at both lexeme and lemma levels through the control of two internal features of words known to

play a role in their accessibility during processes of lexical access. These will be controlled in our experimentations with the interlingual pairs presented and described in Chapter 3.

The first of these features is the Crosslinguistic Neighborhood Density that plays a role in the definition of the lexemic links between interlingual words. The interest of this dimension lies in the word-form overlaps existing between languages and their influence on semantic extensions. Second, the dimension of Crosslinguistic Morphology involves semantic links and etymological relations between words which correspond to the lemmatic levels of semantic extensions. The innovative aspect of this approach is that it would enable us to take into account competing activation levels which in order to expand the focus of analysis of semantic extensions goes beyond the perspective of learned association previously taken into account in order to explain semantic extensions.

1.2.2.1. Crosslinguistic Neighborhood Density

In this section we will describe Neighborhood effects in monolinguals and bilinguals focusing on studies on Crosslinguistic Neighborhood Density. As a departure point, we must ask what happens technically when a speaker visually processes a word. It is generally assumed that this word activates other words that are orthographically similar to the target word in memory. To illustrate this, when recognizing the English word *sand*, in monolinguals, orthographically similar words from the same language are activated, such as *hand, land* and so on (Grosjean, Li, & Bialystok, 2013:82). In the bilingual mental lexicon, the whole process is complexified.

The concept of lexical Neighborhood Density was introduced by Coltheart et al (1977). Grainger & Segui (1990) define neighborhood as the existence of physical similarities between words at the level orthographical and phonological information. We will use the term form overlap to account for neighborhood in general. An orthographic neighbor is a word that differs from the target word by changing a single letter, keeping the same position and word length (e.g., for '*passez': 'passer', 'casser'* in French). The number of neighbors of a word depends on the length and the position of the letters of the words that might be considered as orthographic or phonological neighbors. In the case of phonological neighborhood, neighbors differ with respect to a single phoneme, e.g., '*passes'* [p.a.s.e] vs '*pause'* [p.o.s.e] in French.

Neighborhood density of a given word can be calculated counting the number of orthographically or phonologically similar words that exist when changing one phoneme or one letter of a word. More precisely, the number of orthographic or phonological neighbors can be calculated through three types of procedures: by suppression, substitution or addition of a letter or phoneme. For example, the word *cat* declined in English has a high number of neighbors such as: *pat, that, mat, chat, sat, cut, cot*. Hence, *cat* could be considered as a word with dense orthographic neighborhood in English. In contrast, the word *cry* counts only few similar words, such as *fry, try, dry*, so *cry* has a sparse neighborhood (Costa et al., 2006, p. 144). As can be seen through this example, the meanings of the lexical neighbors may probably be unrelated.

Coltheart et al. (1977) and Andrews (1989) suggest that monolinguals are sensitive to the number of orthographic neighbors of the target words and their frequency in word recognition as well as word naming. In bilingual speakers, interlexical or crosslinguistic word pairs share orthographic forms and phonological features with their counterpart language especially when both languages use the same alphabetic system[10] , which implies that a word pair can be considered as phonological and orthographical neighbor at the same time (de Groot, 2011).

Neighborhood density effects in crosslinguistic studies (Grainger & Dijkstra, 1992; van Heuven et al., 1998) suggest that the size of L2 neighborhood density has an effect on activation of the other language (L1) and vice versa. To illustrate this, a monosyllabic word in Spanish such as *mil* (thousand) has few neighbors in Spanish (e.g., *vil, mal, mis*), so *mil* has a sparse neighborhood in Spanish but it has a dense neighborhood in English (e.g., *kill, chill, gill, bill, till, miss*) (example from Costa et al., 2006, p. 145). Crosslinguistic neighborhood density effects can facilitate lexical access in one of the languages only whilst for words with few neighbors or disperse density such benefits will be smaller (Costa et al., 2006).

In the context of semantic extensions, taking into consideration that both languages are activated simultaneously (Bice & Kroll, 2015), crosslinguistic neighborhood density competition levels are illustrated in **Figure 11** in which other words sharing the same formal links while being semantically unrelated are co-activated in L1 and L2.

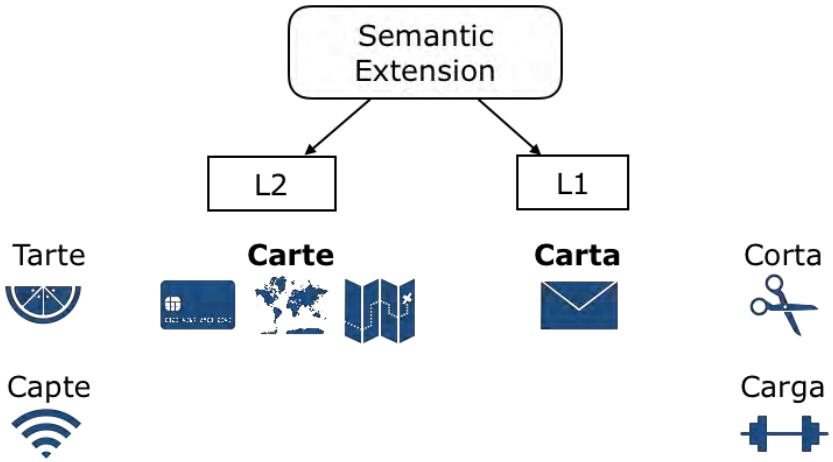

Figure 11 Illustration of interlingual words and their orthographic neighbors in French and Spanish

'*Carte*' in French means '*credit card*' or '*map*', contrary to '*Carta*' in Spanish means '*letter*'. A production containing this example has been observed in the form of "*se me olvidó*

[10] For further information about crosslinguistic effects of inter-alphabet of Greek-French Bilinguals see Voga-Redlinger (2005).

*la carta*del metro*' (I have forgotten the metro card) instead of "*tarjeta del metro*". This semantic transfer comes from the L2 '*carte de métro'*.

An approach taking into account Neighborhood Density is interesting since it allows researchers to explore whether bilingual word recognition in one language will be interfered by or activate linguistic features in another language. In this respect, Grainger and Dijkstra's (1992) study concludes that neighbors influence the non-target language in bilingual performance suggesting that bilingual's word recognition processes involve the activation of both languages systems, even when a part of the linguistic information is treated unconsciously (as in masked priming techniques). The visual word recognition paradigm suggests that orthographic neighborhood is also influenced by the frequency of the words which are part of the neighborhood and that are triggered by the target word (e.g., during a lexical decision task). Hence, if *lame* has 17 orthographic neighbors and some of these have high frequency, then the frequency effect of *lame* will be even more important. Interlingual competition of word representations between the non-target language and the selected language support the idea of an '*integrated lexicon of bilinguals'* (Brysbaert & Duyck, 2010).

Neurolinguistic studies argue for activity in the prefrontal cortex of the brain during processing of homographs, this suggests that bilingual regions of phonological and semantic processing are simultaneously activated leading to interferences (van Heuven et al., 2008). A detailed review of brain imaging studies, see Abutalebi & Green (2007), pointed out that lexical access in the bilingual mind involves control processes rather than language processing. They suggest that a network of structures is active while the bilingual uses one of their languages during linguistic tasks. This network is associated to cerebral areas that are responsible for cognitive control such as executive function, decision making, response inhibition and so on. Following this perspective, the role of homographs in the bilingual memory is probably related to high control levels.

Regarding Phonological Neighborhood Density, studies suggest that the first letter will have salient phonological attributes that allow the speaker to resolve the retrieval process through an initial phonological frame that will trigger retrieval of the following segments in comparison to other syllables or letters positions (e.g., Fay & Cuttler, 1977; Garrett, 1984). Considering this, it can be concluded that the phonological lexicon has a salient structure that conditions activation processes. From this perspective, it seems interesting to deepen our understanding of how this phonological lexicon influences word retrieval and the processing of cross-linguistic words such as false cognate, or words with word-form overlap between languages as suggested by Brown & Knight (1990). It seems that salient phonological attributes facilitate the access of other lexical units. For example, if the beginning of the word is *go* for the retrieval of *goat*, this will probably activate *goal*, *gout, goes, and so on.*

To conclude, the bilingual lexicon shows sensibility to Neighborhood Density across languages suggesting that both languages are interconnected at least at formal levels of representation. We could assume that lexemic links are constructed integrating lexical

resources providing from more than one language. Hence, neighborhood density is a factor that will allow to study semantic extensions in terms of formal or lexemic transfer, suggesting that during lexical retrieval, words with dense neighborhood are retrieved easier than words with dispersed neighborhood due to multiple accumulated sources of activation (Costa et al., 2006). Following this, semantic extension can be explained not only as the result of interlingual learned associations but also as a lexemic transfer in which multiple formal features are activated inducing CLI.

1.2.2.2. Crosslinguistic Morphological Links

The role of morphology in language processing is supposed to be determinant to define the structure of words, influencing semantic, syntactic, orthographic and phonological levels (Sánchez-Casas & García-Albea, 2005).

The study of morphological links is proposed here to illustrate co-activation processes at the lemma level when a speaker is confronted to the processing of semantic extensions in a given language. In contrast to orthographic neighborhood, words with morphological links are generally semantically related (e.g., morphological derivates for 's'*énerver*' in French are '*énervement', 'énervant', 'énervation'*. As mentioned above, morphological structure is very complex and follows different principles depending on the language concerned. Cross language studies are particularly interesting to explore the nature of lexical representations in the bilingual mental lexicon which seem to be defined by morphology (Lalor & Kirsner, 2001). The type of morphological link that the present research focuses on is Morphological Family Size, with the hypothesis that overlap in morphological family size will contribute to crosslinguistic competition (Mulder et al., 2013).

Morphological Family Size is defined as the number of morphologically related complex words in which given word-stem occur as a constituent (De Jong et al., 2000). The latter study has found that Dutch words that are part of large morphological families were processed faster in a lexical decision task than words with smaller morphological family size.

Frost & Grainger (2000) centered their attention on studies touching cross-linguistic morphological processing, in which the replication of the previous study in German, English, Chinese, Arabic reported all for the same results. They argue that independently of the orthographic and phonological intersection between languages, morphology plays a role in processing. In the case of Finish, reading tasks involving analysis of eye movements have shown that morphological family size affects lexical access in word recognition processes of inflected nouns (Bertram et al., 2000). Results suggest that the level of familiarity and the frequency of a word affect early and late stages of word perception, as well as its root morpheme frequency plays also a role in visual word processing.

As put by Mulder et al. (2014, p. 60), "*The family size effect is observed to be predictive over and above other lexical proprieties such as word frequency, morpheme frequency, word length, orthographic neighborhood size, bigram frequency*".

The superiority attributed to Morphological Family Size over formal properties reminds us of the cognate effect (see section 1.2.1.2.1) suggesting that neither form nor meaning by itself offers conclusive account of facilitation effects. The idea that morphological relationships across languages determine the way cognates are represented in the bilingual lexicon is justified by the role that morphological information plays in lexical representations and processes. At the same time, morphology defines the structures of words affecting others levels such as semantics, syntax, and phonological/orthographic properties (García-Albea et al., 1996; Sánchez-Casas & García-Albea, 2005).

Forster & Azuma (2000) suggest that morphology effects happen independently from form similarity. This is the result of the analyses of priming effects between semantically transparent morphologically related words (e.g., *fold-unfold*) sharing bound stems (e.g., *submit-permit*) with semantic opacity, and compared with an orthographic control condition. After reducing orthographic overlap, priming effects of semantically transparent or opaque words remained, which suggests that pure morphological effects do not include form overlap.

Possible limitations of the traditional methods of research may lead to inconclusive results regarding form-meaning dichotomy. Recent ERP studies corroborate the distinction between orthographic neighborhood and Morphological Family Size. Müller et al. (2010) suggest that cross-linguistic family size effect is a stronger predictor than neighborhood size because the morphological links include also semantic representations of the family members.

Mulder et al. (2013) studied the effects of morphological family size in L1 and in L2 in Dutch-English Bilinguals using ERPs to observe whether activation levels of the non-selected language (at the lemma level) are restricted or spread to the morphological family of the lemma concerned. Results pointed out that ERP signals are more sensitive to lemmas with a high family size than low family size lemmas in L1 Dutch. Hence, a larger number of family members converging semantically with each other would facilitate lexical processing. When exploring crosslinguistic effects, these researchers suggest that L1 family size affects L2 family showing faster RTs responses for high family size stimuli than low family size. Thus, activation of the lemma would spread to the family members of the non-selected language.

The effects of the family size would additionally depend on the closeness of the family members. The activation of a morphological family spreads more easily to immediate family members than more distant family members, because the first group is formed by words that are related not only in form but also in meaning. The primary family contains all derived words including compound nouns (e.g., 'horse' – 'horsefly'), the secondary family are two-constituent compound words, e.g., 'horse' would activate

'horsefly' and then secondary family activation would be 'fly' and 'flypaper' (see figure 1 in Mulder et al., 2014, p. 61). The first kind is semantically related to the target and the second is not always semantically related (Mulder et al., 2014). Considering this distinction in our methodological approach will allow us to restrict the morphological family composition to semantically related words only. In other words, selecting primary family members of the target word assures the exclusion of "opaque family members" such as *honey-honeymoon* as described by Mulder et al., (2014, p. 61). Unfortunately, there are few studies associating SLA research and the question of how bilinguals start to develop morphological and semantic relations between L2 words. However, we do know that the construction of bilingual's L2 primary family might differ from the way it is constituted in the L1 which is explained by the specificity of each language. Following these points of view, it can be hypothesized that cross-linguistic morphological learning is a process of interlingual associations. This process might be triggered by an overgeneralization or simplification of rules of two morphological systems.

1.2.3. Chapter Summary

This Chapter focused on current models of the bilingual lexicon on the one hand, and on the other, on the different types of interactions between the knowledge of more than one language with specific attention to the lexical level. The models of the bilingual mental lexicon we reviewed were constructed progressively over the years and the modelling of the place of conceptual and lexical representations has shifted from independent to distributed/shared connections between both levels, but they vary also depending on different theoretical positions and methodological approaches. The issue of of language transfer and restructuring at the conceptual level has received specific attention in these models, leading to the recognition of the dynamics and flexibility of the bilingual's language systems. We propose a theoretical overview of CLI particularly focused on the lexical level to illustrate two distinct levels of transfer: the lexeme and the lemma level. In order to investigate their role, we propose to take into account two psycholinguistic features: Neighborhood Density and Morphological links across language as a valuable perspective to study cross-linguistically lemma and lexeme levels.

2. **Dynamics of the Bilingual Lexicon**

As observed in the previous Chapter, the acquisition of a new language influences the knowledge of other languages at all structural levels: phonetics, lexicon, semantics and syntax. Regarding the lexicon, in bilinguals, interaction between multiple lexical systems attests for language flexibility, involving adaptation processes (such as weakening and strengthening of links between concepts and lemmas or lexemes). In this Chapter, a theoretical framework of bidirectional crosslinguistic influence (CLI) is proposed with focus on the bilingual processes involving the knowledge of two languages. At the same time, the dynamics of the bilingual lexicon is analyzed, for example, in some cases naming patterns of bilinguals in both languages exhibit differences compared to monolingual standards (see Ameel, Storms, Malt, & Sloman, 2005). Indeed, to a certain degree, both languages exert some influence on each other until languages deviate in a new, enriched linguistic competence instead of affecting negatively both linguistic systems.

First, in section 2.1, CLI is overviewed through some studies to illustrate how language traffic not only occurs from the L1 to the L2 (as traditionally observed) but also from other languages to the L1 (L3 to L1, and L2 to L1). In this regard, types of transfer (forward or backward) at multiple levels have been distributed into two main categories: phonological and lexical transfer, special attention being given to CLI at the lexical levels since it constitutes the main focus of the current research. More specifically, the literature centered on bidirectional CLI is developed in section 2.1.2 through two central questions: what bidirectional transfer is and why it does occur. Second, factors affecting the L1 during L2 language acquisition are discussed, such as language use, language dominance, proficiency, and language immersion in an L2 dominant context. In section 2.2.1. we will discuss studies of first language attrition focusing on lexical attrition and first language restructuring, as a perspective on long-term effects of bidirectional transfer. Then, in section 2.4 we discuss the concept of convergence and the role played by structural similarities across languages in the dynamics of the bilingual lexicon. In the final section 2.5 we propose multiple interpretation for a better understanding of the semantic extension phenomena in bilinguals.

2.1. Bidirectional Crosslinguistic Influence

Bidirectional CLI is defined by Pavlenko and Jarvis (2002) as the two way interaction between the linguistic systems of the L2 user. Following the bidirectional perspective, facilitation and linguistic transfer occur equally in both directions (Brown & Gullberg, 2011; Gass & Selinker, 1992; Malt et al., 2015; Pavlenko, 2000). During early stages of second language learning, the L2 naturally shows evidence of traffic arising from the L1 to the L2, then, after development and current use of the L2, the L1 also shows to some degree linguistic traffic from the L2. As described in **Figure 12** CLI touches

multiple linguistic levels including phonology/ phonetics and the lexicon as described in section **Error! Reference source not found.** and 1.2.1, but also grammar and pragmatics.

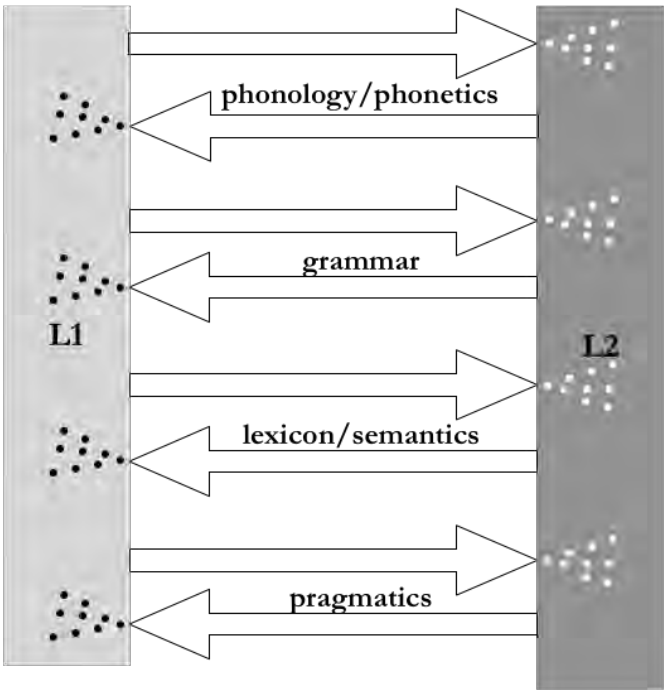

Figure 12 Integrated view of bilingualism (Schmid & Köpke, 2007, p. 3)

Bidirectional CLI has been observed for instance on oral narratives of Russian (L1) English (L2) late bilinguals who exhibit transfer in both directions (Pavlenko & Jarvis, 2002), especially regarding semantic extensions, lexical borrowings and loan translation. The latter study suggests that L1 restructuring is observed not only in the areas of semantics as expected but also at unexpected formal levels such as sub-categorization transfer (see section 1.2.1.3.4), which involves transfer at the grammatical level. This result suggests that L1 restructuring may involve lemmatic transfer with respect to syntactic and semantic constraints.

Jarvis & Pavlenko (2002, p. 198) proposed three paradigmatic [11] categories intervening in CLI:

1) The first category involves linguistic frames, as in the expression of motion events. For example, speakers of Spanish express direction using verbs such as '*entrar*', while English speakers express direction using prepositions such as '*go into*' (see Hohenstein et al., 2006; Hohenstein et al., 2004 for motion event description on Spanish). Following Slobin (1993), crosslinguistic difference of linguistic frames influences the way bilinguals conceptualize ideas in each language, especially when both

[11] By paradigmatic category or dimensions, we refer to a set of structural or linguistic units that compose or form a definable category. For example, the possible inflections a word can adopt.

languages allow both possibilities because this implies that the speaker has to select one of both structures.

2) The second category is word choice, which is subsequent to the linguistic frame category and refers to the selection of a word within the linguistic frame category (1). Following Pavlenko & Jarvis' (2002) perspective, word choice entails the selection of a word (e.g., a verb among multiple competitors of '*entrar*'), here lexical accessibility plays a determinant role.

3) The third category is word inflection which refers to grammatical form, e.g., the conjugated form, or nominal or pronominal case, gender agreement markers relating to the words previously selected (Pavlenko & Jarvis, 2002).

These three categories have their specificities depending on the set of languages the bilinguals know and constitute a resourceful framework to study CLI. Word choice (3) is interesting because this category affects the bilingual lexicon in one of the languages or both.

Thus there is a narrowed relation between semantic extension and lexical borrowing with word choice category; bilinguals' word choice is determined by the way a person's knowledge of one language affects the choice of words in another language (Jarvis & Pavlenko, 2010, p. 88). This may be illustrated through the case of English-Spanish bilinguals who would choose Spanish based verbs to carry motion event descriptions (e.g., *'go', 'come', 'enter', 'cross')* rather than English verbs that carry manner information (e.g., *'run', 'walk', 'skip')* (Hohenstein et al., 2006). Word choice would determine in some cases the level of proficiency of the speaker or even language dominance.

It can be concluded that bilingual word choice is unique not only in the L2 but also in the L1 since the linguistic competence of the bilingual is enriched by two systems, considering the amplitude of the paradigmatic level (which includes linguistic frames, word choice and word inflection). Word choice transfer is a sphere that may be manifested in semantic extension and lexical borrowing. However, to some degree, this kind of transfer is superficial since bidirectional traffic may also be explained through deeper processing category levels such as linguistic frame construction.

Following Pavlenko and Jarvis' study, in the case of semantic extension (see section 1.2.1.3.1) is defined as "*the extension in the use of L2 words and expressions to include the meaning of a perceived L1 translation equivalent*" (Pavlenko & Jarvis, 2002, p. 200), in other words, when the meaning of L1 words follows the 'model' of L2 translations which results in a broader semantic range including the other language meaning. Semantic extensions involve the use of nouns, adjectives and a few verbs in both directions (L1 to L2 and L2 to L1). One of the productions analyzed in the oral narratives of this study was the use of the Russian verb '*vybrat'* which means '*to choose*' or '*to pick out*' which is extended in the L2 use to express '*to take out'*. In the case of lexical borrowing which includes the use of

phonological or morphological elements from one language to the other, the authors observed L2 to L1 influence in words with phonological and morphological adaptation of L2 words (English) into the L1 (Russian) such as in nouns (e.g., '*landlord*', '*appointment*'). What is more, for the syntagmatic category, L1 to L2 transfer is observed for L1 based loan translations which involved Russian figurative metaphors, e.g. '*deep inside herself*', stemming from the Russian '*uiti v sebia*' '*to go inside oneself*' (Pavlenko & Jarvis, 2002, p. 206).

With respect to the conceptual level, bidirectional influence can be manifested through the bilingual semantic accent and meaning nuances (discussed section 1.1.7.2), following Pavlenko (2009) which constitute an illustration of how the conceptual architecture of bilinguals is formed through different layers in which each language makes use of finer and deeper specificities to express a single idea between multiple possibilities. It can be said that in accordance with word choice that L1's semantic nuances affect the way subtle semantic distinctions are made, depending on how semantics are filtered through the L1 or the more dominant language.

In sum, bidirectional CLI is observable in bilingual language use at multiple linguistic levels, specifically during oral or written production showing evidence of interactions between both languages in the bilingual mind. This bidirectionality suggests that both languages are affected by bilingualism to the point that bilingual language use deviates from monolingual standards in the L1 and the L2. At this purpose, the speaker's word choice processes discussed above serve as an illustration of the complexity of conceptual settings that affect bilingual language use explained by the fact that bilingual users manage and juggle with two or more languages at the same time in one mind. The studies presented so far constitute key experimental clues towards the comprehension of bidirectional transfer and the narrowed relationship between word choice transfer and semantic extensions.

2.1.1. The concept of multi-competence

The concept of multi-competence is considered as the theoretical framework that explains the best bidirectional transfer and how languages interact during learning and acquisition.

The term of multi-competence is proposed by Cook (2003, p. 3) who suggests that languages co-exist in the bilingual brain, instead of being separate systems, they interact in permanence. The main focus of the concept of multi-competence is the relationship between different languages in the mind of the language user (L1 or L2). Cook is also one of the first authors that reconsidered the definition of native speaker and bilinguals.

On the one hand, he argues against the misconception that bilinguals are deficient in L2 learning by default (Kasper & Kellerman, 1999). In contrast, he supports that bilinguals should be measured by their achievement in being L2 users, independently of

the L1 native speakers' standards or language proficiency, thus, he introduces the term 'L2 user'. Hence, the L2 user differs from monolinguals in aspects such as:

1) The use of languages which involves different linguistic contexts.
2) The bilinguals' L2 knowledge in relation to the knowledge of the L2 native speakers.
3) Their L1 knowledge in relation to the knowledge of L1 native speakers.
4) The mental processes in each language.

Altogether the multi-competence framework is supposed to explain why bilinguals should be defined through the competence they achieve in being bilingual language users and not by arbitrary comparisons to native speaker standards of language use or high levels of proficiency.

On the other hand, Cook queried the concept of "*so-called native speakers*" and replaced it with the term '*monolinguals*', referring to a population that should be considered as exceptional[12] instead of being considered as the norm or standard of language use. From Cook's perspective, a bilingual is a person who uses an L2 on a daily basis to fulfill communicative needs and does not constitute an "approximation" of a native speaker or monolingual, who is someone who speaks and uses regularly the language he /she learnt in childhood.

Schmid (2011, p. 13) interprets Cook's concept of multi-competence adding an interesting perspective where multi-competence corresponds to the ability of bilinguals to switch and merge their languages, which at the same time they can be used separately and selectively. In the same perspective, Pavlenko & Jarvis (2002) suggest that the multi-competence framework is in line with the bidirectional transfer perspective because it recognizes that L2 language users have a compound state of mind that differs from two monolinguals in one (Cook, 2003; Grosjean, 1989, 2010).

Cook proposed a new perspective that represents the bilingual mental architecture as a dynamic system, in opposition to Weinreich's perspective of bilingualism (see section 1.1.2). Cook's perspective is reflected in the integration continuum represented in **Figure 13**, based on three types of language interaction: separation, interconnection, and integration. Following this perspective, an L2 learner can be placed at anyone of these types of language interaction, which means that for Cook there is no structural hierarchy or that a particular level would precede the other. In other words, this representation does not imply an established direction or evolution, hence the integration continuum would not apply to the whole language system (p., 9) (e.g., an L2 learner might show integration of his/her languages for the lexical level but independence or separation of phonological features for each language (Cook, 2003) , and these distinctions would depends on the

[12] Here 'exceptional' refers to the idea that being monolingual is reductive considering that half of the population is bilingual since most individuals in society are confronted to being second language users at least once in their lifespan.

individual specific characteristics. Indeed, this positioning describes bilingualism and interconnectivity of languages as a dynamic process which differs from other kinds of modeling that are static and limited to some types of interactivities at different degrees.

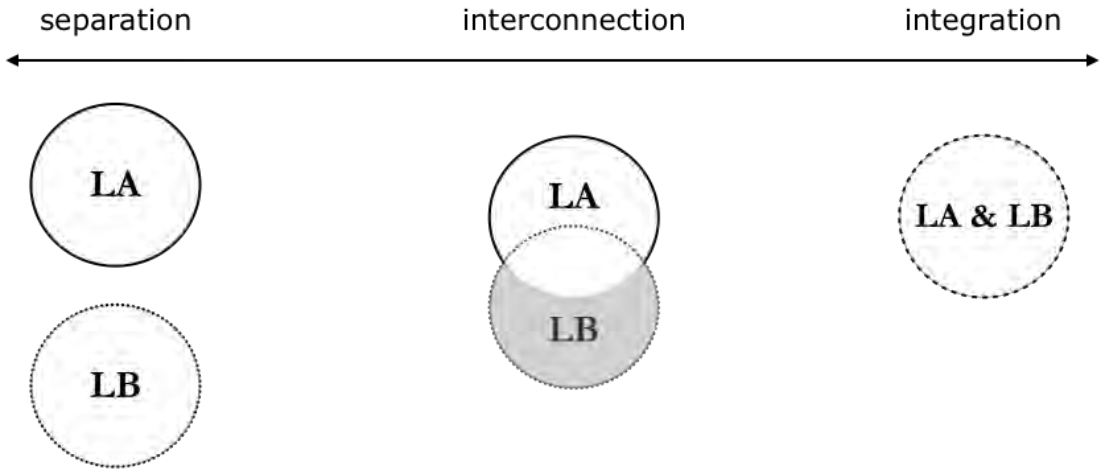

Figure 13 The integration continuum of possible relationship in multi-competence (Cook, 2003, p. 9)

It is worth mentioning that this is the first representation that takes into consideration three aspects: the individual differences of L2 learners, the non-linearity of developmental stages, and the flexibility, dynamics, and dissociation between linguistic levels. Separation, Interconnection and Integration could be undeniable factors that influence the organization of languages in the bilingual mind.

The multi-competence model can also account for other aspects that intervene in defining bilingualism and languages in interaction such as the frequency of the linguistic input. Letting aside the comparison of competence between bilinguals and monolinguals, Dewaele (2018) proposes to investigate variants of the same language (American English and British English) to show that small differences in the linguistic input and output influence users' multi-competence. For instance, when the use of some words in a particular language or variety [13] is reduced, this motivates changes in language use of North-Americans living in the United Kingdom who shown avoidance of the word "Jerk" (Dewaele, 2018). Indeed, factors such as the frequency or the nature of input and the languages activities affected the variety of English they mastered (American) in order to match with a new linguistic competence during the immersion in the other variant of English (UK).It can be concluded that if the interaction of two varieties of the same language exerts such an influence on each other in language use, in the case of bilingualism, the learning of an L2 and the L2 exposure would affect deeply the way the L1 is used in order to match L2 use standards. On this matter, an interesting perspective

[13] The term 'language variety' refers to a form of language that is distinguishable from another one by their use of words/verbs or even intonation and phonetic realizations (e.g. French variety in France vs French variety in Québec).

is proposed by Seliger (1991) in the context of bilingualism, in which an unlearning process together with a lack of accessibility of the L1 would result in growing L2 dominance.

Moreover, Dewaele & Pavlenko (2003) studied whether language and culture affects productivity and lexical diversity[14] in monolinguals and bilinguals. Researchers found that L2 users approached L1 users in both (productivity and lexical diversity) without changing their L1 use. They support Cook's (2001) positioning suggesting that multi-competence should not be seen as a fixed and static end-state, but as a dynamic and ever-changing process, which affects each individual differently. For instance, any language (L1 or L2) may easily be shaped and reshaped motivated by a change of environment, or changes in the linguistic input: such as frequency of use and linguistic activities (exposure to books, films, etc.).

To conclude, it can be said that the model of multi-competence is in line with the bidirectional transfer framework because it explains successfully how languages interact at multiple levels, and differently depending on:

a) individual characteristics of L2 speakers,

b) external factors (e.g., changes in the linguistic environment or language exposure).

This model defines bilingualism depending on the linguistic competences that L2 learners achieved at being language users, disassociating L1 standards from the process of second language learning. Additionally, it proposes a dynamic and non-linear view about linguistic interaction that made research evolve through richer perspectives that go beyond the traditional comparisons between the L1 and the L2 or the analysis of second language learning based on the L2 native speakers' patterns of language use.

2.1.2. Why does bidirectional transfer occur?

The objective of this section is to propose theoretical principles to try to understand why CLI or bidirectional transfer occurs[15] and what is their function in bilingual communication or language mode. To answer these questions, CLI will be examined through linguistic and psycholinguistic factors that may explain different types of transfer and the underlying processes of language learning/use. Special attention will be given to the external factors that influence bidirectional transfer.

---

[14] Briefly, by language productivity the authors refer to the total number of words produced orally and diversity refers to the level complexity that are measured using the calculation of tokens taking into account and word length and frequency of words.

[15] In the valuable review proposed by Schmid (2011) the term « CLI » is replaced with bidirectional transfer since most of the examples proposed here arise from the L2 to the L1. The positioning adopted is that there is no unidirectional transfer as stated by the multicompetence framework.

It is worth mentioning that in some linguistic communities the use of transfer is accepted at different levels. Attitudes towards transfer phenomena such as borrowing, word blending, or codeswitching are variable and it can be accepted as a natural phenomenon that may even carry a certain prestige or, conversely, be banned by the linguistic community. The attitude towards transfer depends on the community's bilingualism settings. For example, in contexts in which both languages are used simultaneously, speakers are more propitious to code switch and the relevance of crosslinguistic influence is recognized; in other settings, linguistic transfer might be banned by the speakers, so that languages can be "kept apart". Hence, extent and nature of bidirectional transfer depends on the community the speakers live in. However, in the present study, the sociolinguistic framework that explains attitudinal or motivational individual's or collective's reasons relative to the production of CLI will not further be developed. Instead, we focus on the role of L2 immersion (Length of Residence) and frequency of L2 use and level. We are also interested in exploring the processing strategies and the way the mental lexicon works, in order to respond to the question of the role of semantic extensions in late bilinguals.

In relation to acquisition, the linguistic context in which a language is acquired can set the way speakers access linguistic information and the way bidirectional transfer occurs. One determinant factor is the linguistic context in which the words are learned. In second language acquisition in an L2 dominant context, some words are associated with specific meanings that fit better in the L2 environment than in the L1 linguistic knowledge or environment. Schmid (2011) illustrates this case analyzing her data of two German-Jewish descendants who immigrated to an Anglophone context for about 60 years during the Nazi holocaust. One of her subjects "borrows" the word '*boys'* during an oral task in German that reflects one of the first words the participant learned during school years in England and that corresponded better to the L2 context than the equivalent word in the L1. Through this example is illustrated the impact of the linguistic context on language use and the production of CLI.

In the same line Schmid, (2011) proposed that in some cases specific words learned for the first time in an L2 dominant context may have no previous correspondence in the L1 lexicon that matches with the semantic information in the L2. Hence, social configurations — such as an active work environment in the L2, the lexicon associated with professional life that has been encountered and learned for the first time in an L2 dominant context — are more likely to trigger interference since this type of vocabulary is L2 context-specific.

Moreover, bidirectional transfer can fulfill a linguistic need, such as expressing some concept or idea that is inexistent in one of the languages. This kind of interference is complemented by disfluency markers including:

a) hesitation markers (e.g. *repetitions, silent and filled pauses*) and/or

b) repair words, which serve to interrupt formulations of speech in order to modify them (e.g., *approximation repairs)* used to adjust previous production for better comprehension,

c) *error repairs* that are lexical, phonological, or syntactic errors and

d) *cover repairs* which are adjusted or corrected before articulation (Lickley, 2015).

Essentially, disfluency markers are associated to stages of discourse planning such as word retrieval in the linguistic repertoire (for a review on disfluency in typical and pathological speakers see Pistono, 2017). In a bilingual context, Schmid (2011) illustrates the borrowing of the word '*involved'* in German (by German speakers living Canada) followed by disfluency markers suggesting difficulties while retrieving an equivalent L1 term.

In other cases, the L1 is no longer updated after a long period of immersion in an L2 dominant context, as shown by the use of terms related to a specific topic that is exclusively used in the L2 since the equivalent L1 word represents a past and outdated form[16]. Pavlenko and Malt (2011) suggest that there are crosslinguistic differences in the structure and boundaries of linguistic categories as household or drinking containers (e.g., '*Chashka*' in Russian corresponds to a narrow category which defines small cups for hot liquids, in contrast '*cup'* in English is a large category that includes plastic, paper and measuring cups and so on). Naming and word knowledge tests in Russian-English bilinguals have shown that English (L2) affects Russian (L1) naming patterns.

A similar example is proposed by Schmid (2011) for the case of household or commodities vocabulary (e.g., '*Kühlschrank'* vs '*Fridge*') that involve generational differences. Considered altogether, this kind of analysis reflects that there are multiple strategies underlying bidirectional transfer and that in most cases their production is related to L1 inaccessibility (Sharwood Smith, 2019) rather than due to formal similarities.

To conclude, the factors that explain the functional purpose of bidirectional linguistic traffic may be of different nature, including:

a) the order of acquisition that influences the way concepts are linked to a particular language

b) language-specificity at the conceptual level, with certain concepts corresponding better to one linguistic context (L1 or L2), and

[16] In my particular experience after living for only three years in France, at my first visit to my native country, native speakers did recognize me as a native speaker of the local dialect (Colombian Caribbean Spanish), however they rapidly noticed that i did not understand some expressions or words that were recently adopted in the colloquial language. This is an illustration of how fast languages evolve and change in a monolingual context.

c) the level of integration of languages (L1 or L2) and the way the concepts match the current linguistic environment or reality.

In the following sections four determinant factors will be discussed to address the initial question of why bidirectional transfer occurs: language use, language dominance, language proficiency and the linguistic environment in which bilingual are immersed.

2.1.2.1. The Role of Language Use

Language use is as determinant for language maintenance as it is for language acquisition, which may be problematic for a bilingual since the active practice of a new language sometimes engages the decrease of language use of another language. Adult bilinguals can be compared to bilingual children who favor to speak the language that is more 'useful' in immediate situations, in most cases, the dominant language of the environment (Grosjean, 2010). The need to use a language is a determinant factor in defining bilingualism: the more a language is needed on a daily basis, the higher it is the fixation level and maintenance of that language.

Learning an L2 is a long process, however, once the speaker has learned an L2, the following stage is language maintenance, which is as important for the L2 as for the L1. Language maintenance is related to language stability in common language use, which includes psychological, social and cultural processes when more than two languages interact in the same linguistic environment (Fishman, 2012). Language maintenance has received a lot of attention in research concerning language attrition (e.g., Klatter-Folmer & Avermaet, 2001). Indeed, language use seems to be a more determinant factor for preventing from language attrition in comparison to the length of time of language exposure or the level of education of the bilingual. Hence, for example, for the immersed bilingual living in an L2 dominant context, the use of the L1 in the work environment seems to favor language maintenance and prevent from language attrition (Schmid & Jarvis, 2014). Following De Bot, (2001) language use is considered to be the most important variable in both acquisition and language loss. Thus, language use and the availability of linguistic elements needed for production are closely related. In that sense, use and availability constitute separate factors to be taken into account as explaining factors in language attrition and language shift. However, the role of language use and availability is still not supported by the empirical results on attrition.

Second language immersion sometimes goes along with a significant reduction in the use of the L1 to learn an L2 (especially when the L1 is a minority language), this is commonly related to linguistic isolation and limited contact with L1 users which can affect L1 maintenance. As discussed in section 1.2.1.3.1, about the Reduced Redundancy Principle, Seliger (1991) suggests that native speakers need regular feedback regarding the use of their languages to maintain the L1, in others words, a native speaker needs confirmation of the 'accurate' use of his/her native language in order to maintain a normative use of the L1. It may be assumed mistakenly that L1 expertise does not depend on negative or positive evidence of other interlocutors, and that linguistic competence

developed over the years (adulthood) should be quite resistant. Nevertheless, input in both languages (L1 and L2) constitutes a potential linguistic resource for evaluating 'language rules' during language use. Linguistic resources are linked to language use because if the L1 input is reduced, L2 evidence may come into play for the L1, and this phenomenon could explain 'deviant' productions in the L1 that follow 'the model' of the L2, especially for speakers immersed in an L2 dominant context.

This lack of exposure of the L1 would result in some cases in L1 changes, and it may be due either to the lack of 'activation' of knowledge in the L1 or related to the influence of the L2 on the L1. Changes in language use can be reliable to changes in the accessibility of lexical information, as shown by the dynamics of languages in interaction, especially in immigrant contexts. Thus, language use determines the way languages are preserved, as well as the number of contexts in which languages are used (see Birdsong, 2014). On this matter, Ecke and Hall, (2013) conducted a longitudinal study that focused on the dynamics of language use in a multilingual speaker and linked the changing patterns during word retrieval studying tips of the tongue (TOT) coming from five languages (L1 German, L2 Russian, L3 English, L4 Spanish, L5 Portuguese) in different linguistic environments. The authors concluded that bilingualism over the lifespan is characterized by stages of instability and stages of balance, thus, accounting for dynamic systems that fluctuate constantly, including the L1 which regains stability rapidly. In the same line, it has been shown that in bilinguals who are currently using two languages, increased processing costs may lead to TOTs or retrieval failures (Gollan & Acenas, 2004) . As discussed in section 1.2.1.1 it has been proposed that TOTs are the result of a dysfunction between the lemma and the lexeme level, whereby the semantic and the syntactic information are accessible, meanwhile, formal levels remain inaccessible resulting in partial inaccessibility to the lexicon. As an illustration, bilinguals would use strategies such as giving a definition rather than the expected target word (e.g., describe the function of an object, or asking for the equivalent meaning in another language).

From this perspective, the regular use of the L1 would contribute to the maintenance of active linguist representations at both levels (lemmatic and lexemic). An explanatory framework linked to the complexity of language use and their relationship to language maintenance and language change (L1 or L2) is the usage-based perspective of language entrenchment. This phenomenon is defined by the strengthening of linguistic representations triggered by the intensity of language processing, which is particularly enhanced through language use (Steinkrauss & Schmid, 2017).

However, how researchers can measure the amount of language (s) used by the speaker remains a complex issue. Language use in bilingual speakers is determined by the context in which the languages are employed, often referred to as domains (e.g., talking to family members or the language use in a workplace environment), this is the reason why sociolinguistic questionnaires are valuable in bilingualism research. For a more detailed description of these domains see section 2.1.2.2.

It is sometimes mistakenly assumed that language use is related to expertise, for example simultaneous bilinguals would benefit from more L1 and L2 use and expertise than late bilinguals. Nevertheless, the model of multi-competence advocates against this supposition because L1 and L2 use of bilinguals differs from monolingual L1 use, as well as their L2 use differs from monolingual's L2 use (Cook, 2003). Moreover, Hohenstein et al. (2006) have found that in late bilinguals, L2 use is impacted by the L1 in the same way the L2 affects the L1, as reflected in atypical results of late bilinguals, attesting for more bidirectional influence in both languages than early bilinguals. Late bilinguals are concerned by the restructuring of the languages, affecting or facilitating CLI at multiple levels, and this, independently of the expertise attained in a given language.

To conclude, the undeniable role of language use is determined by multiple aspects in bilingual research, including L1 and L2 language maintenance, preventing attrition. Its role is essential in acquisition, facilitating this way the stabilization of linguistic representations in the bilingual mind (e.g., language entrenchment).

2.1.2.2. The Role of Language Dominance

Language dominance indicates an asymmetry of skills or use of languages in comprehension, production, and lexical access. In fact, two defining axes compose language dominance: dimensions and domains (Birdsong, 2014).

Following Birdsong (2014), dimensions correspond to a set of linguistic skills in each language, including linguistic competence, production skills (e.g., fluency of speech, accuracy) and comprehension (e.g., lexical and morphosyntactic knowledge). The comparison between the language abilities in each language will predict dominance (e.g. the language in which the bilingual is most proficient in). The other axis is defined by the domains of language use, it refers to a set of activities related to a particular language such as: listening to music, watching TV, interacting in the working place and at home. The calculation of dimension and domains constitutes a resourceful information to define language dominance, the assessment of dominance for each language would point out an index of dominance as put forward by Treffers-Daller & Silva-Corvalán (2016).

Language use would function as an indicator to measure language dominance, for example, the amount of language contact (L1 or L2) allows to differentiate the domains in which a particular language dominant, this may be also crucial while studying language attrition (Hulsen, 2000, p. 22). Language exposure is a relevant factor in defining language use because it determines not only active use in different domains but also shows more or less passive types of language contact such as media input.

Nevertheless, it is important to mention that bilinguals are gradually dominant in different dimensions and domains, which implies that language dominance does not display a cut-off criterium. On the contrary, dominance is distributed gradually and in different degrees, for example the comparison of two bilinguals dominant in the same languages would be different at multiples degrees (Birdsong, 2014).

Several authors have pointed out the misleading assumption of considering the L1 as systematically equal to the dominant language since a number of external factors intervene to define a particular type of bilingualism and dominance. It can be considered that language dominance switches when the linguistic environment changes, so, the use of languages increases or decreases. From this perspective, languages should be considered as changeable and dynamic systems that are subject to restructuring, to forgetting, shift or replacement. Dominance has been shown to play a direct role in lexical access, for instance, Abutalebi & Green (2007) have pointed out an asymmetric switching cost in dominants bilinguals, in other words, it took longer to process linguistic information when they switched back to their dominant language after producing words in their less dominant language. In contrast, when they switched to the less-dominant language costs were inferior. This asymmetric switching cost is interpreted in terms of inhibition processes, in this case, the non-selected language needs more inhibition when it is the dominant language of the speaker (Green, 1998). In the context of trilingual (Aparicio & Lavaur, 2018) it seems that a larger switching cost in lexical decision task is observed when the primes are presented in the non-dominant language condition, suggesting that language dominance plays a determinant role during early language recognition.

Goral et al. (2015) compared dominant and balanced bilinguals in terms of language use and language proficiency and examined different executive functions: 1) inhibition 2) alternating attention and 3) working memory. Among the multiple findings of this study, we will focus on inhibition, assessed with a Simon task[17] pointing out that balanced bilinguals display a higher Simon effect than dominant bilinguals. In contrast, dominant bilinguals (in terms of language use and proficiency) display faster and constant responses, that is, a smaller Simon effect, which suggests that dominant bilinguals benefit from the enhanced need of inhibition, resulting in better control abilities than in balanced bilinguals. In other words, it is harder to inhibit the language, which is used mostly, because it is more *available* than the other languages. There are two innovative aspects in Goral et al.,(2015)'s study . First, the great scale of participants tested in this study (106 participants) which is more representative of the population. Second, the choice to compare bilinguals with bilinguals instead of monolinguals.

The concept of dominance should not be restricted to linguistic competence in writing, reading, or speaking a particular language but also to daily language use. The study of domains (context or activities) and dimensions (linguistic skills) define language dominance regardless of the first language learned which is not the dominant one by default. Moreover, language dominance seems to influence inhibition processes and at some degree the production of CLI, hence, the direction of interference could be defined or predicted following the domain that is more related to a particular language, for

[17] Simon task was based on Simon & Wolf, (1963) is which participants have to press one key (left or right) depending of the stimulus presented within a square color (e.g. blue or red) for example half of the stimulus presented in a red square should be associated to a right key (in Goral et al., 2015).

example a bilingual whose professional life was constructed in L2, talking about work in another language would probably trigger CLI.

2.1.2.3. The Role of Proficiency

Language proficiency is associated to linguistic standards and norms that are calqued on monolingual use of a language. Birdsong (2014) proposes that proficiency is one dimension of dominance that is independent of other dimensions. For instance, a bilingual can be proficient in oral comprehension in L2 and have a lower level of performance in L2 writing compared to the L1. However, the language that the speaker uses mostly in daily life is still the L2, regardless of the levels he/she displayed in writing.

Moreover, it should be considered that the proficiency levels achieved by bilinguals (in their dominant or their non-dominant language) may differ from the proficiency levels displayed by monolingual speakers of each language of the bilingual. As it has been discussed in section 2.1.1, bidirectional transfer affects both languages, resulting in L1 and L2 linguistic competence distinct from monolinguals. On this matter, Jarvis (2019) proposes that high levels of proficiency in the L2 have implications for the accessibility of the L1.

L2 effects on the L1 have been observed in the literature in bilinguals with high levels of fluency and proficiency (Flege & Eefting, 1987; Major, 1992), however only few studies have found CLI at the beginning of language learning, which is probably due to the fact that few people have studied this period. Recent studies (Chang, 2019) on phonetic drift - which is defined as phonetic changes in the L1 due to recent experience with an L2 - suggest that even minimal L2 immersion in intensive L2 courses (5 weeks) may induce phonetic drift in the L1 of beginning L2 learners. Similar results are reported by Kartushina et al. (2016) who observed L1 phonetic drift after only one hour of intensive training in foreign vowels. Nevertheless, other research does not support such results, e.g., Lang & Davidson (2017) question L1 restructuring as the result of L2 experience as proposed by Chang.

In contrast, Bice & Kroll (2015) argue that the L2 influences the L1 also in early developmental stages of L2 learning, in other words, being highly proficient in the L2 is not a requirement for L1 change or L2 influence on the L1. Behavioral and Event Related Potential (ERP) data during a lexical decision task (in L1 English and L2 Spanish) show that an emerging N400 is observed for cognates in L1 at the beginning of SLA processes. This is interpreted as a strong argument for co-activation in early stages of SLA during L1 processing of cognates, suggesting that L2 word-forms are activated during L1 processing.

Regarding linguistic transfer and its relation to the learner's language proficiency, the source language is generally supposed to correspond to the more proficient language, for instance the L1. But the role of proficiency in these processes is questioned by Pavlenko & Jarvis (2002) arguing that semantic transfer occurs independently of

proficiency level of the source language. In the case of semantic transfer, nevertheless, a high proficiency level involving language-specific meanings would be a determinant factor in semantic transfer, In the case of multilinguals, during semantic transfer (termed by the authors as meaning-based transfer) the background language in which the speaker is the most proficient in tends to play the role of source language. This suggests that semantic transfer is to some degree determined by high proficiency levels (Lindqvist, 2012).

To sum up, proficiency levels of bilinguals would differ depending on the linguistic skills involved (e.g., written vs oral proficiency). They also differ from those of monolinguals, specifically with respect to accessibility (e.g., high proficiency levels in one language would affect that accessibility in the other). Moreover, the literature has illustrated two extremes of the effect of proficiency levels on the production of CLI. In some cases, high levels of proficiency do not seem to be a requirement for CLI, and in other cases, high levels of proficiency seem to be necessary for the arising of semantic transfer.

2.1.2.4. The Role of Language Immersion

L2 immersion takes place when an immigrant moves to a new linguistic environment different from their L1 environment. In most cases, L2 immersion is characterized by a reduction of input in the L1, which may lead to changes in the way linguistic knowledge/memories are accessed (Schmid, 2011). Besides, as mentioned previously, the lack of exposure reduces also the positive evidence related to the correct use of the L1 provided by native speakers of L1 (Seliger, 1991). Hence, the change of the linguistic environment influences not only the use of the L2 but also of the L1 any other language previously learned by the speaker. This point will be developed later on.

Linck et al.'s (2009) research suggests that learners immersed in an L2 environment have to inhibit their L1 both in production and comprehension. In their study, two groups of English (L1)-Spanish (L2) bilinguals in two different linguistic environments (L1 dominant and L2 dominant context) were tested in two linguistic tasks: a translation recognition task and a verbal fluency task. The translation recognition task involved two conditions: (1) lexical neighborhood (e.g., '*cara-card*') and (2) semantic neighborhood distractor interference (e.g., '*card-head*'). The results showed that learners immersed in an L2 context were insensitive to perceptual formal overlap in (1) but sensitive to (2). These findings suggest that these immersed learners strongly inhibited their L1 so that they remained unaffected by formal overlap in (1), and that they processed L2 in a deeper way than their non-immersed peers (2). The Verbal Fluency task[18] confirmed the presence of L1 inhibition because learners immersed in an L2 context produced more L2 exemplars and fewer L1 exemplars than their non-immersed peers.

[18] The verbal fluency task used in this study consisted on providing as many exemplars (word) as possible of a given semantic category (e.g., body parts) during a limited amount of time.

Following the explanatory framework proposed by Linck et al. (2009), bilinguals would have reduced functional frequency of lexical representations in each language compared to monolinguals, resulting in weaker links between these representations (see section 2.3. for further information concerning the *weaker links hypothesis*). Likewise, living in an L2 immersion context implies a reduction of L1 and an increase of L2 use, entailing a reduction of the L1 functional frequency. Following the Revised Hierarchical Model (Kroll & Stewart, 1994; Kroll & de Groot, 1997), Linck et al., (2009) suggest that the insensitivity to lexical neighborhood distractors found in the results of their study could be explained by the hypothesis of L1 inhibition in an L2 immersion context favored by stronger lexical-conceptual links, reflecting effects of resistance to L1 lexical competition.

With respect to the reduced accessibility of the L1 in bilinguals, Gollan et al. (2008) suggest that the shared frequency of language use over time between two languages influences the functional frequency of use in the L1, in comparison to monolinguals. This means that the languages in use are divided between different domains while a monolingual uses one language in every domain of his/her daily life. This shared functional frequency is supposed to affect especially the retrieval of low frequency words in the L1. An additional factor that may have an incidence on L1 accessibility in an L2 dominant context is the L2 mental set. Learners shift their mental set inhibiting constantly the L1 and making the L2 more available during the performance of any linguistic task that involves the L2 (Pavlenko, 2002). On this matter, the study of Pavlenko and Malt (2011) has found that after a short exposure to the L2 context, Russian-English late bilinguals show L2 influence on the L1 when naming drinking containers task.

The research of Link et al. (2009) is relevant to our study because it refers to the effect of immigration experience on L1 inhibition at two different levels: production and comprehension. Indeed, Linck et al.'s study focuses on two main principles that have been rarely studied before: language use and language dominant contexts suggesting that external factors such as changes in the linguistic context and configuration play an important role affecting language use and the accessibility of L1 elements in order to facilitate L2 acquisition and use of L2.

2.2. Language Restructuring and Language Attrition

In this section we will discuss some main differences between language restructuring and language attrition. These two topics are considered here as part of the same continuum and discussed essentially through the perspective of lexical change in bilinguals. Language restructuring and language attrition in the context of SLA affect equally the L1 and the L2, here we will focus on L1 attrition and restructuring. For an introduction to and review of L2 language attrition see Mehotcheva & Köpke (2019).

Special attention is given to the lexical level because it is most susceptible to language change (Köpke & Schmid 2007). Besides, it is the first system to be affected by linguistic change, and the one investigated mostly compared to other linguistic systems,

such as phonetics (e.g., Major, 1992) and grammar (e.g., Köpke & Nespoulous, 2001). This vulnerability of the lexical level is justified by its nature: an open class system (Schmid, 2011), which is believed to evolve more quickly than other systems (e.g., diastratic variations [19] such as generational lexical change). An illustration of the permeability of the lexical level is the high level of lexical borrowing across languages, as in French anglicisms such as *'spoiler'*, *'break'*, *'people'*, even though the inclusion of new words is limited by lexicographic policies.

2.2.1. What is Language Restructuring?

This section is aimed at defining language restructuring and proposing a new perspective to explain underlying processes and external factors playing a role in L1 restructuring during SLA.

Language restructuring is a process in which the bilinguals re-analyze the L1 system in accordance with the corresponding rules of the developing L2 system. In this process the existing L1 knowledge is not adapted or modified into the recipient language, but instead it gains a different value (Schmid, 2011, p. 27). For example, Russian (L1)-English (L2) bilinguals associate English concept of *cups* with the existing conceptual representation in Russian *'chashka'* and *'stakan' (glass)*. This association results in a restructuring of the categorization that switches from shape (e.g., 'chashka'/cup with handles and 'stakan'/glass withouth handles) to material (e.g., paper or glass) in the bilinguals' naming patterns of those objects to match with the English conceptualizations which differ from Russian monolinguals (Pavlenko & Malt, 2011) also discussed in 2.1.2).

It can be assumed that the changes that the L1 systems undergoes are subtle since those changes are not explicitly adapted from one system under the influence of the other. A key concept in the restructuring process is the way the L1 system is *used*, in order to match or to be accommodated into the other. On this matter, Schmid (2011) exemplified some caracteristics of language restructuring in a comparison with lexical borrowing, concluding that contrary to borrowing, restructuring is more commonly observed between similar linguistic contents (namely semantic information), even when the lexical form is divergent between languages.

Schmid's comparisons pointed out that restructuring involves mostly abstract concepts that differ across languages in the way conceptualizations are made. To some extent, restructuring may engage mostly high frequency words with unspecific meanings, or with flexible use, e.g., lexical verbs which allow different possibilities of collocations or prepositions (e.g., the extension of the use of '*take'* in English for German '*nehmen*' in attriters see Schmid, 2011, p. 28–29). This type of transfer (discussed in section 1.2.1.3.3 collocational transfer) results in literal translation of the L1 meaning inspired by L2 use.

[19] Diastratic variation refers to language change and its relationship to social factors such as age, social status, gender, etc. (Calvet, 1991).

From her perspective, this kind of lexical restructuring remains unnoticed because the prepositions or collocations that form the complement of the verb do not contribute to the encoding of meaning, so they are not fully processed. We can image that for the bilingual, it is the absence of negative evidence or feedback in L1 (in a L2 dominant environment) that will prevent the speaker to detect this type of deviant production. A second aspect proposed by Schmid (2011) is that restructuring would be an unconscious and involuntary process which is not related to a specific semantic purpose as in the case for lexical borrowing or codeswitching. To conclude, restructuring could be associated to implicit and unconcious adaptations (of abstract or ambigous concepts) which involve the way language is used.

Keeping in mind the definition proposed here, at some point it can be assumed that the number of languages the speaker uses will influence the way the metalinguistic analysis is made to accommodate different systems, or that the mostly used language will have an influence over the other(s). Further research is necessary to answer these interrogations, among the open questions, we may ask:

- *Is restructuring a strategy to compensate for linguistic gaps, because the linguistic knowledge is insufficient or deficient?*
- *Is it caused by the inaccessibility of linguistic information?*
- *Are there external factors that shape the way the L1 is already formed?*

To try to answer these questions at least partially, Pavlenko (2009, 2014) explains that conceptual restructuring is subsequent to a process called destabilization, in which L2 users readjust conceptual structures and boundaries in the L1 in order to match with constraints of the L2. From this perspective, restructuring would not be related to language loss or inaccessibility of L1 structures, instead, it would be related to a particular way in which systems accommodate to each other, as a result of a single conceptual representation that would "*center*" both uses in the languages into a single one (see Dewaele, 2018).

Dewaele (2018) proposes an interesting perspective: when the L1 lexicon is frequently used in a speech community, the L1 and their corresponding semantic representations[20] remain stable (i.e., an L1 user immersed in an L1 dominant context). Nonetheless, when the social configuration changes and the L1 lexicon become less frequent, because another language (L2) or variety becomes more frequent (e.g., in an L1 user immersed in an L2 dominant context) this will affect the stability of the L1 lexicon and the lexical representations, thus, producing a semantic shift, so that the L1 is in accordance with the new speech community use (L2).

[20] Dewaele's definition of semantic representation is related to the definition proposed by Pavlenko (2009) discussed in section 1.1.7.1.

This particular configuration (discussed in section 2.1.2.4 with respect to the role of language immersion) is interpreted as a *Cognitive Restructuring,* i.e., a process in which the language user, or bilingual, accommodates the L1 semantic representations into those of the dominant speech community. Dewaele described this process while studying two English variants -British vs American-, and the way speakers use emotion-laden words[21]. Language specificities will affect the way the speaker expresses emotion-laden words. According to this study, North Americans living in the UK have changed or restructured the use of emotion-laden words to match to the more frequent variety of English privileged by the speech community. This phenomenon is interpreted as the result of L2 exposure and acculturation that precedes destabilization, and cognitive restructuring. Following this study, it can be concluded that during immersion in another cultural community, if speakers who share the same L1 can adapt their variety of this language to correspond to the variety of the speech community, in the same way bilinguals may restructure the L1 to match with the L2 spoken in their speech community.

We agree with Pavlenko and Dewaele's approach to explain lexical restructuring which takes into consideration L2 exposure, language use and recognize that conceptualization processes are dynamic in the bilingual mind. For instance, restructuring is a perfect example illustrating that bidirectional transfer exists, but it goes beyond the loan and the adaptation of L2 words, it also affects cognition and the way L1 conceptual structures can be adapted into wider boundaries in L2. To some extent, semantic extensions can be associated to lexical restructuring, in both cases, subtle changes of meanings are exhibited: words narrowed or extended from one language to another (Schmid, 2011, p. 27). From this perspective, we can distinguish a new type of semantic extension very close to lexical restructuring, when two distant word-forms express a literal translation between two languages, involving relatively commons or ambiguous abstract concepts.

2.2.2. What Is First Language Attrition?

In this section, we describe the concept of first language attrition research and subsequently, we cover the topic of lexical attrition to finally conclude with a brief overview of two theories that illustrate the links between language attrition and language restructuring.

At any stage of second language learning, independently of the level of proficiency achieved by the second language learner, all kinds of L2 language users have experienced L1 changes as a result of second language learning, which constitute a part of bilingual language development (Schmid & Köpke, 2017). On this subject, Kroll et al. (2002)

[21] Those words do not refer directly to an emotional state (e.g., *exited*), instead it would elicit a particular emotion from the interlocutor (e.g., '*jerk*') in Pavlenko (2008).

suggest that L2 learning would add a complementary cost impacting the linguistic processes involved in the L1.

Our definition of attrition goes beyond the traditional perspective of language loss or erosion (Seliger, 1991), in fact, manifestations of L1 attrition can be apparent in multiple ways, such as language interference, reduction or simplification of the linguistic forms (Ahumada-Ebratt et al., 2018; Köpke & Genevska-Hanke, 2018; Schmid & Köpke, 2017). L1 attrition is the process by which the pre-existing linguistic knowledge (L1) becomes less accessible or is modified as a result of the acquisition of a new language (L2 or L3). Such a cognitive reorganization due to the addition of a new linguistic system would affect the L1 at different spheres: production and comprehension (Schmid & Köpke, 2017).

The core of the study of language attrition focuses on what is restructured in attrition but also at explaining how and why changes in the L1 linguistic system occur. In the present study the center of interest is on the interpretation of the changes of the L1. In fact, the study of CLI as discussed in Chapter 1.2 is not fully achieved through the description of types of language interference, but through the understanding of the underlying processes behind language change. Köpke & Keijzer (2019) suggest that bilinguals experience CLI because languages (L1, L2 or L3) are processed through the same mechanisms. Therefore, crosslinguistic interactions and attrition are consequences of the processing of two or more languages through the same mechanisms.

However, language attrition is not exclusively triggered by internal conditions, it can also be influenced by external conditions such as a new language environment (see section 2.1.2.4) or the reduced use of the L1 (see section c)). First, the change of the linguistic environment for an L2 dominant context, or a different variety of the L1 (e.g., Dewaele, 2018) would have consequences on the way the frequency of use of the L1 declines entailing a weakening of mnesic traces, and decreasing language accessibility. Hence, CLI (reflecting a state of constant language competition and inhibition) is closely related to language attrition due to the multiple changes of the conditions of L1 use (Ahumada-Ebratt et al., 2018).

Second, in the same line, the decreased use of the L1 is accompanied by the development and regular use of an L2, resulting, as mentioned before, in interference from the L2 (Schmid & Jarvis, 2014). From this point of view, CLI could be a departure point in restructuring. As suggested by Schmid (2011) the use of structures borrowed from L2 when using the L1 may replace to some degree the L1 equivalents, because they are easier to access than the L1 structures.

Third, language acquisition constitutes an explanatory framework of L1 language attrition since production and comprehension is affected by both systems. The impact of SLA on the L1 is witnessed through difficulties in the accessibility of linguistic information. These changes of language use commonly associated with CLI (e.g., lexical restructuring) are related to the process of being bilingual, leading to the suggestion that

every bilingual experiences L1 attrition as consequence of second language learning (Schmid & Köpke, 2017).

We can conclude that L1 inaccessibility is a current argument in L1 attrition, which confronts the idea that language cannot change, and instead it highlights that L2 users provide evidence of difficulties or failure at accessing information in the L1. As suggested by Linck & Kroll (2019, p. 95) "*the repeated retrieval of the non-dominant language may lead to reduced accessibility of the corresponding lexical representations in the dominant (native) language*", suggesting that language attrition is not related to oversight but to reduced access of linguistic information.

Lexical attrition is defined as structural changes that affect a person's mental lexicon, which implies that lexical attrition is more distinctly observable or measurable through the person's language performance, in particular in the person's lexical skills (Jarvis, 2019). In contrast, lexical attrition is presumed to be due to difficulties in language access rather than loss of lexical representations in the mind (Schmid & Köpke, 2009). It can be assumed that lexical attrition is manifested in the extremes cases as vocabulary loss, involving difficulties with both comprehension and production (i.e. in adoptees in a L2 dominant context). In less extreme cases, lexical attrition can emerge as retrieval difficulties as the result of changes relative to language use and bilingualism (Schmid & Köpke, 2017).

One of the rationales underlying this position is that language production of attriters would be characterized by disfluencies (as discussed in section 2.1.2) including filled or empty pauses, repetitions, lexical substitutions, circumlocutions, and even interference (e.g., codeswitching and borrowing). Actually, L1 attriters report difficulties showing that lexical retrieval can be an effortful process, however, the target word is (usually) attained after using markers of disfluency reflecting mental search of the word (Ecke, 2004). These difficulties during lexical retrieval can be explained by changes of linguistic environment (L2 immersion) and a significant reduction of L1 use by the bilingual.

In order to study L1 lexical skills and lexical attrition, Jarvis (2019) proposes three main spheres: lexical fluency, lexical accuracy and lexical complexity.

First, regarding lexical fluency, defined as "the ability to produce the L2 with native-like rapidity, pausing, hesitation, or reformulation"(Housen et al., 2012, p. 2), several studies (cf. Goral et al., 2008; Schmid & Jarvis, 2014) suggest that attriters or L2 users are slower in recognizing and retrieving the L1 than their L1 monolingual peers, creating disfluencies.

Second, lexical accuracy refers to the accurate knowledge, production and comprehension of the meaning of words, including syntagmatic relationships, usage constraints and connotations (Jarvis, 2019). Attrition in the sphere of lexical accuracy would manifest itself in the use of semantic extensions (e.g., the sense of the Spanish word '*carta*' extended to the sense '*carte*' to refer to '*credit card*' instead of '*tarjeta de credito*'), loan translations (e.g., misleading translations from one language to another such as

'actually' for '*actualmente*' in Spanish which means currently) and in the construction of compound or collocational words (e.g., associations of words such as '*to make homework'* instead of '*to do homework*'). In some cases, replacement with L2 words by attriters can be interpreted as a process in which the L2 lexicon is unconsciously integrated into the L1, resulting in lexical invention or morphologically adapted words inspired by the L2.

Third, lexical complexity is related to lexical diversity (i.e., variety in production), lexical density (the quality of lexical content in a language) and lexical sophistication that involves the use of infrequent words.

Traditionally lexical attrition is studied through verbal fluency tasks and picture naming for lexical accuracy. In our study we propose a battery of tests including a lexical decision task— rarely used in attrition studies—that will allow to observe both lexical processing and lexical accuracy through uncanonical association conditions (i.e., semantic extensions). Indeed, new experimental approaches are necessary to access lexical attrition. Schmid & Jarvis (2014) suggest that a combination of analyses involving lexical access and lexical diversity implying more naturalistic language production would be valuable to explore attrition as the result of managing two linguistic systems in parallel.

Regarding lexical diversity, Schmid & Jarvis (2014) have shown that L1 attrition was reflected in a decreased level of lexical diversity in comparison to the control counterparts. Attriters overused high frequency words and underused less frequent vocabulary. Regarding lexical fluency, bilinguals accessed fewer items of specific lexical categories than controls, suggesting that lexical access is effortful due to the fact that bilinguals have a larger linguistic repertoire. So, it can be assumed that the wider the mental lexicon, the harder are retrieval processes, in which bilinguals have to suppress L2 items in order to access the L1.

To conclude, from our perspective language attrition is viewed as the result of second language acquisition, in which a reorganization of the languages in the bilingual mind affects the L1 (e.g., producing a reduction or simplification of rules or modifying L1-L2 lexical accessibility). Language attrition is affected by multiple external conditions, particularly changes in L1 language use and related to linguistic environment. Studies agreed that frequency of use plays an important role defining bilingual's activation thresholds as predicted by the Activation Threshold Hypothesis which will be develop in what follows.

2.2.2.1. Activation Threshold Hypothesis (ATH)

The Activation Threshold Hypothesis (ATH) was adapted for the first time in the context of bilingualism by Paradis using a cognitive approach and following the neuronal functioning introduced by Hebb (1949). In 1993-2007 this theory assumes that in order to access linguistic information that is stored in memory, the brain appeals to multiple groups of neurons which form neuronal circuits, they are responsible of sending, receiving and transmitting information. For instance, to retrieve a word a number of circuits are

requested to 'activate' that particular word. Frequency of use of any linguistic item will reduce the activation threshold, that is, the effort needed to access the information (Fabbro, 1999). In contrast, when the neural circuit is not frequently used or accessed, the process of activation of the linguistic item is slower, so activation threshold increases. In other words, the more we use a particular linguistic unit (or a particular language), the more effortlessly the reactivation will be. Actually, constant activation will modify structurally and functionally the neural circuits in accordance to frequency of use (Fabbro, 1999). This means that '*repetition makes perfect*' and that frequency of language use will impact considerably the way connections are made in the bilingual brain. But there is another mechanisms that is as important as the activation threshold, that is the inhibitory mechanism, whose function is to block the activation through inhibition of the neuronal activity (Fabbro, 1999).

In the context of language attrition, the ATH provides a valuable perspective that may explain L1 inhibition as the starting point of L1 attrition. That is, in the case of second language users immersed in an L2 dominant context, for whom the most activated language is inevitably the L2, the L1 would have a higher Activation Threshold than the L2. Hence, the L2 will be more easily activated which explains constant intrusion of the L2 during L1 production and use (Paradis, 2004, 2007).

However, it seems that frequency effects do not explain entirely L1 attrition. Studies testing frequency effects in the context of the ATH are not conclusive, for example, grammar and lexicon are not affected to the same degree by frequency of use (Köpke, 2002). Similar results were observed by Gürel (2004) who described selective attrition for aspects of grammatical processing involving mismatch/divergent competition between both languages (in Köpke, 2019). Once again, it seems that the lexical level is sensitive to frequency of use. We can hypothesize that this sensitivity might be related to the strong competition of lexical items, which are rarely subject to flexibility/permeability as is the case for grammar.

Another point is that sensitivity to variation in AT may also depend on language typology, wherein cross-linguistically closer lexical items and translation equivalents are probably more prominent for competition. Considering this, instead of relying on frequency of activation by itself, a more interesting perspective to approach L1 attrition is to consider the competition mechanisms with the added value of frequency effects (Schmid 2011).

2.3. The Weaker Links Hypothesis

The Weaker Links Hypothesis (Gollan et al., 2008) builds on frequency effects and the idea of weaker conceptual links in bilingual speakers compared to monolinguals. Since bilinguals use less frequently each of their languages than monolinguals, it is assumed here, that an L2 dominant context leads to reduced accessibility of the L1

explained by the weakening links between the conceptual and the form representation in the L1 (Gollan & Acenas, 2004).

In comparison to monolinguals, the linguistic disadvantage of bilinguals would be associated to the way multiple languages are used. The fact that bilinguals split their time and contexts of language use between both languages, causing a disparity that will configure differently connections for each language with different strengths between semantic and lexical representation in the L1 or L2 of bilinguals, compared to monolinguals whose connections are limited to a single language (Gollan, Montoya, Fennema-Notestine, & Morris, 2005). In other words, the bilingual disadvantage is translated in the constitution of weaker links in each language compared to monolinguals, which is argued by the fact that monolinguals do not have to share domains of language use or dimensions of competence between languages.

Consequently, this difference in the construction of intra-language connections will be an indirect effect of bilingualism on lexical retrieval, which will be more strongly influenced by language use or frequency effects. For instance, reduced L1 language use would influence word frequency because it would significantly affect the existing links between the phonological, semantic and lexical representation of each language, thus, affecting lexical accessibility.

2.4. Convergence Between Languages

This section seeks to define convergence and review the main differences between convergence and restructuring, both phenomena being anchored perfectly in the framework of our research in view of a better understanding of semantic extensions.

Convergence is defined in Ameel et al. (2009, p. 271) as "*the enhancement of inherent structural similarities in the two systems*" a process of interaction between two different languages in the bilingual mind reflected through the influence over time of a language over the other, and evidenced at different linguistic levels: syntax, semantic and phonology. Following Pavlenko (2000) it results from the outcome of the existence of similarities in two languages.

As the term suggests, the concept of convergence refers to the merging or integration of two systems that results in a new creation that differs from both source languages. One of the conditions that distinguish convergence from other kinds of transfer is that the same linguistic unit or structure must be present in both languages: crosslinguistically shared features are necessary for convergence, which is not the case for restructuring, wherein transfer can rely on only a common content. The difference between restructuring and convergence relies on the idea that restructuring refers to partial modification of the existing conceptual categories (Jarvis & Pavlenko, 2010). In restructuring new links and association can be developed as the result of second language influence such in the case of semantic extension. In contrast, convergence would be

determined by increasing similarities between languages in which shared features are necessary to be 'convergent' across languages.

Several authors (Bice & Kroll, 2015; Dijksta, 2002) argue that the L2 converges with the L1 in the case of cognate translations, in other cases, it has been suggested that convergence would involve false friends or false cognates [22] (e.g., Schmid, 2011). Phonological similarities between languages may also trigger convergence as observed by Sharwood Smith (1983). At this point we wonder whether the formal similarity of typologically related languages plays a role in language convergence or if convergence constitutes a strategic way to confront both concepts and converge into a common representation. Indeed, crosslinguistic convergence is interpreted to be part of a normal stage of second language acquisition, independent of the L2 learner's proficiency levels (Brown & Gullberg, 2013). Riehl (2019) proposes convergence to be a compromise strategy which enables the L2 user to integrate lexical components from different linguistic repertoires.

2.4.1. Structural Ambiguity between Languages

This concept proposed by Müller (1998) suggests that the direction of crosslinguistic transfer of a particular structure is produced from the less ambiguous structure to the more ambiguous structure. In other words, the structure that is easier to use in a particular language context is going to be placed over the structure that has multiple exceptions and that, for instance, cannot be used in every linguistic context. An example illustrating this is proposed by Schmid (2011) in attriters of pro-drop languages such as Italian and Spanish who had English (not a pro-drop language) as an L2. For these speakers, the production in L1 is characterized by an overuse of anaphoric pronouns, which are not necessary in pro-drop languages. This can be interpreted through the theory of structural ambiguity since the use of pronouns in Italian or Spanish depends on the contexts in which the structure is used with both options being possibly. In the L2 (English), pronouns are obligatory, and zero pronouns are not permitted. In this case the option that does not allow for variation (English as a non-pro-drop language) will win the competition and operate over the more ambiguous one (Spanish a pro-drop language) leading to overgeneralization from the L2 rule to the L1.

Structural ambiguity can easily be related to the Reduced Redundancy Principle (Seliger, 1991) sharing the same explanatory basis: the L2 intervenes in the L1 when the latter is ambiguous, resulting in CLI. However, the Reduced Redundancy Principle includes also the role of native speakers at defining negative and positive evidence that play a major role on L1 and L2 linguistic competence and awareness.

[22] Defined as crosslinguistic words that share formal features but have different meanings.

Both principles can be associated in the study of Grosjean & Py (1991), proposing that the L1 competence of long-term bilinguals is governed by a set of grammatical criteria arising from the L2—which have allowed the authors to detect which L1 rules were restructured following the Reduced Redundancy Principle coming from the L2 —, in consequence, grammatical structures or norms were replaced in the L1 under strong influence of the L2. In the case of semantic extensions such rules were not applicable since the use of lexical units is not conditioned by grammatical structures.

Considering that lexical ambiguity is harder to account for in order to determine which structure prevail over the other, than is the case for grammatical structures, we propose to focus on activation levels following the ATH. Constant L2 activation levels over the L1 in the context of immigration would imply that all lexical structures will be equally affected by language attrition (in terms of reduced availability in the mental lexicon) or language restructuring (in terms of partial modifications at the conceptual levels), however, this is not the case. Semantic extensions cannot be entirely explained by formal features nor by exclusively semantic convergence. In order to investigate this complex phenomenon, this research integrates two psycholinguistic factors that are known to influence the mental lexicon on monolingual and adapted them in the context bilingual research: Neighborhood Density and the Size of the Morphological Family of words in L1 and L2 (see section1.2.2).

2.5. How to Interpret Semantic Extensions

Semantic extension can be seen as the result of lexical restructuring of the L1, triggered by constant L2 influence. In this case, L1 semantic extensions are framed in the dominant L2 and influenced by factors such as frequency of use, dominance and proficiency, playing a role in the transformation of the L1. In this process, the L1 plays an important role as well as discussed in section 2.1.2.1, L1 language use has a key role preventing from restructuring. However, in the present study, language restructuring is not viewed as a regression process leading to language attrition, but as an illustration of the dynamics of the mental lexicon.

When the L1 is restructured at the lexical level, (e.g., the replacement in Spanish of '*subir las escaleras'* for '*montar las escaleras'* triggered by the French verb '*monter les escaliers'* -go up the stairs-) L1 reestablishment to standard production is possible. Hence, the use of semantic extensions is not irreversible. A recent literature review proposed by Köpke (2020) suggests that re-immersion of L1 attriters in an L1 linguistic environment even for short periods (one or two weeks) is sufficient to reverse attrition affects, for example with respect to production and on-line processing of pronouns. In the same perspective, Linck et al., (2009) suggest that L1 competence may be reestablished after having been immersed in an L1 dominant environment again.

From the perspective of restructuring, semantic extensions do not involve formal 'adaptations' into the recipient language or phonological change — the sphere that is

'*affected*' here is the meaning in the way it is extended or calqued from another language—. Restructuring occurs between interlingual word-pairs (e.g., cognates and false cognates), however, the formal overlap does not seem to lead automatically to semantic extensions. If this were the case, for second language learners, it would be challenging to learn languages typologically related to their L1, instead, similarities between linguistic systems, for example L1 and L2 language distance seem to facilitate in some cases (e.g., in comprehension) the learnability of an L3 (Schepens et al., 2016). Hence, an essential aspect in SLA is the awareness of crosslinguistic similarities to avoid interference between typologically related languages, implying that semantic extensions would be less marked in bilinguals who speak typologically unrelated languages. However, semantic extension occurs between unrelated and related languages as well (see Ringbom, 1987, 2001). Considering that the overlap of formal features is not the only factor triggering semantic extensions, and that bilinguals' linguistic awareness of interlingual similarities prevents from overgeneralization, a further perspective considered here includes the conceptual network existing between the systems (Abutalebi & Green, 2007).

A second factor explaining semantic extension is the semantic or conceptual[23] interaction between languages, which can be illustrated by the way conceptual specificities are structured in each language (see Pavlenko, 2009). From this perspective, the current proposition follows the Reduced Redundancy Principle (Seliger, 1991) that we apply here to semantics extensions: bilinguals would disfavors the language in which the semantic ambiguity is more pronounced and in which multiple exceptions exist, and favor the language in which semantic constraints are more consistent. Hence, conceptual restructuring will be produced in the language that involves more inconsistency, especially if the corresponding conceptual structure in the other language is more reliable.

As an illustration (see **Figure 14**), we refer to the hierarchy proposed by Heredia and Brown (2006, p. 240) with an example of the verb '*amar*' vs '*love*' (discussed in section 1.1.5) following the Reduced Redundancy Principle. As it can be seen in **Figure 14**, the verbs '*amar'* and *'love'* are used differently with animates and inanimates in Spanish and English. '*Amar*' in Spanish has multiple constraints, represented by rectangles corresponding to inanimate things. In English, the conceptual structure is less complex, because the verb may be used for living entities and non-living entities (Heredia & Brown, 2006, p. 240). Hence, the use of *amar* in Spanish may adopt some patterns of language use arising from the L2 such as "*amo el vaso*"/ "*I love the glass*". This adaptation is supposed to be motivated by the fact that semantic ambiguity is quite complex in Spanish and more reliable in English.

[23] Here we do not make the distinction between semantic and conceptual levels proposed by Pavlenko (2009) because from our perspective semantic extensions can attain both types of representations. In some cases, the conceptual structure is attained (in the sense that concepts are learned conjointly) and in other cases processing occurs in superficial levels that may be explained under the scope of crosslinguistic misleading associations between languages.

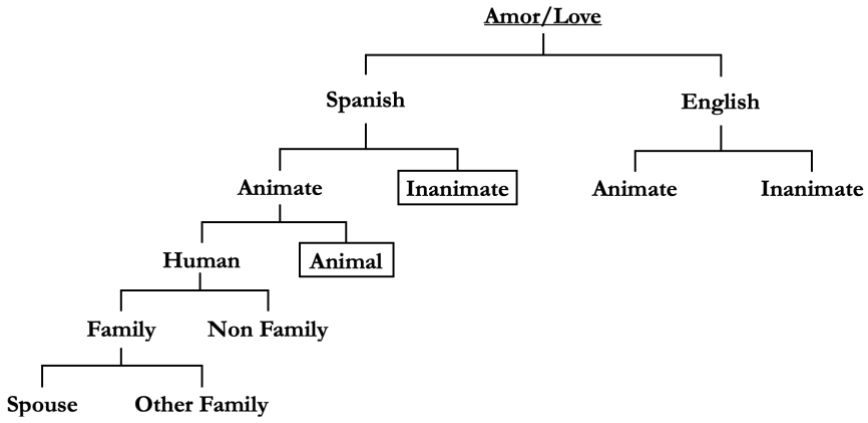

Figure 14 A schematic description of amor and love (Heredia & Brown, 2006, p. 240)

Moreover, semantic extensions can also be defined under the scope of first language attrition, since in some cases the use of semantic extension is marked by the inaccessibility of concepts in the bilingual lexicon rather than the loss of the linguistic information. In these cases, it is possible that the use of semantic approximations between the L1 and the L2 such as '*amo el vaso'* by Spanish-English bilinguals is more available in the bilingual lexicon than the actual standard form in the L1 '*me gusta el vaso'*.

The current research seeks for a better understanding of the stage(s) that precede(s) language restructuring and language attrition, which seem to affect the lexical level at the very beginning of SLA, a topic that needs undoubtedly further investigation. In next sections we propose three plausible framework that may allow us to understand the nature of semantic extensions, as a cognitive strategy, as the result of L2 acquisition and L1 inhibition.

2.5.1. Cognitive Strategy

One of the open questions with respect to the possible reasons for the arising of semantic extensions is whether they due are to a cognitive strategy facilitating language production and comprehension.

Basetti et Cook (2011, p. 10) explain how an additional language affects cognition, assuming that the acquisition of an L2 influences codability of languages and the way the speaker expresses habitual thought. Codability is related to the way a particular concept is lexicalized (expressed in vocabulary) and grammaticalized (expressed in syntax or morphology) in a given language. For example, a concept may be labeled immediately in one of the bilingual's languages and be inexistent in the other, e.g., the Spanish concept of '*estrenar*' ('*wearing something for the first time*') is inexistant in English. The way concepts are categorized depends also on language specificities such as the color spectrum (e.g., the Italian label '*azzuro*' meaning '*light blue*' in English). Form and material of household objects (Pavlenko & Malt, 2011) are categorized differently depending on the languages

and at different degrees in each language (e.g., '*chashka'* in Russian and *'cup'* in English, see section 3.1.2).

Another perspective is that the habitual way of coding reality will influence conceptualization at the preverbal level. For example, features such as number or evidentiality (Chafe & Nichols, 1986) are encoded differently in each language and put different demands on speech planning depending on the language. This suggests that processing in bilinguals and monolinguals might be significantly different not only in (oral or written) language production, but also from the onset of thinking what to say and how to say it in each language, i.e. conceptualizing.

Taking into consideration that bilinguals' concept coding processes are different, and that non-linguistic information affects the way the speaker thinks and expresses ideas, we hypothesize the use of semantic extensions in the L1 as due to the arising of a single semantic representation that is merged from two different languages. This can be considered as an economical cognitive strategy because connections between word forms and referents are gathered into a single set of connections involving fewer cognitive resources (Ameel et al., 2005). Otherwise, another possibility would be that representations are partially shared between the two languages, leading to convergence.

Another important aspect regarding semantic transfer is that semantic extensions do not replace the L1 lexico-semantic knowledge, a double performance is frequently observed: errors as well as canonical productions in the L1 suggest that new semantic knowledge is added to previous ones. This is in line with previous studies suggesting that semantic boundaries in the L1 and in the L2 show simultaneously influence of both systems (Graham & Belnap, 1986).

2.5.2. SLA and L1 use

Another issue of interest concerns the way L2 acquisition impacts L1 use. We hypothesize that SLA affects the way bilinguals cope with the decreased accessibility of the L1 which is caused by an L2 dominant environment (Linck et al., 2009). During SLA, the already existing languages are not "erased" in the brain, on the contrary, they serve as tools to learning and using the new language. As a consequence, the new knowledge of the L2 impacts the L1 (Schmid, 2011).

Concerning L2 learning, Malt, Jobe, Li, Pavlenko, & Ameel, (2016) proposed a number of factors that may explain some difficulties that L2 learners face while shaping progressively the L2 word meanings into native-like targets. The struggles during SLA would originate from initially poor corrective feedback and a certain lack of sensitivity towards native-like use (e.g., when L1 constraints are mapped onto the L2). Following the framework of lexical network construction (Murphy, 2004), an additional struggle is the parallel use of two languages. Because both are used on a daily basis, distinct representations are constructed, rendering the shaping of L2 word-form meaning mappings difficult. Another determinant factor (as discussed in section 2.1.2.4) is L2

exposure and the L2 dominant environment which affects word-meaning connections in both languages. For instance, limited L1 use, and retrieval affects L2 representations in the same way as active L2 use and retrieval affects the L1, which may explain the arising of semantic extensions in L1.

Following Ameel et al. (2005), L2 use causes the reactivation of associated words in the L1, resulting in changes of meaning patterns in the L1. More specifically, these authors observed that category boundaries of French-Dutch bilinguals influenced both languages, that is, bilinguals expressed category boundaries for drinking containers differently from Dutch and French monolinguals, which exemplifies Cook's multicompetence model (2003). For example 25 objects named as '*fles*' (bottle) by Dutch monolinguals can be divided into 2 different categories in French by French monolinguals among which 13 objects are categorized as '*bouteille*' and 10 objects as '*flacon*'. In other cases, there was no correspondence between categories across languages: a single category of containers in Dutch (e.g., *bus*) could be spread into 6 different ones in French (*bouteille, flacon, spray, bidon, brique,* and *bombe*). As it can be observed, the same objects that are categorized in a single category for one language (e.g., 25 objects categorized for a single category '*fles'* in Dutch) may be categorized in different categories in the other language (the same 25 objects are categorized in two categories '*bouteille*' and '*flacon*' in French) (Ameel et al., 2005, p. 67).

Once bilingual data is compared with monolingual data, a clear difference between bilinguals and monolinguals is established, especially in the category of bottles and dishes, suggesting that both languages permeate the mental representations of bilinguals. Ameel et al. (2005) explained these results through the *'one pattern hypothesis'* which suggests that naming patterns of bilinguals converge into a single one, which implies a degree of interconnections between the word forms of the two languages and their referents. Since these interconnections are limited to one set, the *one pattern hypothesis* suggests that it would be '*cognitively economical'* and less demanding in terms of resources in memory than a '*two pattern hypotheses'*. In the latter, two naming patterns would converge toward one common naming pattern that would not match perfectly across languages.

But the results observed in this study could be interpreted as well as the restructuring of the mental L1 and L2 mapping due to learning an L2. This may be the case of '*bouteille*' whose concept is extended to conceptual features corresponding to the Dutch word '*fles*' which are used as interchangeable concepts by bilinguals (Ameel et al., 2005). An additional interpretation of the data is that this phenomenon might be the consequence of SLA: when advanced learners acquire new vocabulary in the L2, this affects the mappings of L1 structures, which allow progressively the naming patterns to match into a single one that would converge between both systems.

Taking into account the effects of L2 acquisition on the L1, we may interpret that semantic extensions could be the result of a convergence point that involves lexical mapping and representations. This would not mean that once bilingualism restructures the way interconnections between languages are linked the L1 remains deviant, instead it

is enriched with both representations. Hence, bilingualism would be a linguistic advantage over monolingualism, and learning a new language would help to reinforce the linguistic competence of the L1. The cumulated knowledge a speaker has acquired in the L1 is helpful in order to construct an interlingual development that may accelerate or delay the progress of developmental pathways[24] (Ortega, 2013). This process is called interlingual identification which is constructed considering three factors:

1) The nature of the specific L2 phenomenon and the universal forces that shape its natural development,
2) The distance perceived by the learners between the L1 and the L2 and their intuitions of what is transferable or not; and
3) Their relative proficiency level" (Ortega, 2013: 33-34).

From this perspective, L2 acquisition constitutes a way of progressively training linguistic abilities such as comparing distances and similarities between systems. Such linguistic awareness involves also construction strategies that help to enrich linguistic concepts such as in the use of borrowings or even codeswitching.

Semantic extensions constitute an illustration of this process of interlingual identification in which the learner's knowledge contributes to the emerging lexicon that converges between both languages. Hence, lexical transfer would not define a path of development but crosslinguistic similarities. For instance, L1 and L2 similarities would facilitate learning of L2 structures, as suggested by Jarvis (2002), who studied accuracy rates of English-Swedish bilinguals in their use of definite articles. The results suggested that the similarities between the L1 (Swedish) with English facilitated the acquisition of definite articles of late bilinguals. However, bilinguals with other L1 backgrounds would favor different kinds of strategies, such as overgeneralization of rules, that may result in errors in the L2. This is the case of Spanish-English bilinguals who tend to overuse definite articles in English, c.f. "*me gustan las patatas fritas*" vs. "*I like Ø French fries*" cited by Ortega (2013, p. 36).

Many other examples are provided by Kellerman (1995) to illustrate that crosslinguistic similarities, such as cognates, in an L2 typologically related to the L1 can be useful when exploring a new language system or at the beginning of second language learning. Indeed, discovering and taking advantage of those similarities allows the learner to understand and express non-acquired elements in the L2 on the basis of the L1.

[24] Pathways of development here refers to the natural progress in language acquisition such as the order of acquisition of grammatical rules or syntactic constrains that are acquired depending of the learners' age and the language concerned.

2.5.3. L1 inhibition

There are subtle differences in the way concepts or ideas are expressed in a given language, as can be seen when contrasting for example the expression *pay attention* in English with other languages, e.g., in Spanish or French attention is '*lended*' (*'prestar attention'* or *'preter attention'*) and not paid. Considering that lexical representations are formed from word associations rather than unities, word meaning and use pattern can be viewed as a lexical network that conforms more complex systems. Thus, different degrees of association strength are made forming conceptual layers between the two languages (Malt et al., 2015). If semantic extension is considered under the scope of feature association and strength between representations, the manner in which both languages interact with each other would influence inhibition processes. This process would also depend on the number of shared associations existing between both languages.

Following the Inhibitory Control framework (Green, 1998), the strong influence of the L2 over the L1 can be explained by changes in L1 accessibility. The reduced accessibility of the L1 may be explained by L2 immersion, an L2 dominant context enhancing L1 inhibitory mechanisms. On this matter, studies by Linck et al., (2009) (see also Linck & Kroll, 2019) have found that reduced accessibility of the L1 is observed in comprehension and production tasks, when the participants are immersed in an L2 dominant context. Indeed, these authors have reported contrastive results of L2 learners in an L1 verbal fluency task made before and after L2 immersion. L2 learners immersed in an L2 context showed a reduction in L1 verbal fluency, suggesting strong L1 inhibition in an L2 dominant context. It can be assumed that in semantic extensions, strong L1 inhibition and predominant L2 activation can account for L1 changes at the lexical level.

In sum, this research suggests that the capacity of accessing linguistic elements in the L1 is affected by constant inhibition of the L1 in order to allow the learner a better adaptation in an L2 dominant context. This can be supposed to facilitate L2 acquisition.

As discussed in section 2.1.2.2, constant inhibition of the L1 in parallel with SLA would allow the learner to reach better L2 mastery because of high frequency of L2 use, as this language is becoming progressively more dominant. As can be noted, L2 dominance may be associated with crosslinguistic influence at multiple linguistic levels. This type of L1 changes would also be supported by the ATH. However, as discussed above, a clear cut-off difference between L1 language attrition and L1 restructuring due to SLA effects is hard define, and we have argued that these two processes would be part of a same continuum (see also Köpke & Keijzer, 2019). Indeed, this would not mean that L2 acquisition affects L1 knowledge negatively, instead languages would inevitably exert influence on each other until they attain some stability (Ecke, 2013).

2.6. CHAPTER SUMMARY

The present Chapter illustrates the interacting levels of languages with special attention to the lexicon based on the assumptions that systems are dynamic structures that complete themselves grow together during language acquisition. We hypothesize different factors influencing transfer such as language use, language dominance and language immersion. We locate semantic extensions into other theoretical paradigms explaining language change including language restructuring and first language attrition. Finally, we propose to interpret the case of semantic extensions as the result of a cognitive strategy, the interface between SLA and L1 use and as the result of L1 inhibition processes.

RESEARCH CONTRIBUTION

3. **Experimental Design**

The current study seeks to investigate the phenomenon of semantic extensions in Spanish (L1)-French (L2) late bilinguals who are immersed in an L2 dominant context. We compare late bilinguals with monolinguals in their use of L1 (Spanish) and hypothesize the restructuring of the semantic system of the L1 during the first years of L2 immersion. We tested their L1 use regarding semantic extensions from different levels including oral production, processing of isolated words, and sentences.

The focus on late bilinguals was motivated by the following reasons:

Pavlenko and Jarvis (2002) propose that there are two sources of research bias in the field of bidirectional transfer. First, the lack of studies on adult late bilinguals. Actually, the field of crosslinguistic influence has mostly explored the case of L2 influence on the L1 in childhood bilinguals (or heritage language speakers, see Montrul & Polinsky, 2019) and it is frequently assumed that in adult late bilinguals the native language is 'mature' enough to resist to changes, or that only highly proficient bilinguals will show signs of first language change.

The second research bias we want to challenge concerns the little attention that has been paid to individual patterns of L1 language change, i.e., in the first generation of immigrants. In the past, crosslinguistic influence has been studied through the sociolinguistic scope of language change, commonly referring to second-generation speakers who learnt a contact variety of L1 from a generation that deviates from native speakers' standards.

Concerning the lack of studies on adult late bilinguals, our position agrees with the assumption that the L1 might be subject to restructuring at any time of life of the bilingual and that languages are dynamic and not stationary or rigid systems. In the same way as a monolingual speaker enriches his/her linguistic abilities during his lifetime, the lack of language contact, the limited language use and language input in an L2 dominant context may interfere with the L1 enrichment in late bilinguals. In this particular configuration, the positive evidence (Seliger, 1991) that a speaker normally receives from native peers in an L1 dominant context will also contribute to language development, the latter being not necessarily completely attained in adulthood, instead it is a life-long process. The investigation of late bilinguals is particularly interesting because it will allow us to illustrate the flexibility of the L1 which may even in late bilinguals be changed, restructured, and adapted into a new linguistic environment. In an L2 dominant context the L2 is a major source of enrichment, competing with but also improving the L1 competence, at the same time as the L2 proficiency level increases. Pavlenko & Jarvis (2002) have reported semantic extension in L1 Russian post-puberty learners of English, whose L2 affected the L1 in all areas of language and especially at the lexical level. Likewise, Grosjean & Py (1991) observed such phenomena in long-term immersed bilinguals Spanish-French through two tasks: Acceptability Judgment and an Attestation Task, used to attest the existence and to evaluate L1 ungrammatical sentences. Since a double task may influence the other, we chose to implement an Acceptability Judgment Task to evaluate sentences involving SE.

The current research focuses on 40 L1 Spanish - L2 French bilinguals, in order to explore the changes that occur in the early years of L2 immersion in France (mean LOR 4,8 years) at the lexical and semantic level through the study of semantic extensions in the L1. In this research we explore the concept of multicompetence, which holds that linguistic competence of bilinguals in the L1 will be affected by the L2 in the same way as the L1 affects the L2, a process that results in a linguistic competence that differs from the corresponding monolinguals in either language.

Concerning the individual patterns of L1 language change, independently of the importance of social factors related to L2 acquisition in an immersion context, we argue that the first generation and their use of the L1 at the beginning of immersion constitutes the beginning of language change in opposition to the incomplete acquisition (see Montrul, 2011) related to limited Spanish input between generations that will affect the quality of the L1 acquisition in later generations.

Focusing on individual patterns of language change instead of societal change will allow us to better explain CLI at the beginning of immersion, but this will also contribute to the understanding of future intergenerational language change. Recently immersed bilinguals are of great interest in the investigation of CLI in the L1, which is supposed to be stabilized in adulthood. Moreover, the study of semantic extensions will provide us a better understanding of the processes underlying CLI production in bilinguals.

This research focuses particularly on the lexical level because as we have seen in section 1.2.1, even though the lexicon is often claimed to be the most vulnerable system in language restructuring and change, and in spite of the permeability of the lexicon, there are few studies that focus on the influence of the L2 lexicon on the L1 (e.g., Jarvis & Pavlenko, 2002). As discussed by (Hohenstein et al., 2006), further research concerning bidirectional effects on the lexical level are needed for a better understanding of bilingual's patterns of language use.

The difference between the grammatical and the lexical level is that the first is governed by rules and syntactic constraints, which facilitates the detection of irregularities and crosslinguistic influence between languages. In contrast, the lexicon is governed by semantic constraints instead of specific rules. The idea of semantic constraints is less structured motivates our prediction that the lexical level is more fragile than the other levels. For instance, contrary to the grammatical level, the irregularities at lexical levels cannot be detected at first glance, instead subtle changes can be detected when the context in which the language is used is considered and compared to patterns of language use and by different groups of speakers.

Grammatical competence is more likely to be measurable in terms of respect or violation of rules, while the lexicon is observable rather through production measures (in spontaneous speech, written production or elicited speech) and processing measures (Response Times, Event Related Potentials, etc.). However, there is no general criterion that governs lexical patterns of language use in order to detect deviation from standard

language use, since the lexicon is the subsystem that evolves the most (e.g., diastratic variation) and in short time notice.

In order to study CLI at the lexical level, researchers do not benefit from specific language indicators (e.g., rules), which implies that the detection of transfer needs to take into account other psycholinguistic factors that influence and trigger CLI as indicators to facilitate the analysis: morphological links, semantic links and formal links between languages. These indicators are taken into consideration in lexical access modeling for monolinguals, as for example in the dual mechanism model which accounts for morphological facilitation (e.g., word stem and suffix facilitation between primes and targets). With respect to neighborhood density (Davis & Taft, 2005; Forster & Taft, 1994; Van Heuven et al., 2001), this factor included in the Dual Route Cascade (DRC) model of Visual Word Recognition (Coltheart et al., 2001).

Nevertheless, these psycholinguistic indicators are hardly taken into account in models of bilingual lexical access (described in Chapter 1). This is the reason why we propose to investigate the role of these indicators and compare two groups of Hispanics: a bilingual group immersed in France and a monolingual group in Colombia. Participants performed 3 experimental tasks in their L1 (Spanish):

1) Word Production in a Gap Completion Task
2) Word Identification in a Lexical Decision Task
3) Comprehension in Acceptability Judgment Task

As mentioned before, among the factors that govern the use of the lexicon we will focus specifically on two psycholinguistic measures CLI: Neighborhood Density and Morphological Family Size. For this purpose, stimuli tested were chosen with either weak or Dominant Neighborhood Density and weak or Dominant Morphological Family Size in the L1 and the L2. This configuration allowed the investigation of semantic extensions through two levels: crosslinguistic word-form overlap (at the lexeme level) and crosslinguistic semantic overlap (at the lemma level). It is assumed that high neighborhood density will facilitate lexical access in one of the languages whilst for words with few neighbors such benefits will be smaller (Costa et al., 2006). Meaning overlap as manifested through morphological family size will contribute to crosslinguistic competition as well (Mulder et al., 2013) with semantic activation increasing with higher morphological family size of the targets(De Jong et al., 2000).

As discussed in section 1.2.1.1 semantic extension is traditionally defined as a lemmatic transfer, which implies that activation levels are not supposed to intervene in the linguistic transfer following Jarvis (2009). In opposition to this perspective, we hypothesize that crosslinguistic word-form overlap explains at least part of the production of semantic extensions, implying that activation and competition levels between languages are a valuable explanatory framework for semantic transfer.

Following the Reduced Redundancy Principle and taking into account activation levels, we hypothesize neighborhood density and morphological family size play a

complementary role in semantic extensions. Hence, when the L1 features involve strong neighborhood density or morphological family size or both in comparison to the L2, the L1 will show resistance with respect to language restructuring, replacement, and inaccessibility because these features will help to maintain strong activation of the lemmas/lexemes. In contrast, when neighborhood density or morphological family size or both are stronger in L2 than L1, the latter will be more vulnerable to change and deviations may be observed at the lexical level, in particular the semantic spheres will rely more on L2 and give rise to semantic extension, because the L2 will benefit from higher levels of activation in comparison to the L1.

3.1. General Hypothesis

We propose 2 general hypotheses that articulate the current research:

H1: Late bilinguals or advanced French learners/speakers will be sensitive to semantic extensions during processing and use of the L1 (Spanish) in comparison to Monolinguals (Spanish speakers living in a L1 dominant context).

H2: Neighborhood Density and the Size of the Morphological Family of the words will affect crosslinguistic links between languages of bilinguals, producing interferences in production, processing and judgment.

With respect to the three experimental tasks (detailed in section 3.6) we propose the following predictions for each of the three experimental tasks.

H3: In the Gap Completion task, bilinguals will produce L1 semantic extensions since the L2 is activated while the L1 is processed and produced.

H4: In the Lexical Decision Task, bilinguals will process semantic extensions ~~d~~ (as shown by RTs) depending on accumulated sources of activation triggered by the Neighborhood Density and the Morphological Family Size (in the L1 and the L2). We predict that a strong activation of L1 in comparison to L2 will prevent occurrence of semantic extensions (as shown by shorter RTs), while strong activation of L2 will enhance interference of L2 and result in semantic extensions in L1 leading to slower RTs.

H5: In the Acceptability Judgment Task, bilinguals will be more flexible in their evaluation of L1 sentences that express a semantic extension. We hypothesize that bilinguals may accept to a higher degree and maybe reject the corrected meaning of semantic extensions (Condition 3), the type of sentence accepted (condition 3 and 2) will indicate whether semantic or conceptual restructuring is rooted in the L1.

3.2. EXPERIMENTAL STRATEGY

This Chapter describes the characteristics of the population involved in this research, the rational related to their recruitment and procedures. Moreover, a detailed description about the constitution of the linguistic material is proposed, including the databases used and their stages of construction. Finally, we present the three experiments developed and the aims of each experiment. Experimental tasks will be presented separately, including its time setting, structure, and procedure during experimentation.

3.3. POPULATION

For this research, we needed to constitute two groups matched by age, L1 (Spanish) and by educational instruction background (secondary school graduates or bachelor's degree level).

Our target population was composed by a group of French bilingual with L1 Spanish, we will refer to this sample as *Group 2*. Our control population was composed by a group of Spanish without any instruction in French that we will refer as *Group 1*. *Group 2* was immersed in France and *group 1* was living in their native country (Colombia), thus, we have opposed L1 dominant context population to L2 dominant context (France) population.

3.3.1. Inclusion Criteria

All participants were asked to answer a few questions before the experimentation to assure that the profiles of participants fitted with the inclusion criteria of the research.

The inclusion criteria for *Group 2* comprised:

- Aged between 18 and 45 years to assure homogeneous L1 use
- L1 Spanish speaker native from Latin America
- Late acquisition of French as a foreign language (who have begun to learn a second language once their first language was fully acquired)
- Immersion in France after the age of 20-years
- Being student at the university or actively involved in French language learning
- Minimum lengh immersion in France of one year
- Use of French and Spanish on a daily basis

The inclusion criteria for *Group 1* included:

- Aged between 18 and 45 years to assure homogeneous L1 use
- L1 Spanish speaker native from Latin America
- Being student at the University

3.3.2. Parameters of Exclusion

For bilinguals (*Group 2)* parameters of exclusion were:

- Being early bilinguals, the age of acquisition and immersion (i.e., after 18 years) were restricted to adulthood only.
- Being a graduate form a French college in Latin America or international school since this type of early L2 schooling is aimed to form simultaneous or early bilinguals.
- Having a learning disability, visual or hearing impairment.

For monolinguals (*Group 1)* parameters of exclusion were:

- In order to avoid sensibility to French, participants had to be naïve in French as a second language.
- having learning disability, visual or hearing impairment.

We have considered equally for both groups two parameters of inclusion: age, and impairment. Age criteria was restricted to 45 years old because native speakers over the years develop their L1 linguistic knowledge, for example by improving L1 lexical richness through language use, besides, age is a predictor of *Diastratic Variation*[25]. Visual and hearing impairment was a parameter of exclusion because optimal abilities were needed to complete the linguistic experimental tests presented.

3.3.3. Recruitment

Participants were recruited through announcements at the University of Toulouse (France) for group 2 and at the University of Barranquilla (Colombia) for group 1. We posted an online announcement through the University websites for Masters and Bachelor's Degree students at both universities, and websites pages of Spanish Speaking Associations of Toulouse. All volunteers deliberately decided to participate in the study and contacted the experimenter to get a better understanding of the study before experimentation. All volunteers received goodies as compensation for their participation and snacks were provided during the experimentation test breaks.

[25] i.e., when different generations use differently their lexicon, particularly, regarding pronunciation or the choice of grammatical structures that change over the years or in between generations.

Prior to the experiment, participants from group 2 filled in an online questionnaire **(see Appendix 1)** and participants from group 1 filled in a paper questionnaire (see **Appendix 2**) Different questionnaires and sections were presented depending on the group, further description in section 3.3.3.1.

3.3.3.1. Questionnaire for Group 2:

The questionnaire completed by the bilingual group was longer than for controls because for this population there were some additional variables to be taken into account, including LoR, mastery of L2 (French) and the knowledge of other foreign languages different from French.

In the first section, we gathered their personal information, including the participants' nationality. In the second section, we gathered information about the educational background of the population, current studies in France, previous studies and schooling in France or elsewhere, and in their native country. We also collected information about the languages in which all studies were pursued to determine the language that was used the most during schooling years. In this part we also controlled that secondary and primary schooling was in L1.

The third section was called *Language*, in which we collected information about the L1 of the population, their parents' L1, their average skills and possible language impairment. Additionally, we have asked for the length of years living in France without interruptions of more than three months, in other words, the Length of Residence (LoR). We have also reported for the L2 mastery (in months and years) of SL learning before and after their arrival in France. We asked whether participants had tested their level in French language following the CEFR (*The Common European Framework of Reference for Languages*). Additionally, a detailed description of skills in both languages was collected through the completion of a self-assessment grid involving production and comprehension/written production and comprehension.

The fourth section involved the *Description of Language Use*, it included L2 and L1 use, we asked specific questions to evaluate the language preferences in different contexts of language use and to report the daily activities in each language. Moreover, a subsection compared frequency of language use in the L1 and the L2 using scales from 1 to 7. Additionally, we asked participants to evaluate their daily use of language adopting percentages as estimations. Finally, we asked again for self-assessments using multiple-choice questionnaires and open questions about L1 and L2 mastery, L1 and L2 changes and emotional aspects regarding each language.

The fifth section included information about other languages besides French and Spanish, the same subsection scales are used one per languages in order to observe the frequency of use and to determine the context(s) in which the languages are specifically used. The questionnaire was completed online, data was anonymized using a coded number to refer to the participant.

3.3.3.2. Questionnaire for Group 1:

The questionnaire completed by the monolingual group included four sections. In the first section we have collected personal information such as age, sex, origins, educational background, and laterality.

The second section called *Language* gathered information about participant's L1, their parents' L1, their educational background since this might have an impact on the language performance. This section checked also that there were no issues of learning and language disability, hearing or visual impairments.

The third section was related to L1 use, it allowed to define the frequency of reading, writing, and language use in general. The fourth and the last section focused on learning of second languages, the number of years or month of formal instruction, and their levels following the CEFR. The formal support was a questionnaire anonymized for each participant.

3.3.4. General Procedure

The general procedure of the experimentation started with the completion of the questionnaire and the signatures of two consent forms; subsequently, participants could complete the three experiments in L1. All participants gave written informed consent (information letter and consent form may be found in **Appendix 3**.

The experiment lasted for about 2 hours (see **Figure 15** for procedure of experimentation). Group 1 was tested in Colombia in the Computer Lab of University del Atlántico (Barranquilla, Colombia), that was exclusively used by the experimenter so that multiple participants could be tested simultaneously. For group 2, the experiment took place in France in the residence of each participant, and they were tested individually.

After participants had signed the forms, experiments were run in Colombia on a computer ThinkCentre Model A70Z Core i5) and in France on a laptop (HP EliteBook 8470p Core i5) both groups wore a Lapel microphone and headphones (DT770 PRO OHM) in order to facilitate concentration, further description are provided in each task description in section 3.6.

Gap Completion Task (GCT), participants were told to listen to oral stimuli of incomplete sentences. An audible signal indicated them, that they had to complete the sentences orally as fast as possible, using one or two words maximum. Complementary instructions were read before the task begun. The average time invested for this task was 8 minutes, each response was timed for 5.000 milliseconds.

For the Lexical Decision Task (LDT), stimuli were presented on the computer screen, participants were told to choose between yes/no key options for L1 existing and non-exiting words. Complementary instructions were read before the task begun. The average time for this task was 40 minutes, each response was timed for 4000 milliseconds.

For the Judgment Acceptability Task (AJT), participants were asked to read carefully 3 of list of sentences presented randomly on the screen, they had to give especial attention (in term of semantics) to the target word stimuli presented within the sentences in capital letters, they had to evaluate sentences as acceptable or unacceptable using a 1 to 7 scale. Responses for this task were not timed, and the average time for an average reader was 30 minutes.

The whole experiment lasted around 1hour 45 minutes to 2 hours, depending on the length of the pauses taken by the participants and their reading skills, among others (e.g., the length of each pause or the participant's reading skills).

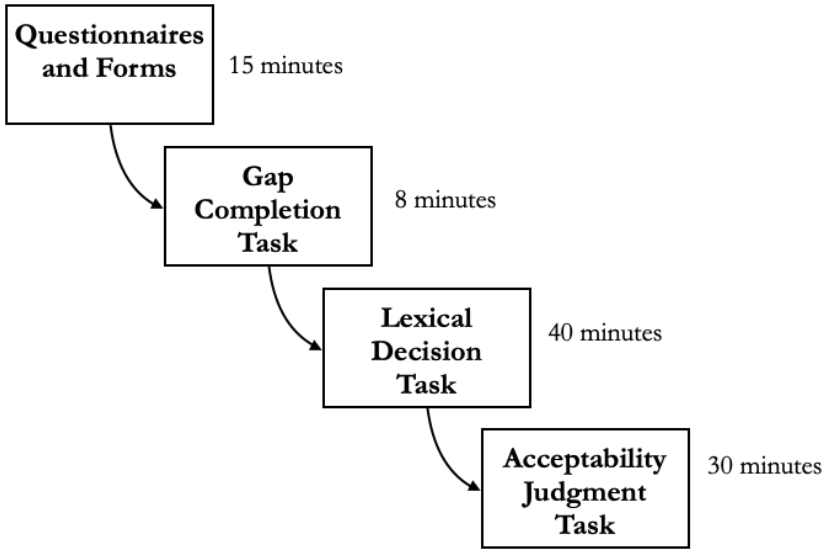

Figure 15 Procedure during experimentation

The order of the experimental section was the same for all participants proceeding from the less to the most constraint in terms of linguistic material (GCT+LDT+AJT). This was motivated by the consideration that the bilingual group (group 2) was supposed to show some sensibility to French word-form and aimed at avoiding induced interference from the stimuli presented in the Lexical Decision Task.

For monolinguals (Group 1) in some exceptional cases, the order of task was inverted (LTD + GCT + AJT) due to restriction in availability of the experimental material (Microphone, Edirol and headphones) for all participants simultaneously. Experiments are detailed in section 3.6.

3.4. Description of the Population

The experimental Group (2) of the present study involves 40 bilinguals. Their L1 is Spanish and their L2 is French, the foreign language that the group used the most in

their daily basis was French for a mean of 0,64/1[26] (see **Appendix 4**). They all lived in France in an L2 dominant context at the moment of testing. For the experimental group the average age was 30,74 years, ranging from 21 to 43 at the date of experimentation. 30 females and 10 males participated (see **Table 2**).

The control group (1) involved 53 monolinguals Spanish living in an L1 dominant context in Barranquilla (Colombia) at the moment of testing. For this group the average age was 22,04 years ranging from 18 to 37 years at the date of testing. There were 28 females and 25 males (see **Table 2**).

Table 2 Distribution of Sex and Age of the Participants

Group	Females	Males	Mean Age	Minimal Age	Maximin Age
Bilingual	30	10	30,74	21	43
Monolingual	28	25	22,04	18	37

3.4.1. Experimental Group

3.4.1.1. The Length of Residence LoR for Bilingual Group

The mean Length of Residence (LoR) corresponding to immersion of the population in France is 4 years and 6 months (see **Table 3**). The maximum LoR is 13 and 11 years that correspond only to 5% of the population. The minimum LoR is one year that relates to 10% of the population. LoR in this study is shorter than in other studies referring to L1 restructuring, where an extended period of bilingualism of more than 15 years of immersion would explain L1 restructuring (Grosjean & Py, 1991). We propose to study restructuring processes in the early years of immersion.

Table 3 Distribution of the LoR of for group 2

Descriptive Statistics	
	LoR
Valid	40
Mean	4,63
Std. Deviation	2,72
Minimum	1,00
Maximum	13,00

3.4.1.2. Nationality

The bilinguals were born and lived in Latin America for almost 20 years, a period corresponding roughly to Primary and Secondary Schooling. Their nationality involved a

[26] Mean score was attributed at calculating the number responses (1-5 scale) provided by the participant in the questionnaire presented (14 questions) describing the contexts in which a particular language was used. The same questionnaire in presented for the L1, the L2 and others foreign languages learnt by the speaker.

majority of Colombians (40 %,) and Venezuelans (25 %), as well as 12,50% Mexicans, 10,00% Chileans, 5,00% for Argentinians, one Bolivian, one Paraguayan and one Guatemalan represent 2,50 % each (see **Table 4**).

Table 4 Origins of the Bilingual Group

Group	Nationality	%	Participants
Bilingual	Colombian	40,00%	16
	Venezuelans	25,00%	10
	Mexican	12,50%	5
	Chilean	10,00%	4
	Argentinian	5,00%	2
	Paraguayan	2,50%	1
	Bolivian	2,50%	1
	Guatemalan	2,50%	1
		100,00%	40

It was not possible to restrict origin of participants to Colombians because there is a limited number of Colombians living in Toulouse (France) that fulfilled the inclusion criteria defined at the beginning of the study. However, during the construction of the linguistic material for the three experiments, regional variants (e.g., diatopic variation) were excluded and a standard variant of Spanish from Latin America was followed. Additionally, a neutral accent was chosen during the recording of auditive stimulus by a professional singer [27], the accuracy of the sentences was tested before the final experimental stimuli selection was made.

[27] We would like to thank Fiorella Mancilla for her time and dedication during the recording of the auditive stimuli presented in the Gap Completion Task.

3.4.1.3. Educational Background

The level of instruction of the Bilingual Group is detailed in

Table 5. This group was mostly formed by Master graduates or Graduates currently involved in professional life. Master students represented 30,00%, graduates in France represented 25,00% of the participants, PhD Candidates constituted 15,00% and bachelor's degree students represented 15,00%. 12,50% had finished secondary school without pursuing education and 2,50% had dropped-off University before arrival in France.

Table 5 Educational Level Background of the Bilingual Group

Group	Educational level	%	Participants
Bilingual	Master Students	30,00%	12
	Graduates	25,00%	10
	Bachelors	15,00%	6
	PhD Students	15,00%	6
	Secondary School	12,50%	5
	Dropped off	2,50%	1
		100,00%	40

3.4.1.4. French Language Mastery

Language assessment and mastery was mainly determined by the CEFR (The Common European Framework of Reference for Languages). The levels of mastery of French are described in **Table 6**.

Globally, half of the population had an intermediate level according to the CEFR, B2 Level representing 50,00%. 22,50% of the group were advanced learners of French with C1 Level. 2,50% represented C2 level; 5% had attested for B1 level and 5% for A2 level. 15,00% of the participants did not have a formal CEFR test to define their level of French. The self-assessment of these participants indicated a B2 level, the mean scored obtained by this subgroup was 0,68 /1 (ranging from 0,55 to 0,80).

However, the official level in French of the participants is not fully representative of language use or language mastery itself. LoR plays a determinant role, as well as the activities of the participants during their immersion and the languages involved in activities such as studying French at the Language Center of the University, preparing a Bachelor or master's degree in the L2, or their current professional activity in France. We gathered information about daily activities in L2 (reading and writing habits, watching TV, cultural activities, languages preferences, etc.). These factors taken together provide a more comprehensive view of L2 use and mastery.

Table 6 CEFR L2 Mastery Distribution

CEFR L2 level	%	Participants
A2	5,00%	2
B1	5,00%	2
B2	50,00%	20
C1	22,50%	9
C2	2,50%	1
NA	15,00%	6
	100,00%	40

Additionally, we took into account information that works as indicators of L2 mastery including the length of study of the L2 before the participants' arrival in France and once they arrived. In all cases, the participants learnt French either in their home-country or in France, in some cases, in both. Participants received formal language instruction in French in languages institutes such as *Alliance Française*, through University Language Programs and Languages Learning Associations in Toulouse.

3.4.1.5. Other Foreign Languages

We have also collected data concerning the knowledge of other foreign languages (**Table 8** and **Table 7**). As it can be noted, the mean score of use of English (0,49/1,0) was inferior than the use French as a Second Language (mean score of 0,64/1,0) (see **Appendix 5** for a detailed description of use of English and length of language instruction).

Table 7 Summary of languages learnt by group 2

Language	Percentage	Participants	Mean Use
English	77,50%	31	0,49
Portuguese	5,00%	2	0,43
Italian	2,50%	1	0,31
None	15,00%	6	
Total	100,00%	40	

As detailed in Table 7 English is the additional language learnt by most participants, representing 77,50% (31 participants) of the target group. Their level in English is detailed in **Table 8**, participants corresponded mostly to B1 with 22,50% and B2 with 17,50%. A1 level represented 12,50% of the group; A2 corresponded to 5,00%. Advance learners C1 represented 12,50% and C2 represented and 5,00%.

15,00% of English learners have attested that their level corresponded to Basic/Elementary and Secondary School level, their levels were between A2 to B2 (6

participants). For the 7,50% of the group (3 participants) the length of studies of English was 1 year, the minimum time of formal language instruction in our data.

Table 8 Description of English CEFR Levels

	L3 English	Participants
A1	12,50%	5
A2	5,00%	2
B1	22,50%	9
B2	17,50%	7
C1	12,50%	5
C2	5,00%	2
NA	2,50%	1
Total	77,50%	31

Regarding more advanced learners that corresponded to C2 and C1 levels, the maximal value for formal language learning was 20 years for 2 participants (5,00%) and the minimal value was 9 years for 1 participant. Regardless of the level of mastery of English reported by the participants, the language that participant attested to be used the most was French.

5 % of the participants reported learning of Portuguese as a foreign language **Table 9**). The self-assessments grids suggested two levels, A1 and B2 and a length of instruction of 3 months and two years. Following the information provided by the participants, the use of Portuguese in France is restricted to few contexts with a mean score of 0,43/1. Italian was represented by 2,50 % for 1 participant, self-assessments grid pointed a C1 level for this participant, however, language use of Italian is restricted for 0,31/1 score of contexts of language use (**Table 7**). 15% of the population did not report knowledge of other foreign languages than French.

Table 9 Other L3 CEFR Levels

	NA L3	L3 Italian	L3 Portuguese	Participants
A1			2,50%	1
A2				
B1				
B2			2,50%	1
C1		2,50%		1
C2				
NA	15,00%			6

3.4.2. Control Group

The participants of this group were born and lived in Colombia at the time of the experiment. They had never staid in a non-Hispanic country for more than one month, their native language was Spanish. Moreover, there weren't any indigenous descendent participants, their parents' L1 was Spanish, and their nationality was Colombian.

3.4.2.1. Educational Background

Participants were currently students at the University of Atlántico (Barranquilla, Colombia), where they were enrolled in different semesters and programs, most of them in their 5th semester (ranging from 3rd to 10th semester). 5 participants had already obtained their bachelor's degree, which represents only 9,00% of the population. However, they were still active in the academia because at the moment of testing they were studying for an extra year in a postgraduate specialization.

The programs in which the participants were recruited are detailed in Table 10. Engineering programs (industrial, mechanic) represented 30,19% of the group. 9,43% were involved in Law and Chemistry Engineering programs. Arts and Spanish Literature programs corresponded to 7,55% per program. History and Sociology programs represented each 5,66%. Business, Health assistance, Management and Especial Education are less representatives Programs with 3,77% each.

More disperse programs with 1,89% included: Architecture, Preschool Teaching, Economics Philosophy and Human Sciences Programs.

Table 10 Educational Level Background of the Monolingual Group

Group	Educational Program	%	Participants
Monolingual	Engineering	30,19%	16
	Chemistry	9,43%	5
	Law	9,43%	5
	Arts	7,55 %	4
	Spanish Literature	7,55%	4
	History	5,66%	3
	Sociology	5,66%	3
	Business	3,77%	2
	Health assistance	3,77%	2
	Management	3,77%	2
	Especial Education	3,77%	2
	Architecture	1,89%	1
	Preschool Teaching	1,89%	1
	Philosophy	1,89%	1
	Economics	1,89%	1
	Human Sciences	1,89%	1
		100,00%	53

This group was considered as monolinguals because they fulfilled the criteria described in section 3.3.2. They had no knowledge of French and were currently living in an L1 dominant context, with no use or contact with French as an L2.

Nonetheless, 40 students had learnt other (s) foreign languages that are listed in **Table 11**. English was studied by 66,04% of the population (35 students), which is due to the recent educational government policy for Public University Programs that aims at

training bilingual professionals. Bilingualism in this context is still developing and the process might take longer than expected from the government policies.

When we analyze frequency of foreign language use of these students, questionnaire responses indicated that they used very rarely English in their daily life, only 3,77% (2/35) attested that they used 'frequently' or 'very frequently' the language, this minority have studied the language for 2 to 3 years and context of use is mostly limited to the classroom and rarely in real life situation.

Table 11 Distribution of L2 formal instruction

L2 instruction	%	Participants
English	66,04%	35
English and Japanese	3,77%	2
English and Portuguese	3,77%	2
English and German	1,89%	1
none L2	24,53	13
	100,00%	53

Other languages are underrepresented, as complementary second languages besides English. Japanese was represented for 3,7% specially in listening categories, Portuguese corresponded to 3,7%. In contrast, 24,53% of the group has reported for no knowledge of second languages. This piece of information supports previous statement about the passive bilingual process regarding Postgrad educational policies in Colombia.

3.5. LINGUISTIC MATERIAL

Grosjean & Py (1991) have shown the relevance of measuring the resistance and flexibility of change of different linguistic structures. But contrary to the grammatical axes proposed by Grosjean & Py', our study, with focus on semantic and not on grammatical aspects, will investigate internal features of the lexical item that form our material. Thus, we have selected 2 internal features of the cross-linguistic pairs of words that are involved in semantic extensions of the L1. The words were analyzed and described in both L1 and L2 in order to calculate the Orthographic Neighborhood Density and Morphological Family Size in French and in Spanish. Comparing and opposing these factors will allow us to study the lexemic and lemmatic levels during processing of semantic extensions.

The selection of semantic extensions that constitute our linguistic material builds on attested semantic extensions of French (L2) patterns in Spanish (L1). Our main source is the paper of Quilis et al (1982) where a list of semantic extensions called "*Semantic Gallicism*" in Spanish was established. This work is unique of its kind because it involved a longitudinal study focused on language changes (mapped in the L2 French) in the productions of migrants living in France. We prioritize the semantic extensions that were also reported in Quilis et al.'s study.

The construction of the experimental material was based on a list of attested stimuli in Quilis Study, among which final stimuli controlled for two main criteria, i.e. Density of Orthographic Neighborhood and Size of the Morphological Family in L1 and in L2, were selected. The final selection of stimuli is detailed in next section.

Each task involved specific constraints with respect to stimulus adaptation and presentation. For the GCT, the linguistic material were incomplete sentences to be orally completed and the stimuli were targeted to integrate the sentences as a complement response. Hence, in this task, the targeted responses to be used for completion corresponded either to a canonical response or to the expected semantic extension. For the LDT, the linguistic material were primes matched with target words presented in different conditions (outlined below 3.6.2.2.2). The target word could be a semantic extension or not. For the AJT, the semantic extensions were integrated into sentences following different conditions sentences.

3.5.1. Stimuli Construction

The initial list of stimuli was reviewed for a deeper understanding of the internal features in each language, final stimuli selection will be described in the next section 3.5.4.

At a first stage, we conformed our stimuli taking into account their Orthographic Density in French and in Spanish. These were calculated through the multilingual database *Clearpond* (Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities, Marian et al., 2012) that is a freely available for researchers. *Clearpond* constitutes a valuable resource to calculate and compare

crosslinguistic orthographic and phonological neighborhood density. This database is available in five languages: Dutch, English, French, German and Spanish, and allowed us to compare our target languages French and Spanish.

Additionally, this database provides information about word frequency, length of words, the type of neighborhood (orthographical vs phonological), the relation between the source word and the neighbor (i.e., addition, deletion or substitution). Corpora are derived from television and film subtitles to capture spoken word frequency of the 5 languages. The main matrix of the corpora is SUBTLEX in Dutch, in English, French, German and Spanish.

Furthermore, the developers of this database paired language misspellings and language intrusion in the target language corpora which were verified and corrected through reference dictionary in each language. Homographs in French were reduced to a single entry and their frequencies were cumulated per million, e.g., for the homograph '*est*' meaning 'east' as well as the third person singular of the verbe '*être*' (to be).

Following Marian et al., (2012), corpus was equated by word frequency, which includes 27,775 most frequent words in five languages above a frequency threshold of 0,34 per million, this way, enabling corpora equation and comparable average frequencies across languages. Additionally, a normal distribution of data was respected during the construction of the Clearpond data base.

We will present the constitution of the corpora in Spanish and French focusing on how the orthographic neighborhood measures were calculated in Clearpond. There are some specificities about these corpora, namely, French shows a greater number of phonological neighbors in comparison to the other four languages (Marian et al., 2012, p. 8). Phonological and Orthographic neighbors in Spanish gathered more similar entries than the other languages. Regarding orthographic neighborhood of french words, it gathered the greatest number of neighbors obtained by addition and deletion than the other languages (Marian et al., 2012, p. 10).

For our study, we choose to not take into account phonological neighbors in French or Spanish for two reasons. First, in terms of phonological features, the orthographic consistence is higher in Spanish than it is in French, i.e., in Spanish a single phoneme maps generally into a single grapheme, which is not the case for French. Second, the multiple homographs and phonological neighbors of French will represent a great discrepancy with Spanish where homographs and phonological neighbors are more limited. This typological difference between the two languages suggests that French's phonological neighbors differ significantly from Spanish. This was noticed when comparing stimuli in each language using Clearpond, where the number of phonological neighbors was systematically higher in French than in Spanish. The relation between orthographic and phonological neighbors is not as easy to settle, indeed, it depends on the language's specificities. For example, not every orthographic deletion will lead to phonological neighbors. The relation between orthographical and phonological neighbors depends largely on orthographical transparency and consistency of each language.

The Clearpond database was chosen for stimulus selection because it was constructed in order to allow for cross-linguistic analysis. Clearpond is in fact comparable, its corpora are controlled and equated across languages. At the beginning of our study, we have witnessed difficulties during the constitution of linguistic material when using two different databases for Spanish and for French. Their unbalanced corpora size and the use of different corpora sources constituted a methodological bias. The use of Clearpond has facilitated crosslinguistic analysis, reducing at the same time methodological bias since corpora have the same derivations. Besides, the values for word frequency and language use can be considered representative of real spoken language.

There were however some words that were not available in Clearpond, for these we have used Spanish SUBTILEX for word frequency. Since SUBTIBLEX is based on the same corpora as Clearpond, methodological bias was limited. Additionally, for Spanish we have searched on ESPAL database (Duchon et al., 2013) for neighborhood, since this corpus is derived from online written data and subtitles, we reported for specific stimuli using a different coding. Espal provides the same information as Clearpond for Spanish, including neighborhood type (by addition, deletion or substitution), frequency, and neighborhood density. Regarding the Morphological size of words, we calculated frequency using SUBTILEX for French and Spanish.

3.5.2. Neighborhood Density Stimuli

In section 1.2.2.1 we have discussed the role of neighborhood density during bilingual lexical access. Taking into account that activation levels will be stronger in the language that has the highest neighborhood density (Costa et al., 2006), we assume that lexical selection may be faster in one of the languages of the bilingual.

To match our stimuli based on the list of attested semantic extensions by Quillis et al (1982) with respect to neighborhood density, they were searched in Clearpond and the list was reduced depending on their densities in L1 and L2. Then, we compared the neighborhood density of crosslinguistic words (e.g., *enerver* and *enervar*) involving semantic extension using Excel and created our own data (see **Appendix 6** describing neighborhood density of the stimuli selected). In the next phase, the words were categorized into two types of stimuli. The first type had dispersed neighborhood in L1 (Spanish) and a dense neighborhood in L2 (French), i.e., *dominant neighborhood in L2*. The second type had a *dominant neighborhood in L1*, with disperse neighborhood in L2 (French), and dense neighborhood in L1 (Spanish). The main objective of this stimuli construction stage was to determine in which language density was stronger, and to constitute a data base adapted to our experimental design.

3.5.3. Morphological Family Size Stimuli

The main objective during this stage of construction of the stimuli was to be able to compare and oppose stimuli with respect to Morphological Family Size in both languages as was done for Neigborhood Density. However, contrary to neighborhood density, we were not able to find a database available to automatically generate morphological family size of words. We have created, as previously, our own database using the available resources. We have searched manually for our attested list (Quillis et al 1982) using two etymological dictionaries per language, for French (Dubois et al., 2011) and for Spanish (Corominas & Pascual, 2008). A research trainee participated in this stage of stimulus construction. Dictionary query provided an exhaustive list of derivate words that were etymologically linked to the target words.

However, the values obtained were not entirely representative of language use since they included low frequency words and very distant family members. In order to reduce the list into more familiar data, based on the results of the dictionaries query, we conceived an online test used to filter the members of the morphological family that were actually semantically related to the target stimuli. This has enabled at the same time the exclusion of less frequent derivate words. In this online test (see **Appendix 7**), native speakers of French and Spanish judged a list of word derivates. They were asked to strikethrough from the list of words the derivates that were not semantically related to the target word. Knowledge and frequency of use of the native speakers was also taken into account in order to exclude infrequent words.

6 French native speakers and 6 Spanish native speakers participated. These judges were bachelor's students and PhD candidates in linguistics, their educational background contributed to an accurate evaluation of the words. (See **Appendix 1** in which one test per language was completed). Subsequently, we calculated word frequencies of the derivates to obtain the cumulated frequency of our corpus. **Appendix 8** describes Morphological Family Size data comparisons across languages (crosslinguistic word pairs) of the stimuli used in our experimental tasks. Using the available information about Neighborhood Density, we compared dominance with respect to both factors in L1 and L2 (Morphological family Size vs Neighborhood Density) in order to conform our final stimulus selection detailed in **Appendix 9**, this list is going to be detailed in section 3.5.4.

3.5.4. Final Stimulus Selection

Linguistic material was presented differently depending on the experiment as described in section 3.5. Here we are going to describe the stimulus selection that was obtained for each experiment.

For the GCT, we selected 42 target responses for the incomplete sentences. These took into account mostly neighborhood density criteria and in some morphological family criteria, since it was difficult to adapt the potential hypothetical responses at the end of the gap sentences constructed. (See example provided in **Appendix 10** in for an online task used to target responses for the GCT). The targeted sentences presented as linguistic material was not perfectly equated in either Neighborhood Density nor Morphological Family Size criteria. However, some of the stimuli integrated were controlled for Neighborhood measures (see **Appendix 1**).

For the LDT, the selected stimuli took into account both criteria: Neighborhood Density and Morphological Family Size for L1 and in L2, resulting in the establishment of the following stimulus categories, see **Appendix 13** for the description of the final selection of stimuli presented.

- 11 *Equal MN:* Equivalent Neighborhood Density and Morphological Family Size in Spanish and French
- 8 *High M Low N_Sp* : Stimuli of High Morphological Family Size and Equal Neighborhood Density in the L1 Spanish
- 11 *High M Low N_Fr*: Stimuli of High Morphological Family Size and Equal Neighborhood Density in the L2 French
- 12 *High MN_Fr* : Stimuli of High Morphological Family Size and Neighborhood Density in L2
- 4 *High MN_Sp* : Stimuli of High Morphological Family Size and Neighborhood Density in the L1 Spanish
- 8 *High N Low M_Fr*: Stimuli of High Neighborhood Density and Equal Morphological Family Size in L2 French
- 7 *High N Low M_Sp*: Stimuli of High Neighborhood Density and and Equal Morphological Family Size in L1 Spanish

There was a total of 61 stimuli. In order to balance the number of stimuli per category we added 23 distractor words for this task. (See **Appendix 9** for a detailed description of targets and primes, distractors, words presented during Experiment 2).

For the Neighborhood Density, CLEARPOND automatically selects neighbors that have the highest frequency, providing the number of neighbors corresponding to each stimulus. During the selection of the final stimuli, we opposed dense and dispersed neighborhood across languages. For the Morphological Family Size, we calculated the number of morphological derivates, for which we have also calculated the cumulated frequency.

The interlingual word pairs were categorized as dominant in L1 or L2 considering both factors. For instance, the word '*coraje'* in Spanish has 2 morphological derivates and the French equivalent '*courage'* has 6 morphological derivates, hence we would consider that the pair '*coraje'*/ '*courage'* has an L2 dominant morphology. The minimal difference value is 2 because the number represents the cumulated frequency for morphological information and for neighborhood of the most frequent words in the CLEARPOND database.

For the JAT, we needed 65 stimuli; most of these were already presented in the previous tasks. Stimuli distribution is different due to some restrictions of this task: stimuli must fit in a sentence, in some cases stimuli were presented multiple times, in order to explore different meanings see section 3.6.3.2.

3.6. Experimental Tasks

Each of the three experimental tasks was presented to the participants using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). They were presented using an ELITE HP laptop for the Spanish-French bilinguals living in France and on 9 desktop computers ThinkCentre Model A70Z Lenovo for Hispanics living in Colombia. As mentioned before, the experimental tasks were presented and completed in Spanish (L1) in the same order for all participants as follows:

1. GCT
2. LDT
3. AJT

The stimuli were presented differently depending on the task as discussed in the previous section. Stimuli could be produced as a response, presented as a target or as an isolated word in a sentence. In this section we will detail the procedure of each task, the main objectives, and the specific procedures of each experiment.

3.6.1. Experiment 1 Gap Completion Task

The GCT is a traditional written exercise in foreign language teaching that consists in completing a sentence with a missing word. It is used to evaluate specific contents, mostly vocabulary and verbs. This experiment was adapted in order to gather oral

production of semantic extensions. This allowed us to compare oral production in L1 of the two groups.

As explained previously each task corresponds to a specific skill involved in language learning, for the GCT we focused on oral production and particularly, on the levels of lexical activation during L1 production. We aimed to investigate whether the competition level between the languages will allow the L2 to interfere during speaking when completing the CGT in L1. Since the sentences could be completed in L1 or L2 without semantic ambiguity nor grammatical restriction regarding crosslinguistic structures, L2 productions may be considered "accurate" within the sentence context.

The incomplete sentences were pronounced by a female professional singer, sentences were recorded and orally presented in E-Prime. One to two-word oral responses were registered during the experiment using two supports: directly with E-Prime 2.0 recording for 5,000 (ms) and with a professional recorder (Edirol 24 Bit Wave Recorder and Play). Responses were analyzed using PRAAT (Boersma & Weenink, 2020) in which responses times invested by participants were collected.

3.6.1.1. Objectives

The aim of this task was to observe L2 influence on the L1 in controlled speech. In fact, semantic extensions appear mostly in oral production since the processes involved during speech are online.

3.6.1.2. Linguistic Material

The examiner and a research trainee (both Spanish native speakers) have constructed the sentences conforming the linguistic material presented in the GCT. The sentences were constructed one by one starting from the target response arising systematically at the end of each incomplete sentence. Expected responses could be of two types: neutral or semantic extensions (as the result of L2 lexical activation during an L1 task). During the construction of the linguistic material (see **Appendix 11**), we avoided ambiguous sentences and provided contextual elements to facilitate participants to successfully complete the missing word.

e.g. Aún no he salido de casa y mis padres ya me están preguntando a qué hora voy a _____.

*(I haven't left home yet and my parents are already asking at what time am i going to*_____*.)*

The canonical response expected for this stimulus was '*regresar*' ('*to come back*'). The semantic extension expected here as a complement for the incomplete sentence was '*entrar*' ('*to come in',* or, *'to get inside somewhere'* in Spanish).

'*Entrar*' is mistakenly associated with the French word '*rentrer'* which means '*to go back in'*. Here, the sense of '*rentrer'* in French is extended in Spanish '*entrar*', which is restricted to '*to come in',* or, *'to get inside somewhere'* and not '*to go back in'*. The variation of '*entrar*' expressing the meaning of '*rentrer'* is very subtle and it is non canonical in this particular context. The impact of these subtle differences expressing a semantic extension (e.g., the use the use of '*entrar*' instead of '*regresar*') will also be evaluated in the AJT (see section 3.6.3).

3.6.1.2.1. Sentence Construction

When constructing the stimulus sentences, the length of the words was equated, and target responses were always placed at the end of the sentences in order to facilitate the measurement of the response times. Some semantic extensions were difficult to place as a response at the end of a sentence and were excluded. Sufficient informational context was provided for each sentence to ensure that participants might find the target response easily.

In order to control that the targeted responses were homogeneous and accurate, we conducted a pre-test using online forms. Two groups of monolingual participants (French and Spanish) were asked to complete a series of incomplete sentences in a written online exercise, using one to two words. Then, following the responses gathered, (see **Appendix 10**) modifications were applied to the linguistic material. Subsequently, we selected the final version of the linguistic material, resulting in 42 sentences in (see **Appendix 11**). The linguistic material was constructed respecting the two possibilities described previously (canonical and uncanonical productions, e.g., '*entrar*' vs '*regresar*').

A Spanish native speaker who is a professional singer recorded the final sentences with an Edirol 24 Bit Wave Recorder and Play. During the recording section, sentences were carefully pronounced and articulated. Intonation and punctuation were taken into consideration as well. A beep generated by the phonological research team at the University was added at the end of each sentence to mark the end of the stimulus and the beginning of each participant's oral production. This beep also triggered the calculation of the response time during data analysis in PRAAT.

3.6.1.3. Procedure During Experimentation

Before the experiment started, headphones DT770 PRO OHM, a microphone and the Edirol recorder were settled by the examiner on the computer. At the beginning of the experiment the participants read brief instructions (see **Appendix 14**) that were previously orally explained. Each trial started with a fixation cross in the center of the screen, then an image indicated the presentation of the stimulus sentence, and finally the

beep invited the participants to complete the sentence. The experiment involved 42 stimulus sentences. A pause was provided in the middle of the experiment; each participant defined the duration of the pause.

3.6.1.4. Time settings

Stimuli were presented in three blocklists that were composed of 14 stimuli each. Groups of stimuli were organized in the blocklists depending of the duration of the oral stimuli, the first block lasted 5889 milliseconds (**Figure 16**) the second 6583 milliseconds (**Figure 17**) and the third 7522 milliseconds (**Figure 18**).

This experimental design created in E prime 2.0 involved 7 samples (1 cycle x 7 sample/cycle) which means that 1 cycle equals 7 samples, each blocklist had 2 cycles for a total of 14 samples whose oral input durations were the same as indicated above. The samples were the number of stimuli and the cycles allowed to automatically randomize the order of a particular group of stimuli or samples. In fact, each participant had a unique stimulus order; the random selection was repeated after reset, which means that the random cycles were always reinitiated in each blocklist (**Figure 16**). For instance, the order of presentation of stimuli was semi-controlled. In **Figure 16** , **Figure 17** and **Figure 18** we describe the duration of each blocklist and the time settings of each oral input and output recording.

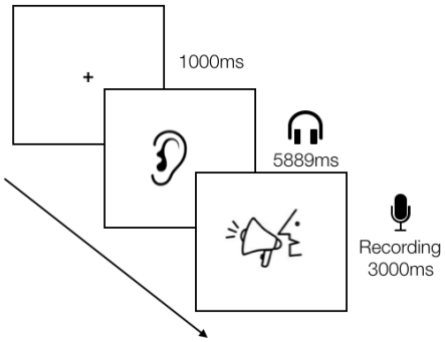

Figure 16 Time settings of blocklist 1

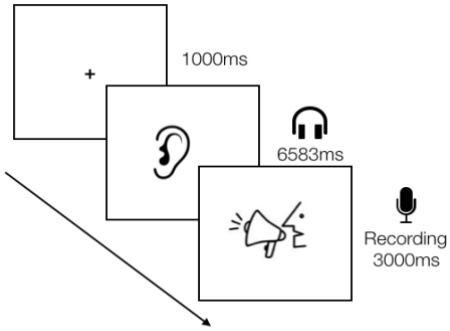

Figure 17 Time settings of blocklist 2

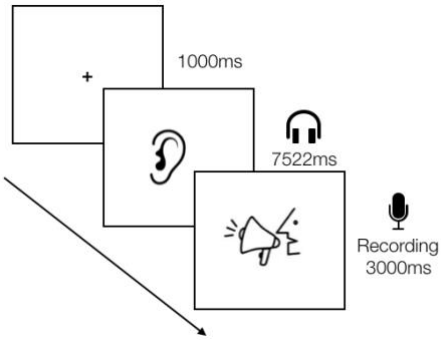

Figure 18 Time settings of blocklist 3

3.6.2. Experiment 2 Lexical Decision Task (LDT)

The Lexical Decision Task (LDT) is an experimental paradigm used extensively in psycholinguistic research in order to measure the time necessary to process a word. The central focus of the LDT is the priming effect which is based on the assumption that phonetic, semantic or formal features presented in a word will affect the processing of another word. The study by Meyer & Schaneveldt (1971) was one of the first to evidence semantic priming effect in monolinguals, bilingual studies involving cognates and crosslinguistic translations were multiplied in the late 90's and 20's (e.g., de Groot & Nas, 1991; Grainger & Segui, 1990; Lalor & Kirsner, 2001) until recent work (Dijkstra et al., 2010; Mulder et al., 2014).

The procedure of an LDT involves a prime word (e.g., *carta*) and a target word (e.g. *marta*). The participants must decide whether the target words exist or not in a particular language, in this case Spanish. The primes words are presented previous to the target words during a short time span, in this experiment we choose 50ms, involving that information will be in all cases processed unconsciously by the participants (Sánchez-Casas & García-Albea, 2005).

In an LDT different association between primes and targets are possible. In our particular case, different primes were used depending on condition. This procedure allowed us to compare the priming benefits between conditions, in other words, to know whether participants respond faster or slower to a stimulus depending on the condition they belonged to.

3.6.2.1. Objectives

This LTD was constructed to observe three different levels of processing of semantic extensions. Therefore, we repeated the same target in three conditions, associated with different primes.

1) In condition 1, we aimed to observe whether the word-form overlap between languages facilitated activation of the L2 during processing in L1.
2) In condition 3, we tested whether the semantic overlap between L1 and L2 would have an effect on processing.
3) And finally, in condition 2, we asked whether semantic links between the primes and the targets are determined only by the L1 in opposition to the previous condition. More details about the rationale of the different conditions of the LDT are presented 3.6.2.2.2.

A second level of analysis is focused on the type of target presented independently of the condition. Does L2 Dominant Morphological Family Size influence negatively L1 processing? Or does L1 Dominant Morphological Family size facilitate processing in L1? Otherwise, are morphological criteria better predictors of L2 transfer than Neighborhood Density? Does L2 dominant Neighborhood density influence negatively L1 processing? Or does L1 dominant Neighborhood density facilitate processing in L1?

3.6.2.2. Linguistic Material

The linguistic material of the LDT was comprised of targets, primes, pseudo-word and distractor words. They were constituted by the researcher and a research trainee during his research internship in linguistics[28]. The research trainee was a student of linguistics at the university of Toulouse 2, he is a native speaker of Spanish and highly proficient in French and other languages.

We have constituted the primes following the length of their respective targets. The targets are described in section 3.5.4 and defined as the result of the database queries and contrasting criteria in L1 and L2 (Morphological Family Size and Neighborhood Density).

Stimuli were presented automatically randomized in three lists in order to control the distance between repeated targets. Each blocklist had an equal number of words, pseudo-words and distractors. In each case we presented primes and targets. The statistical analysis will focus exclusively on the data of primes and targets words, distractors were excluded from the analysis.

We created 3 blocklist for this task, in which we presented 168 samples (1 cycle x 168 samples/cycle) meaning that 1 cycle equals 168 samples including words, pseudo-word sand distractors words. The random selection was made automatically by the software for each blocklist. Thus, 3 random orders were created, one per blocklist. In each of the three lists composed of 168 samples, each target was matched with a prime (**Appendix 15Error! Reference source not found.**), for the Architecture of Experiment 2).

3.6.2.2.1. Targets

There were 84 stimuli in L1, 61 target words and 23 distractors words. Target words were divided in different categories depending on the internal characteristics of the crosslinguistic words such as Neighborhood Density and the Size of the Morphological Family in L1 and L2 (see the column Target in in **Appendix 9** for Experiment 2).

3.6.2.2.2. Primes

Each target was associated to three types of primes that matched to three conditions:

Condition 1; The primes presented in this condition were words that were formally (orthographically and phonologically) related to the target word. Here semantic relationships between primes and targets were excluded. For example, for the Target

[28] We would like to thank Alejandro Hernández Jaramillo for his invaluable contribution to this research during the construction of the linguistic material of Experiment 2 and 3. His constructive suggestions were a key element during this stage of construction research.

'*Coraje*' ('*anger*') we presented as a prime the stimulus '*Rodaje*' ('*shooting*'), thus, prime and target had a word-form overlap without semantic overlap in L1.

Condition 2: The primes in this condition were semantically related to their respective target words, for example, for the target '*Coraje*' ('*anger*') we presented as prime '*Enojo*' ('*anger*'), both words being semantically related in L1.

Condition 3; In contrast, the primes in the last condition were associated semantically to the targets in case of semantic extension from the L2 to the L1 only, here the meaning in L1 was uncanonical and opposed to the meaning association expressed in the primes presented in condition 2. For example, for the target '*Coraje*' ('*anger*') we presented as prime the word '*Valor*' ('*bravery*'), the meaning of French '*Courage*' being extended to '*Valor*' in Spanish.

Prime words were chosen to match with the number of vowels and the length of the target word that were associated to it. We selected primes from the data obtained during the construction of the linguistic material, neighbors' candidates that overlap with the word form of the target stimuli were selected as primes for condition 1. Morphological family member candidates related to the targeted meaning were selected as primes for condition 2. To constitute the uncanonical semantic relationship between primes and targets for condition 3, we used the underlying explanations of Quilis et al. (1982) of semantic extensions and adapted the stimuli to match with the target words associated (see Columns Primes cond 1, Primes Cond2 and Primes Cond 3 in **Appendix 9** for Experiment 2).

3.6.2.2.3. Distractor words

During final data selection, we opposed two internal words features: Neighborhood Density and the Size of Morphological Family. However, there were not enough stimuli to fully complete the 7 categories (listed earlier) and set in an equal number. Consequently, we had to include distractor words in order to balance de total number of stimuli. Thus, 23 distractors words were added to the 61 target words to pair the stimuli set and equalize the number of stimuli per category, with a result of 12 stimuli per category. From this basis, we created target pseudo-words and their respective primes detailed (see distractor stimuli **Appendix 9**). Distractors words and pseudo-words were excluded from analysis.

3.6.2.2.1. Pseudo words

Pseudo words were constituted manually respecting the phonotactic rules of Spanish following the same procedure as for prime words. All pseudo words were based on existing words in Spanish, the meaning and the words became pseudo words after we replaced syllables, consonants, or both. Accents were not modified, and the lengths of words were always matched with the length of the target word. We verified whether the words existed in the dictionary (RAE, 2020). Pseudo words were presented as target and primes respecting the priming categories described in section 3.6.2.2.2. For example, for

target '*Coraje*' the target non-word is '*Tarajo*'. For condition 1, the Pseaudo words presented as prime 1 (word-form overlap) is '*Marajo*'. For condition 2 we used '*almijo*' and for condition 3 '*cabolla*'. Since semantics or orthographic referents are inexistent for Condition 2 and 3, pseudo-words were less calqued from existing stimuli.

3.6.2.2.2. Conditions

As explained previously, the main principle of the LDT is to compare different prime conditions in order to establish whether there is a priming advantage depending on the specific features of each condition. The stimuli schema was divided in three blocklists.

The three conditions corresponded to the three types of prime:

The first condition represents a prime word that overlaps with the form of the target word. E.g. The target word '*exprimir' (to express)* in Spanish (*DOMN_FR*) is associated with the prime word '*reprimir*' *('to repress'*). This condition was called *form overlap* (see **Figure 19**).

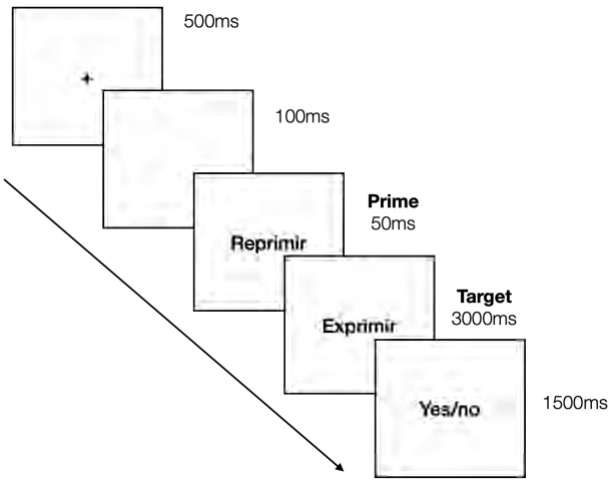

Figure 19 Condition 1: Word-Form Overlap

The second condition is represented by a semantic prime matched with the same target presented before (*'exprimir').* In this case, the prime was '*apretar*', a synonym of '*exprimir',* both words mean 'to squeeze' in Spanish. This condition was called L1 meaning, since it describes a semantic priming using canonical L1 meaning (see **Figure 20**).

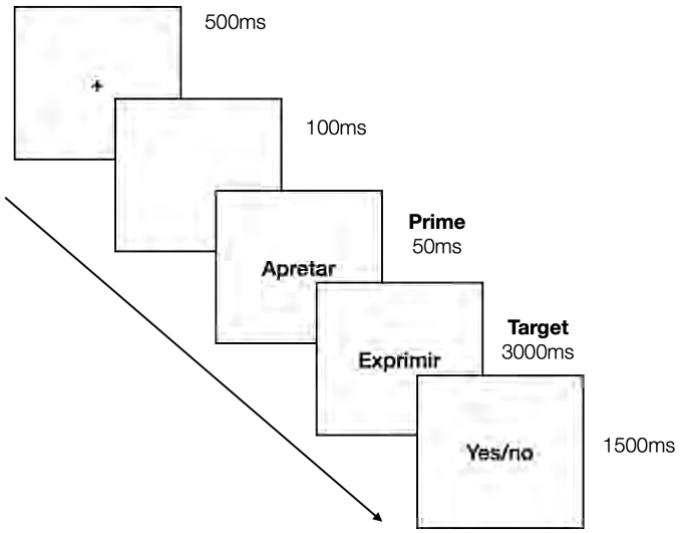

Figure 20 Condition 2: L1 Meaning

The third condition corresponded to the target word associated with a prime expressing a non-canonical or semantic association, which is the meaning transferred from the L2 French to the L1 Spanish. In this condition, '*exprimir'* is associated with the prime '*expresar*'. The meaning of '*s'exprimer'* in French (*to express*) is extended to the meaning of Spanish '*exprimir*' (*to press)*. This condition is called L2 meaning overlap because the target word '*exprimir*' (associated to the sense of *to express* of French) describes a semantic overlap with the prime '*expresar*' in Spanish (see **Figure 21**).

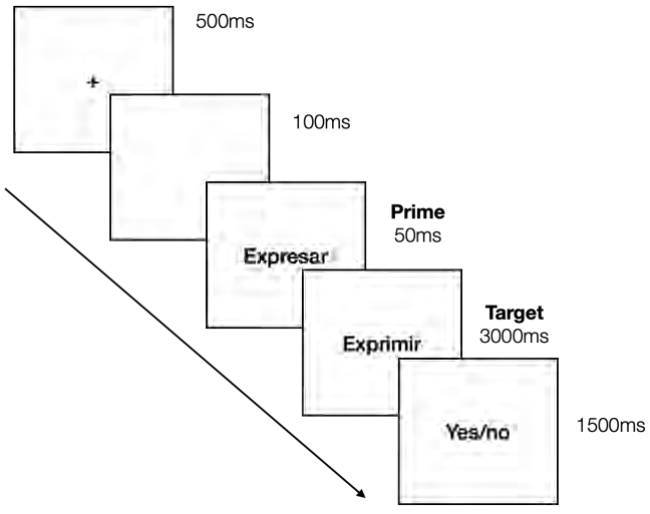

Figure 21 Condition 3: L2 meaning overlap

To summarize, the LDT enables us to determine for condition 1, whether primes associated with targets that reveal some form overlap produces a priming advantage to recognize words. For condition 2, whether primes associated with targets that overlaps with an L1 meaning facilitates recognizing words or whether it is the case for primes associated with targets that overlaps with L2 meaning for condition 3.

3.6.2.3. Time settings

All words, pseudo words and distractors involved a matched prime, one per condition. Only correct responses were analyzed. The time setting time of the presentation of targets and primes was 50 ms, primes were treated unconsciously by participants during this time setting (Sánchez-Casas & García-Albea, 2005). **Figure 22** describes the setting times in the LDT.

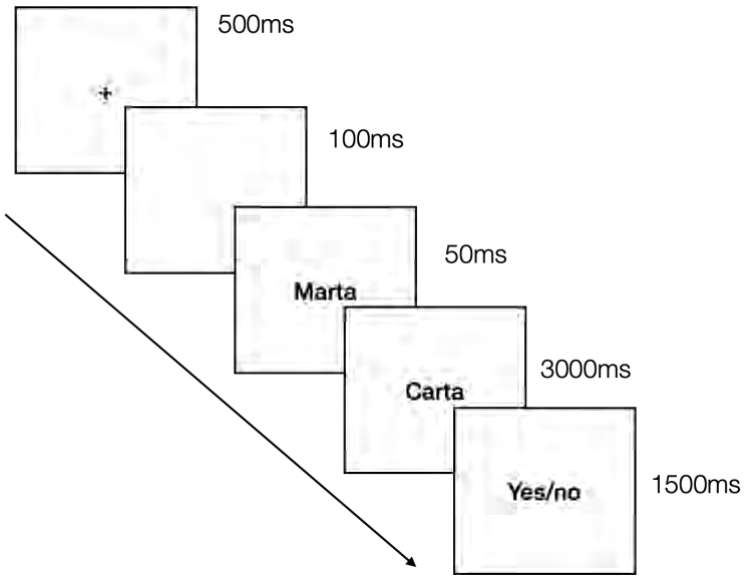

Figure 22 Time settings of experiment 2

3.6.2.4. Procedure During Experimentation

Participants were shown a brief instruction about the tasks, and then a practice trial was performant before the experimental task started. Both sections (training and test) were presented as described in **Figure 22**: A fixation cross was shown for 500 (ms) to catch the participants' attention, then, a blank screen was shown for 100 (ms), the prime appeared for 50 (ms), consecutively, the target was presented for 3000 (ms) and participants had 1500 (ms) to respond. Feedback was given as soon as the response was provided.

In order to articulate the LDT, targets and primes stimuli were presented along with pseudo-word targets and their respective prime. Participants were told to decide

whether the words presented on the screen exist or not in Spanish (L1). If a word existed in Spanish the participants had to type 6 in the keyboard (YES), if the words does not exist, participants had to type 7 (NO) on the keyboard. The right hand or left-hand were placed in equal distance from the 6 and 7 keys. The keyboard was covered in black and only the keys 7 and 6 were available for typing options. The feedback was either '*excellent*', '*try again'* and '*no answer was detected*'. Once the response was given using the participants dominant hand, it had to be replaced in the departure point to start all over again with the next word. A pause was placed before the next blocklist appeared.

3.6.3. Experiment 3 Acceptability Judgment Task (AJT)

The AJT is a suitable tool used in Linguistics to access both language use and frequency of use of certain structures. AJT are widely used in different areas such as L1 or L2 acquisition, communicative disorders and L1 attrition studies. AJT allow the research to assess structures that are infrequent in spontaneous speech and to assess grammatical as well as ungrammatical structures, including also structures that participants have tendency to avoid because of uncertain grammatical knowledge regarding those structures (Altenberg & Vago, 2004).

In this study, we consider that AJT acknowledges the awareness level of language correctness when native speakers judge sentences as acceptable or unacceptable. The sentences were judged using a scale from 1 to 7, in which the minimum value represents the least native like production and the maximum value corresponds to a fully accepted production. The use of a scale, rather than a binary response, enables us to use wider possibilities to assess the stimuli, since the sentences include subtle semantic differences depending on the context of the sentence. Multiple degrees of responses were in line with the subtle meanings presented in the stimuli.

3.6.3.1. Objectives

One of the main objectives of the AJT is to replicate the study by Grosjean & Py (1991) to compare it with more recent results. Thus, it could be possible that some of the L1 structures presented in the current study rebounded with Grosjean & Py's study. This would suggest that the L1 competence is certainly affected by L2 language use of immersed bilinguals. The second main objective of the AJT is to compare the levels of acceptance of the L1 dominant context group and the L2 dominant context group in the three types of sentence conditions.

The expected results are that the bilingual group would be more flexible about their acceptability level of the sentences than the monolingual group, especially in condition 1. In the same line, the monolingual group would be less flexible regarding condition 1 that refers to sentences in which a semantic extension is produced. Since

Condition 1 is a deviant production resulting from lexical transfer, the monolingual group is expected to disagree with the bilingual group, redrawing this condition in particular.

Additionally, for the bilingual group it would be interesting to observe the difference of acceptance levels within the three conditions. This comparison will allow us to relate the level of awareness of a deviated production in L1 with language use (condition 1), regular language use (condition 2) and the corrected version of a deviant production (condition 3).

Over all, the AJT is aimed at analyzing the different spheres of the participant performance as observed by Jarvis (2003) in a case study of a Finish woman immersed in the USA, in which the participant judged certain structures as incorrect, and nevertheless used these same structures in spontaneous speech. In this experiment, the conditions presented were supposed to oppose the normative use of the language with its actual use through a metalinguistic task, asking the participants to evaluate different propositions (conditions) involving deviated productions and the corrected version of these structures.

3.6.3.2. Linguistic Material

The linguistic material used in the AJT were sentences that were judged as acceptable or unacceptable on a scale from 1 to 6. The target words used to construct the sentences were based on a reduced selection of the stimuli of the LDT (see section 3.5.). The sentences were constructed by the researcher and a research trainee, they were pre-tested by two Spanish native speakers. The sentences were constructed considering the linguistic context in which the target word appeared.

The construction of the sentences was inspired from the work of Quilis et al (1982) who described an exhausted list of '*Semantic Gallicism*' that we have renamed sematic extensions. Authors described the linguistic context in which the sentences were observed in spontaneous discourses of bilingual children. (See the list of extensions used **Appendix 16**). This document was taken as a fundamental basis to construct the sentences that correspond to Condition 1, L1 deviated structures influenced by the L2.

Condition 2 was created in order to lead the bilingual group to realize a better and suitable lexical choice that would be "more" L1 native like, the standard meaning that is transferred in condition 1.

Condition 3 was constructed as a complementary category aimed at indicating a suitable context for the stimuli used in condition 1. In other words, this category corresponds to the meaning of the word that is mistakenly used in condition 1, in its canonical context. Keeping in mind the conditions briefly defined so far, we hypothesized that the sentences that express the semantic transfer (in condition 1) would be more accepted by Spanish speakers living in an L2 dominant context, than by those who lived in an L1 dominant context.

The three conditions were constructed jointly and were modified gradually to create a corpus according to the criteria established by each condition. Enough context

information was given to the reader to ensure optimal understanding of the objectives of this task.

In the AJT response times were not collected since the final objective was to gather an introspective and rational response regarding L1 use. The length of the words and the main structure of the sentences was equated by stimuli during the stage of sentence construction. The place of the target word in the sentence is always the same, it was presented in capital letters. The frequency of the words used was not equated because in this exercise the context was more important than word frequency since frequency was already equated during the construction of the stimuli in section database.

3.6.3.2.1. Conditions

As mentioned before, three types of sentences were constructed, each type of sentence corresponding to a condition. The focus of each sentence were the same target words as in the LDT.

The condition 1 expresses a semantic extension in L1, which is the meaning transferred or calqued from the L2 meaning.

' Jesús me ENERVA muy fácilmente con su manera tan prepotente de hablar'

'Jesús ENERVATES me very easily with his arrogant way to talk'

In this sentence, the target word '*enervar*' expresses a transferred meaning from the L2: '*to annoy'* or '*to irritate'* of the word in French *'s'énerver'*. In Spanish the canonical word that expresses this feeling is "*molestar*" ('*to bother'* or '*to disturb'*). Despite the fact that the verb '*enervarse*' actually exists in Spanish, the meaning expressed here is '*to exasperate'* (RAE, 2020) , indeed, the meaning does not match entirely with the sense of *s'énerver'* in French.

The condition 2 indicates the canonical meaning of the target word in L1:

'*Me MOLESTA mucho María con su falta de tolerancia hacia los demás*'

'*María ANNOYS me a lot with her lack of tolerance towards others*'

'*Molestar*' ('*to annoy'*) in this case is the most suitable lexical choice for the context of the sentence. '*Molestar*' is also more frequent than '*enervarse' .'Molestar*' would be the equivalent meaning of the French *'s'énerver'*.

The condition 3 reveals the L1 suitable meaning of the target word that was presented in Condition 1. In other words, it refers to the '*corrected*' meaning of '*enervarse*' in Spanish.

'*Me ENERVO al tomar este medicamento y pienso no tomármelo más*'

'*Taking this medicament ENERVATES me and I don't want to take it anymore*')

In this sentence the meaning of the target word '*enervarse*' matches with the context meaning expressed in the sentence, which is *to collapse the nerves* or *to weaken*, in opposition

to the semantic transfer (*'to annoy'*) coming from French *'s'énerver'* presented in condition 1, and considered as an uncanonical sentence.

In some cases, the target word used to express the semantic extension had two spheres (or more) of meaning, in some cases, the target words were presented twice, expressing this way different transferred meanings.

Three different lists were created in order to better randomize the condition sentence in E-Prime. Then a different random order was automatically generated by E-Prime per participant, following the order of stimuli presented in the lists in order to avoid repetition and keeping stimuli distant. The same randomizing procedure was applied as for the LDT.

3.6.3.3. Procedure During Experimentation

Two slides of instructions were shown to the participants **Appendix 14** followed by the criteria or principles of judgment expected by the researchers. Participants were told to read the sentences and to attribute to each sentence an evaluation as acceptable or unacceptable using a scale from 1 to 6.

The instruction was to evaluate the sentences taking into account the context and focusing particularly on the word presented in capital letters. In each sentence the target word was presented in capital letters to facilitate the participants to identity the focus of the sentence. Before attributing a response, it was recommended to respond the followings questions:

- '*Qué significado tiene dicha palabra en el contexto en el que sé encuentra ?' / 'What is the meaning of the word in that context* ?'
- *'Dicha palabra corresponde al contexto ou oración?'/'Does the word match in the sentence*?'
- '*¿Existe otra palabra que sea mas aceptada para expresar dicha idea, en lugar de la palabra presentada en mayúsculas? / 'Is there any other word more suitable to express that particular idea instead of the word presented in capital letters?*'

We asked the participants to exclude other criteria such as grammar or coherence. The main aspect to focus on was the meaning of the words and the linguistic context in which the target word was presented. The responses obtained in the ACT were not timed, the sentence remained on the screen as long as the participants wished before choosing a number representing their judgment. As soon as a sentence was evaluated, the sentence disappeared, and another sentence was displayed instead. The task was organized in three lists, and the participants could do as many pauses as they wanted. No feedback was given after evaluation.

3.6.3.3.1. Time settings

Instructions were presented on the computer screen, with the slides lasting as long as the participants wished. Before presenting the stimuli, a fixation cross appeared for 500

ms, after the 500 ms the sentence appeared. As mentioned before, the ACJ was non-chronometric, implying that once the sentence was shown the participant decided how long the sentences remained on the screen until it was judged by the participant. As soon as the participant typed a number from 1 to 7 on the keyboard, the next fixation cross appeared, as well as the next sentence. The list was organized in 66 samples by list (1 cycle x 66 samples/cycle). A pause was placed between the 2th and the 3th list (see the architecture of the experiment in **Appendix 15**).

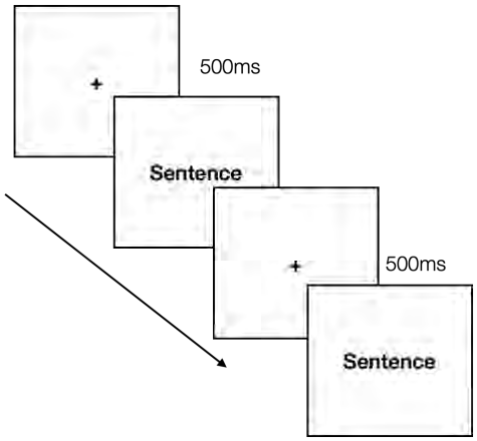

Figure 23 Time settings of experiment 3

3.6.4. Conclusion

This section was aimed to illustrate the experimental design, the stages of stimuli constitution, to present the linguistic resources used (Duchon et al., 2013; Marian et al., 2012) to construct the linguistic material and articulate the experiments. Additionally, this part was aimed at describing the characteristics of population, their linguistic profile, and educational background.

The experimental design was constructed following three main objectives: the first objective is related to the qualitative production of bilinguals; we have gathered controlled spontaneous oral productions in which semantic extensions might be potentially produced in some sentences. This is the rationale of the first experiment 1: the GCT.

The second experiment, the LDT, focused on the linguistic features of words regarding neighborhood density and morphology family size targets to compare processing and perception levels of bilinguals and monolinguals. The objective of the LDT is to determine whether form or meaning overlap plays a major influence during processing and word recognition. In the same line we wonder about how the L2 dominant morphology and neighborhood facilitates L1 lexical activation. On the opposite, whether L1 dominant morphology and neighborhood would have some resistance effects on semantic extensions. This is the rational of the LDT.

The third experiment, the AJT is aimed to observe the L1 comprehension and metalinguistic awareness levels. Particularly, to observe if bilinguals differ from bilinguals in their judgments, whether the bilingual's evaluations are more flexible than those of monolinguals. Differences across the condition sentences evaluated are expected to illustrate the degree of acceptance of semantic extension in its corrected version (condition 2) and regular language use of the is misleaded meaning presented in the semantic extension (condition 3).

An additional advantage of the current design is that our experiments go beyond the traditional bilinguals-monolingual comparison, we improve our approach complementing with comparisons in between tasks comparisons and studying the L1 without alternating languages. In the same line, we have tested the same stimuli observing different psycholinguistic levels in different experimental tasks. Indeed, we aimed to study semantic extension from different perspective.

4. **RESULTS AND DISCUSSION**

4.1. STATISTICAL APPROACH

We have adopted different statistical approaches depending on the type of data collected and the hypothesis exposed in section 3.1. In this section each statistical choice will be detailed separately since each corresponds to a particular experimental task. As described in the previous Chapter, the protocol comprised three main experiments:

- Experiment 1 is a *Gap Completion Task* whose objective was to collect L1 *Oral Productions* of controlled elicited language reduced to one or two-word responses. For the data treatment we used a Linear Mixed Effect Model to analyze RT (response time) using PRAAT. We then conducted a qualitative analysis in order to examine L1 response patterns in each group.
- Experiment 2 is a *Lexical Decision Task* whose objective was to observe word recognition processes depending on the nature of the stimuli presented. Again, we used a Linear Mixed Effect Model to analyze RT in three conditions that were experimentally manipulated and to compare the responses of the two groups of Hispanics.
- Experiment 3 is an *Acceptability Judgment Task* whose objective was to observe processes involving metalinguistic knowledge to evaluate language use and the degree of acceptability of non-canonical sentences. For the data analysis we used an Ordinal Logistic Regression Model to examine ordered responses on a scale from 1 to 7. The responses obtained in each group were compared.

As can be noted, globally, in each experiment we compare bilinguals and monolinguals in their L1. It should be stressed that all stimuli were presented in L1 Spanish, those that were considered as L2 (French) dominant in *Neighborhood Density* or *Size of the Morphological Family* were presented through their closest translation equivalent in L1 (Spanish). For example, the word '*Courage*' in French has more morphological family members (6) than '*Coraje*' (anger) in Spanish (2). Thus, in the category *L2 Dominant Morphological Family* it is represented by the translation equivalent (i.e., '*Coraje*') with few or dispersed morphological family members in L1 (Spanish).

In sum, Experiment 1 related to L1 production processes, Experiment 2 on word recognition processes, and Experiment 3 aimed at observing linguistic awareness and specificities of languages acceptance and use, all observed across the monolingual and bilingual groups.

Each of the three experiments involved different modalities of the linguistic material used. In Experiment 1, the expected target responses (produced or not by the participant) were the stimuli presented in Experiment 2 and 3. In Experiment 2, stimuli were presented as isolated words with three priming associations. In Experiment 3, these same stimuli were presented contextualized in a sentence.

Respecting methodological proposes, the order of experiment will be presented following the order of analysis: beginning with Experiment 2 and followed by Experiment 3, both experiments involved L1 and L2 Neighborhood Density Dominance and Morphological Family Size Dominance as variables. We will end by presenting Experiment 1 in which we have excluded those factors and included complementary qualitative analysis involving semantic extensions and CLI productions. It should be noted that the 7 Targets categories as described in section 3.5.4. were renamed into two main dominances Dom_N and Dom_M, this have facilitated the statistical analysis in R. This new tag allowed to equalize the number of our categories while at the same time opposing our two main factors (see new categories to tag stimuli as in the **Appendix 17**).

4.1.1. Linear Mixed Effects Models

The main characteristic of Linear Mixed Effects Models is that they take into consideration both fixed effects and random effects (Bates et al., 2015). Fixed effects correspond to the independent variables, and random effects represent an aleatory sample of a larger population or group of stimuli of study. Linear Mixed Model allows to take into account the variation of aleatory or random factors. (e.g., the population that is randomly sampled by a group representing the bilingual population, or a selection of stimuli representing a dominant L1 or L2 stimuli categories. Fixed effects or independent variables which are expected to produce an effect on the dependent variable, corresponds to the behavior measured in the task, that are RT values in experiment 1 and 2.

The main advantage of Linear Mixed Effects Models is the treatment of missing values or measurements (NA). In traditional statistical approaches (repeated measures and mixed ANOVAS), incomplete values have to be discarded excluding then, the variables within values or measurements are missing (e.g., a participant that forgets to respond) (Levshina, 2015). Further details involving the advantages of using Linear Mixed Effects Models are going to be presented in section 4.2.2.

4.1.2. Ordinal Logistic Regression Model

Globally, Ordinal Logistic Regression Models are Ordinal Fits Cumulative Link Mixed Models with one or more random effects via the Laplace approximation or quadrature methods (Christensen, 2019).

The functioning of Linear Mixed Effects Models is very similar to the Ordinal Logistic Regression Models, the main difference lies in the nature of the data and the modeling analysis. These models are used for ordinal/ordered data, and the modeling analysis is focused on proportional odds or odds-ratio, which measures the possibilities of recurrence of a variable. Otherwise, odd ration as explained by Szumilas (2010, p. 1) as "*The odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure.*"

Further details concerning the implementation of the Ordinal Logistic Regression Model in experiment 3 are going to the presented in section 4.3.1.

4.1.3. Qualitative Analysis

Qualitative data were used as complement in experiment 1(Gap Completion Task) to try to explain the production of semantic extensions in some cases. The results of the oral production were analyzed in two steps. First, we calculated the response rates in different categories. Second, we analyzed the productions of both groups in terms of the most currently used words as responses for each stimulus. Additionally, in experiment 1, extralinguistic data was included into the statistical analysis to explain the production of Semantic Extension and CLI by the bilingual group.

4.2. LEXICAL DECISION TASK

The objective of the Lexical Decision Task was to differentiate two main levels influencing levels of processing of semantic extensions: word-form overlap between languages (through Dominant Neighborhood density stimuli) and semantic overlap between languages (through Dominant Morphology Size stimuli).

Figure 24 summarizes the experimental design of the task describing the groups studied, the stimuli and the conditions that were presented. The task was performed in the L1 (Spanish) of participants and completed by two groups: bilinguals and monolinguals. Participants categorized words and pseudo-words through a Lexical Decision Task. The stimuli presented were either dominant in the L1, in L2 or neutral for the Morphological Family and the Neighborhood Density. They were presented in three conditions: SP, FR and None.

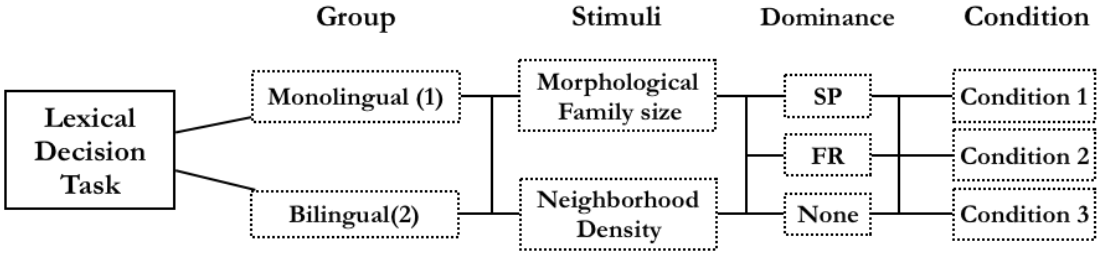

Figure 24 Experimental design of the Lexical Decision Task

In this section, we will describe data analysis, including the procedure of RT data experiment cleansing of RT, and the statistical procedure that was conducted. Final results will be reported and interpreted jointly for a better understanding of the statistical output obtained.

4.2.1. Data Cleansing

Before analyzing the data collected that resulted from the LDT, it was necessary to carry out an experiment cleansing in order to avoid biased data and to facilitate analysis. Experiment cleansing was conducted between stimulus (including distractor stimuli responses, NA, incorrect responses), between conditions and between groups. Data Cleaning stages were the following:

For the bilingual group (40 participants), 504 trials per participant were tested, resulting in a total of 20160 responses[29], half of which were words in Spanish (10080), and the other half non-words.

- In a first stage, we suppressed the 2520 responses of distractors stimuli.
- Second, from the 8040 remaining responses, the data from 28 responses were suppressed and considered time-outs because participants failed to answer within 5000ms.
- Third, we removed 206 incorrect responses that participant categorized as words when stimuli were actually non-words, or when pseudo-words were considered as words. Experiment cleansing resulted in 7326 responses.
- The final stage of data cleansing included the suppression of outlier RT that were too long or too short depending on the condition presented. Table 12 describes the surpassing threshold used for data suppression by condition. We suppressed 901 responses to the remaining data (7326 responses) giving as final data 6425 responses. The global percentage of suppressed data during all stages described so far is 31,87% of the total data collected.

Table 12 Data Cleansing by condition for the bilingual group

	Condition1	Condition2	Condition 3
RTmean+2,5*SD	1520,61	1521,87	1532,00
RTmean- 2,5*SD	584,874	610,32	587,36

For the monolingual group we used the same procedure.

The monolinguals group (53 participants) 504 trails by participant were tested, resulting in a total of 26712 responses; half of them concerned words (13356) and the other half non-words.

- First, we removed 3339 responses that corresponded to the stimuli distractors responses.

[29] By reponse we refer to the entry provided by the participant which indicates wether the stimuli was considered as an L1 word (6) or a non-word (7).

- Second, we suppressed 75 NA responses considered time-outs when participants failed to answer before 5000ms.
- Third, we removed 520 incorrect responses, which were words that participants had mistakenly considered as non-words, or as words. Experiment cleansing resulted in 9422 RTs.
- The final stage of data cleansing included the suppression of responses that had surpassed the threshold of two and a half standard deviation from the average per condition (described in Table 13 suppressing 1412 responses to the remaining data (9422 responses) giving as final data 8010 responses. The global percentage of suppressed data during all stages described for group 2 is 29,99% of the total data collected.

Table 13 Data Cleansing by condition for the monolingual group

	Condition1	Condition2	Condition 3	Condition1
RTmean+2,5*SD	1475,39	1510,16	1493,19	1475,39
RTmean+2,5*SD	548,62	529,08	533,43	548,62

In this experiment data cleansing procedure was conducted by condition and by group to take into consideration the differences between the participants, and to adapt the RT for each response following the characteristic of the conditions presented. On one hand, conditions had different nature, for example the RT invested to recognize word-form related stimuli would take shorter RT than condition 2 (words that are semantically related) and 3 (cross-linguistic words that are semantically related). On the other hand, the fact that bilinguals usually take longer to respond than control monolingual groups (in naming picture tasks e.g., Ivanova & Costa, 2008), is a strong argument for such data cleansing procedure (by group and by conditions). Data was collected through E-prime (Psychology Software Tools, Inc, 2016) and data cleansing and analysis were made using the software R (R Core Team, 2017) which is an opensource language for statistical computing.

4.2.2. Statistical procedure

In the Lexical Decision Task or LDT, one of the main objectives was to test whether possible variation in the reaction time responses[30] of participants is related to multiple explanatory variables or independent variables. In the experimental design, the independent variables were: the group, the type of stimulus and the conditions presented in the experimental task. The dependent variable is the RT response of the participants.

[30] coded as **Targets.RT**

There were four types of independent variables, which we will call *factors* because each of these is divided into different levels. The explanatory variables are:

1) **Group**: divided in two levels: Bilingual-Monolingual.
2) **Condition**: divided in three levels:
   Condition 1 indicated word-form overlap between primes and targets words.
   Condition 2 indicated canonical meaning association between primes and targets words.
   Condition 3 indicated a non-canonical association between primes and targets words.

Additionally, two independent variables concerned the types of stimuli presented in this experiment:

1. **Dominant N** for Neighborhood Density of words and
2. **Dominant M** for Morphology Size of words.

Factors 3 and 4 involved multiple combinations corresponding to the language in which the dominance is manifested in French (Fr) or in Spanish (Sp), this division enables to investigate independently the factor dominance in relation to the language concerned.

The statistical tool used to define the influence of these factors on the RTs is a Linear Mixed Effects Models, a regression analysis performed with R (R Core Team, 2017) which enables us to measure the impact of multiple factors while controlling other variables. As it was mentioned in section4.1, Linear Mixed Effects Models Regression takes into consideration both fixed effects and random effects (Bates et al., 2015).In in the Lexical Decision Task the aleatory factors or random effects were the group studied and the targets presented.

The advantage of implementing a Linear Mixed Effects Models (LMEM) is that contrary to previous processing methods (hierarchical or multileveled models) that assumed random effects as nested, Linear Mixed Effects Models allow to include subjects and items as independent and crossed effects. Among other improvements of LMEM, we find the dealing of missing data, to analyze repeated measures and its previous insufficiencies related to statistical power (Baayen et al., 2008). Besides, LMEN also facilitates the creation of different models which include multiple explanatory variables to explore possible interactions between them and the dependent variable or other(s) explanatory variable(s). (Levshina, 2015). In this experiment, we proceeded with the *Lme4* package (Bates et al., 2019) and *lmerTest* (Kuznetsova et al., 2019) to obtain p-values.

4.2.3. Results of the Linear Mixed Effects Model

The Linear Mixed Effects Model was constructed starting with a basis model, in which we have included all possible predictors and interactions between variables; this allowed the experimenters to explore wider possibilities to analyze Lexical Decision Task's results. This model was reduced progressively by suppressing (one by one) the

factors that were not statistically significant, beginning from the most complex interactions to main effects.

The simplification of the model has resulted into a more specific and targeted model that would explain successfully the variability of responses (reaction time) for each group of participants (in the three conditions and depending on the stimuli dominance). The explanatory factors including this mode were:

- Group: labelled in two levels bilinguals (Group 2) and monolinguals (Group 1).
- Conditions: labelled in three levels (cond1, cond2, cond3).
- Dom_N: labelled in three levels: French (Dom_NFR) Spanish (Dom_NSP) and (Dom_NN) None).
- Dom_M: labelled in three levels: French (Dom_MFR), Spanish (Dom_MSP), and None (Dom_MNN).

We have assessed the normality of residuals of the final model visually residuals were spread across the fitted values and graphically looked like a blob suggesting that the model distribution data follow normality (Winter, 2019).

Table 14 describes the models that we compared in between using ANOVAS. The final model that we have chosen to explain the different factors that influenced RT's between groups is model 5.

Table 14 Summary of the analysis of comparison between models

Factors	X^2	df	p- values	AIC
MS5 without Target as random effect				194064
MT5 without Subject as random effect	0.0000	0	1.0000	197009
M5: Group*Condition + Group*Dom_M + Group*Dom_N + Condition*Dom_M + Condition*Dom_N + Group*Condition*Dom_N	3943.8922	1	0,000	193067
M4 Interaction suppressed: Dom_M*Dom_N	0.5121	4	0.9723	193074
M3 Interaction suppressed: Condition*Dom_M*Dom_N	4.5755	2	0.1015	193074
M2 Interaction suppressed: Group*Dom_M*Dom_N	2.0351	4	0.7293	193080
M1 Interaction suppressed: Group * Condition * Dom_M * Dom_N	0.9565	2	0.6199	193083
MC Complete model with all possible interactions	1.8014	4	0.7722	193089

As it can be seen in **Table 14**, model 5 is significantly accurate than the other models used; having the lowest AIC (193067), it is also significant different from the previous all models ($X^2(1)= <3943.8922$, $p <0,000$).

The Akaike Information Criterium (Akaike, 1974) is used for selecting models. It allows to estimate the likelihood between models while controlling the number of recorded measurements and the number of estimated parameters, a weak AIC indicated that the model fitted the best to explained the phenomenon of study (Mohammed et al., 2015).

In Model 5 in we have excluded all interactions that were not significant in previous models as detailed in **Table 14**. The specificities of model 5 are described in **Table 15**. We have also controlled for random effects by creating two supplementary models, MS5 in which we have suppressed subject as random effect, and MT5, in which we have suppressed Target as random effects. Results suggest that the model 5 predictions are independent for target or subject effects only.

Table 15 Model 5- Final Linear Mixed Effects Model

factors	β	Std. Error	df	t value	Pr(>|t|)
Intercept	977.334	23.250	162.000	42.036	< 2e-16 ***
Group2	13.439	26.292	113.000	0.511	0.610253
Conditioncond1	-3.026	9.291	14256.000	-0.326	0.744703
Conditioncond2	4.596	9.210	14257.000	0.499	0.617774
Dom_MFR	-22.978	20.301	70.000	-1.132	0.261525
Dom_MSP	-19.992	23.078	70.000	-0.866	0.389277
Dom_NFR	-63.847	20.879	77.000	-3.058	0.003059**
Dom_NSP	-26.961	24.056	78.000	-1.121	0.265814
Group2:Dom_MSP	26.724	10.669	14260.000	3.041	0.002365**
Group2:Dom_MFR	27.419	7.732	14257.000	3.546	0.000392 ***
Group2:Dom_NSP	38.039	14.908	14259.000	2.552	0.010735 *
Group2:Dom_NFR	53.651	12.871	14254.000	4.168	3.09e-05 ***
Conditioncond1:Dom_MFR	-27.982	9.432	14256.000	-2.967	0.003014**
Group2:Conditioncond2	14.768	11.341	14254.000	1.302	0.192864
Group2:Dom_MFR	27.419	7.732	14257.000	3.546	0.000392 ***
Group2:Dom_MSP	26.724	8.789	14260.000	3.041	0.002365 **
Group2:Dom_NFR	53.651	12.871	14254.000	4.168	3.09e-05 ***
Group2:Dom_NSP	38.039	14.908	14259.000	2.552	0.010735 *
Conditioncond1:Dom_MFR	-27.982	9.432	14256.000	-2.967	0.003014 **
Conditioncond2:Dom_MFR	-10.309	9.365	14256.000	-1.101	0.271033
Conditioncond1:Dom_MSP	7.831	10.682	14257.000	0.733	0.463493
Conditioncond2:Dom_MSP	-26.027	10.669	14257.000	-2.439	0.014721 *
Conditioncond1:Dom_NFR	25.627	12.355	14255.000	2.074	0.038073 *
Conditioncond2:Dom_NFR	15.008	12.218	14255.000	1.228	0.219322
Conditioncond1:Dom_NSP	7.939	14.450	14257.000	0.549	0.582722
Conditioncond2:Dom_NSP	7.274	14.379	14259.000	0.506	0.612964
Group2:Conditioncond1:Dom_NFR	-41.092	17.877	14253.000	-2.299	0.021541 *
Group2:Conditioncond2:Dom_NFR	-22.354	17.797	14254.000	-1.256	0.209120
Group2:Conditioncond1:Dom_NSP	-13.574	20.842	14254.000	-0.651	0.514873
Group2:Conditioncond2:Dom_NSP	-11.909	20.693	14254.000	-0.576	0.564953

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

4.2.3.1. Main Effects

Data analysis have shown a significant main effect of Dominant Neighborhood when it is higher in French (Dom_NFR) in merged conditions and groups, β= -63.847, SE=20.879, t=-3.058, p=0.003059 (see **Table 15**) This result suggests a facilitation effect

in comparison to other factors such as: dominance neighborhood in Spanish, or the dominance of the Family Size of either French or Spanish. Interaction effects would allow a better understanding of this particular output (see 4.2.3.4.4).

We had predicted differences in between groups during stimuli processing. However, the model did not report for such as main effects: β= -13.439, SE=26.292, t=-0.511, p= 0.610253 for group factor, suggesting that there was not a significant difference between group and their responses (all conditions mingled).

Nevertheless, the model has reported for multiple interactions effects, they involved specific conditions and depended of the type of stimulus presented (dominant in L1 or in L2) detailed in section 4.2.3.3 and 4.2.3.4.

4.2.3.2. Interaction

We have found significant interactions on the variation of the reaction time values of the bilingual group depending on the type of stimuli presented (Dominant Neighborhood and Dominant Morphological Family size), also between the condition presented and the type of stimuli presented. Furthermore, we have found a three-way interaction between: group, condition and neighborhood density of word. We will describe and interpret separately each type of interaction in this section.

4.2.3.3. The Effects of Neighborhood Density and Morphological Family Size in Bilinguals

As we have detailed in (section 3.6.2) we have manipulated the stimuli presented in two categories: Neighborhood Density and Morphological Family Size of words. Each category is opposed as dominant in French or in Spanish. A neutral condition is used as control (in which the words were equivalent for both languages in terms of neighborhood and morphological measures).

As it can be seen in Table 16 bilinguals are slower than monolinguals in both dominances: for French and Spanish in the two categories: Neighborhood Density and Morphological family size.

Table 16 Bilinguals group estimates responses in comparison to monolinguals group for Dom_N and Dom_M (extracted from Model 5)

Intercept	Estimate	Std. Error	df	t value	Pr(>|t|)
Group2:Dom_MSP	26.724	10.669	14260.000	3.041	0.002365**
Group2:Dom_MFR	27.419	7.732	14257.000	3.546	0.000392 ***
Group2:Dom_NSP	38.039	14.908	14259.000	2.552	0.010735*
Group2:Dom_NFR	53.651	12.871	14254.000	4.168	3.09e-05***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

4.2.3.3.1. Dominant L1

Bilinguals are 26.72 milliseconds slower than monolinguals, for Morphological words whose Family Size is higher in Spanish, this difference is significant β = 26.724, SE=10.669, t=-3.041, p<0.002365. In the same language but for the category of high neighborhood density, bilinguals' response rates are slower than monolinguals by 38.039 milliseconds, giving as result: β= 38.039, SE=14.908, t=2.552, p= 0.010735.

4.2.3.3.2. Dominant L2

The effect size for dominant L2 categories (Morphological Family Size, and Neighborhood Density in French) is more significant than for the L1 dominant categories (see Table 15). This suggests that the bilingual group would be sensible to dominant L2 categories in comparison to the L1 non-immersed group.

Regarding the RT values of bilinguals for L2 dominant Morphological Family Size (whose family size is higher in French), reaction responses are slower than monolinguals by 27.419 milliseconds: β= 27.419, SE=7.732, t=3.546, p<0.000392.

If we compare this output with their counterpart in Spanish the processing of L2 Dominant Morphological Family Size by bilinguals is more delayed than the L1, as it was shown by the effect size. Hence, the spreading activation level of the Morphological Family Size delayed responses in both languages for all conditions merged. Our interpretation for this outcome is going to be discussed in section 4.2.3.5.

Regarding the RT values of bilinguals for L2 dominant-Neighborhood Density (whose neighbors are higher in French), reaction responses are slower than monolinguals by 53.651 milliseconds: β = 53.651, SE=12.871, t=4.168, p<0.0000309 [31]. As reported previously, the processing of L2 Dominant Neighborhood Density was more delayed for the L2 than the L1. This output suggests that the processing of high L2 Dominant Neighborhood Density words is more effortful for bilinguals than for monolinguals and, than the Morphological Family Size category. Here, French activation levels would interfere during the Lexical Decision Task, in terms of formal features, even thought that task was performed in Spanish see section 4.2.3.5 for a detailed discussion on this output.

4.2.3.4. Interaction between Dominance and Conditions

Dominance in L1 or in L2 of Morphological Family Size and Neighborhood Density was predicted in relation to the factor condition which is divided in three levels: cond1, cond2 and cond3.

1. Condition 1 refers to form overlap between primes and prime. e.g., '*exprimir*' with the prime '*reprimir*' (to squeeze- to suppress).

[31] Equivalent de 3.09e-5.

2. Condition 2 involved a canonical meaning association e.g., '*exprimir*' with prime '*apretar*' (to squeeze, to tighten).
3. Condition 3 indicates a non-canonical association, as the result of L2 influence on the L1, e.g., '*exprimir*' associated with the prime '*expresar*' (to squeeze, to express).

Here, the semantic extension in condition 3 implies that the meaning of '*exprimir*' in Spanish is extended to the meaning of '*s'exprimer*' in French ('*to express)*. This non-canonical meaning involves misleading meaning in the L1 such as '*no logro exprimirme correctamente*' ('*I cannot **squeeze myself properly*') instead of the canonical form '*no logro expresarme correctamente*' ('*I cannot express myself properly*').

Our general hypothesis concerning the LDT is that bilinguals would process semantic extensions differently because of the accumulated sources of activation in the L1 and the L2. We predicted that for bilinguals, L2 dominant categories would produce or trigger faster latencies rates, especially for stimuli presented during condition 3. However, none of such effects were observed in the data analyzed. Interaction of the variation of latencies rates between condition 3 (semantic extension associations) and the bilingual group are not significantly different: β= -14.7683, SE=11.3409, t=-1.302, p= 0.192864 in **Appendix 18**. Nevertheless, results suggest that L1 dominant categories play a particular role in the variation of the latencies rates of both groups of participants as it is detailed in section 4.2.3.4.1.

4.2.3.4.1. The Role of L1 dominant Morphological Family Size in Condition 2

Data have pointed out an interaction between the Morphological Family Size dominant in Spanish for all groups merged showing a significant interaction for condition two in which a facilitation of -26.027 ms was observed: β= -26.027, SE=10.669, t=-2.439, p<0.014721 (see **Table 15**). This interaction suggests that there is a facilitation of meaning overlap during word recognition processes. Condition 2 involves L1 canonical meaning associations when primes and targets are semantically related in L1, respecting its canonical or standard relationship (e.g., the target words '*exprimir*' associated with the prime '*apretar*' both synonyms of squeeze). This facilitation affects equally both groups: bilinguals and monolinguals (see **Figure 25**), here Spanish is the L1 of both groups.

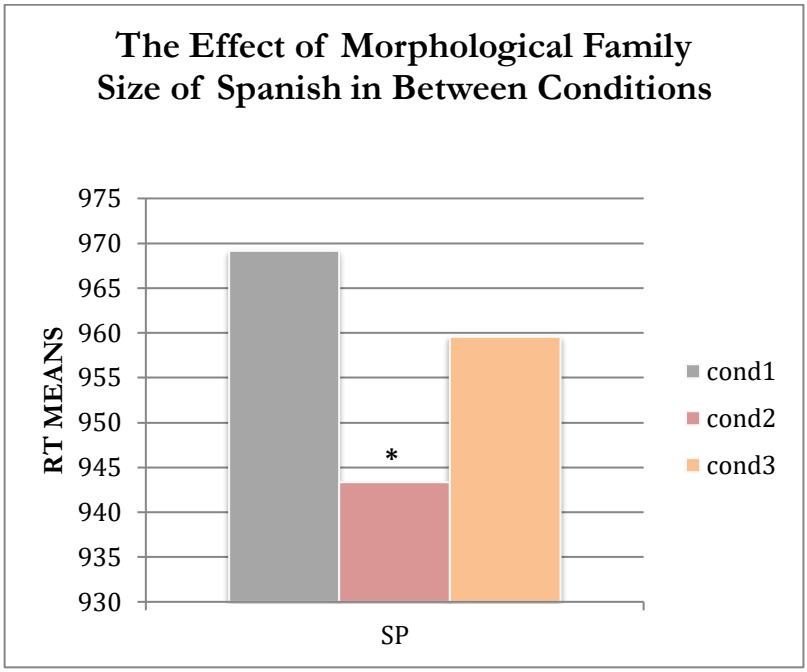

Figure 25 Effect of L1 Dominant Morphological Family size in between conditions

4.2.3.4.2. The Role of Word Form for L2 Dominant Morphological Family Size in Condition 1

This interaction suggests that there is a facilitatory effect when the Morphological Family Size (presented in the L1) is dominant in French for all groups mingled in condition 1, which stands for word-form overlaps (e.g. the target '*exprimir*' - associated with the prime '*reprimir*') β= -27.982 , SE=9.432, t=-2.967, p<0.003014 (see Table 15) This output suggests that word-form overlap plays a role in priming independently of the dominance (L1-L2) presented, here participants are sensitive to the form overlap. The fact that L2 dominant Morphological Family Size in condition 2 is not significant (β= -10.309, SE=9.365, t=-1.101; p=0.271033, in Table 15) for either group suggests semantic representations in the L2 would play a less important role in morphological processing than word-form.

4.2.3.4.3. The Role of the Neighborhood Density Between Conditions

Neighborhood Density dominant in L1 produced no significant effect in neither of the conditions presented. β= 7.939, SE=14.450, t=0.549; p=0.582722 for condition 1; β= 7.274, SE=14.379, t=0.506, p=0.612964 for condition 2 and β= -7.2739, SE=14.3794, t=-0.506; p=0.612964 for condition 3 (**Appendix 18**).

Neighborhood Density dominant in L2 produced a significant effect exclusively in condition 1 β= 25.627, SE=12.355 t=2.074, p<0.038073. This output indicates that

latencies rates are longer for 25.627 milliseconds all groups merged when Neighborhood Density is dominant in L2. No significant effects were found for condition 2 (β= 15.008, SE=12.218t=1.228, p=0.219322) (see **Table 15**) nor for condition 3 (β= -15.008, SE=12.218t=-1.228, p=0.219322) (**Appendix 18**).

4.2.3.4.4. The L2 Faciliatory Effect of Neighborhood Density on Word-Form in the Bilingual Group

A three-level interaction in the bilingual group is observed in condition 1 when the Neighborhood Density is dominant in the L2. In this word-form overlapping condition, facilitatory effects are observed during word recognition for -41.092 milliseconds: β= -41.092, SE=17.877t=-2.299, p<0.021541 as it is shown in **Figure 26.**

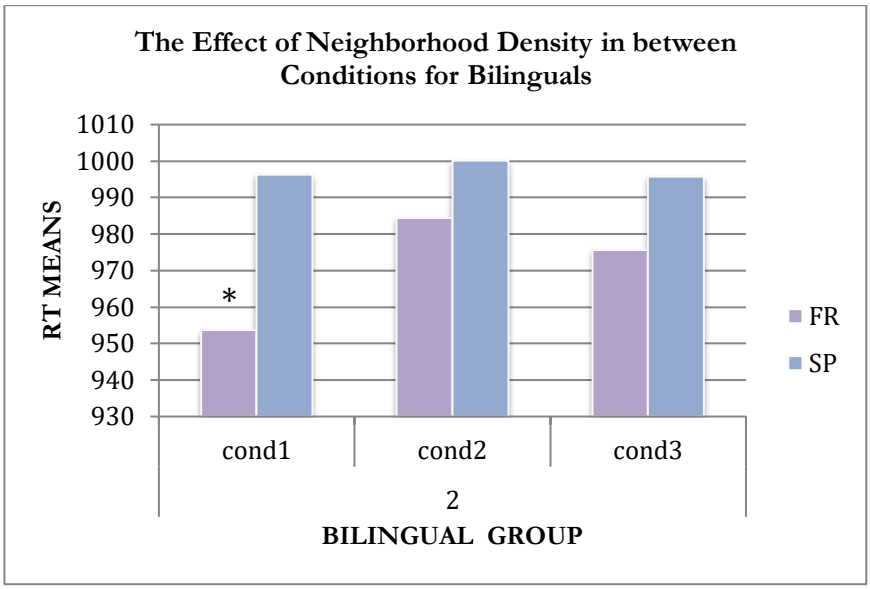

Figure 26 Effect of L2 dominance Neighborhood Density in between conditions for Bilinguals

4.2.3.5. Discussion

The initial objective of the Lexical Decision task was to oppose two interfaces: Neighborhood Density of Words and Morphological Family Size between languages. The main focus was to explain Semantic Extensions as the result of the word-form overlap between languages (using Neighborhood Density Measures) or as the result of meaning overlap between languages (using the Morphological Family Size as measure).

As shown in section 4.2.3.3.1 the effects of L1 Dominant Neighborhood Density and L1 Morphological Family Size in bilinguals produced slowed down reaction responses in comparison to the monolingual group. Bilinguals behave distinctly different from monolinguals in the way they process L1 dominant stimulus (Dom_N and Dom_M)

which may support the concept of multicompetence (Cook, 2003, 2013) see section 2.1.1.). Multicompetence suggests that L2 users differ in the way they use and processes language, this refers not exclusively the L2 but also the L1. Both language use and processing would be distinct for bilinguals L1's and L2's from the way monolinguals use and process language.

If we focus our attention to the effects of L2 Dominant Neighborhood Density and L2 Morphological Family Size in bilinguals RTs are more affected (displaying significant delayed responses) by the Dominant categories in L2 than in the L1 for all conditions merged (see 4.2.3.3.2). Regarding L2 Dominant Morphological Family Size, the absence of facilitation (longer RTs) of the Morphological Family Size effect task may be due to the activation of semantic representation simultaneously with the lexical forms of stimuli presented across languages. A similar effect has been observed in the past (Grainger & Segui, 1990; Schreuder & Baayen, 1997) when progressive demasking tasks[32] were implemented, which allowed to include more 'central' levels of lexical processing involving semantic representations, slowing down latency rates. In our study, primes were not masked, the condition presented (condition 2 and condition 3) involved semantically related words in the L1 and the L2 translation equivalent presented in L1 (e.g., '*carta*' was presented with '*tarjeta*').

Regarding L2 Dominant Neighborhood Density the effect size effect (delayed RT) was more important than L1 Dominant Morphological Family Size for the bilingual group (and in comparison, to the monolingual group). This output suggests that the processing of high neighborhood density words (dominant in French) is more effortful for bilinguals than for monolinguals and, than the other categories. Here, French activation levels would interfere during the Lexical Decision Task, even thought that task was performed in Spanish. We recall that we have presented primes in Spanish, as translation equivalents. For example, when a word has high Neighborhood Density in French such as '*habiter*', the word prime presented was the matching interlingual word in Spanish: '*habitar'* which has by default a disperse Neighborhood Density in Spanish (and few Family Size in French or in Spanish) and a dense Neighborhood in French.

This sensibility of dense Neighborhood Density in French presented in its L1 equivalent goes in line with Thierry & Wu (2007) study in which event-related brain potentials have revealed implicit access to the L1 when bilinguals read words in their L2 (that presented formal links with the L1) suggesting that lexical analysis in one language task not exclude analysis on the L2 or any other language composing the mental lexicon of the bilingual. Thierry & Sanoudaki (2012) suggest that when the bilingual accessed lexical information in any particular language, linguistic features (lexical, phonological, syntactic semantic) that are related in other(s) language(s) can be activated. This co-activation of languages will affect language performance, facilitating or affecting

[32] In a desmaking task stimuli are demasked progressively until they are entirely aparent to the participant during experimental trials.

negatively lexical access. However, this would be dependent on the nature of the psycholinguistic task that is involved (Dijkstra et al., 1999).

Global results suggest that bilinguals are slower than monolinguals, which may be due to multiple linguistic resources available for the bilingual providing from more than one language. Therefore, bilinguals are *"less performant"* in linguistic tasks because of their use of two linguistic systems that will have an effect on the way they process information. In terms of Gollan et al., (2008) because of the weaker conceptual-lexical connections of L2, it makes bilinguals slower than monolinguals, because of a less frequent L2 use in comparison to a L2 monolingual speaker or an L1 monolingual speaker.

The significant facilitatory effect produced by L1 dominant Morphological Family Size and the RTs obtained in Condition 2 (e.g., '*expimir*' associated with '*apretar'*, both synonyms of squeeze) in both groups (see 4.2.3.4.1 tells us about the role of semantic links during L1 recognition. We interpret that this facilitatory effect is reinforced by L1 strong existing lexical networks between lemmas and concepts for both groups. The fact that a significance difference in Morphological Family Size dominant in Spanish was observed in condition 2, suggests that semantic links in the L1 are strongly rooted and that semantic representations have been reinforced over time, since Spanish is the language that represents important periods of time life such as childhood, primary and secondly schooling. Following the concept of entrenchment of the L1 (Steinkrauss & Schmid, 2017) described 4.2.3.5 in linguistic representations will be strengthened due to language use, in this particular case, language entrenchment (due L1 language use over the years) would affect the intensity of language processing, which is reliable to the facilitation effects observed in condition 2.

Finally, our three-way interaction involving the Bilinguals group at processing L2 Dominant Neighborhood Density in condition 1, where primes and targets are related in their formal features (e.g., '*exprimir*' associated with '*reprimir'*) suggest that that phonological and orthographical features of the L2 were highly activated whilst the experimental task is performed in the L1. In this case, L2 levels of activation of multiples neighbors will play a role during early processing of the word in L1. This sensitivity of word-form for a highly activated L2 network is absent in the monolingual sample studied.

If we consider the nature of the word recognition tasks, LDT will involve bottom-up parallel activation of word-form between languages including phonological and lexical features (Kroll et al., 2010). In this experiment, this effect is reinforced by a dominant neighborhood density in the L2, shared lexical features create a '*resonance*' in the activation levels across languages (Kroll et al., 2010).

An alternative explanation to this sensibility to word-form is the context of language use, which is dominant in French. We recall that recency of language use may trigger language activation levels (Kroll & Tokowicz, 2005; Paradis, 1997). Indeed, French corresponds to the language used the most by bilinguals in a daily basis, as it was indicated by the sociolinguistic survey the mean of French use corresponded to 70,13% of a regular

day. The frequent use of L2 could explain the particular sensibility to lexical word-form featured by the bilingual group, we assume that the L2 forms are highly activated because the stimuli presented (Dominant N) coactivates more easily words in the L2 than in the L1.

To conclude, the facilitatory effect observed of L2 Neighborhood Density on word-form by bilinguals suggest that Semantic Extensions do involve formal features and activation levels between languages, thus, facilitating latencies during a task conducted in L1. More importantly, semantic extensions are not exclusively the result of crosslinguistic learned associations, as suggested by Jarvis (2009). The role of the Morphological Family Size in the L1 indicates that meaning mapping and networks with lemmas is reinforced over the years for both groups and facilitating this way L1 word recognition.

4.2.3.6. Further analysis

Future applications of this research can be orientated in foreign language teaching approaches. First, second language teaching of vocabulary could be orientated taking into consideration the Morphological Family Size of words in order to reinforce concept mapping and networks construction between lemmas in the L2. Second, the implementation of teaching techniques giving special attention to L2 word-form related to the L1 could facilitate learning processes and prevent from semantic transfer.

4.3. ACCEPTABILITY JUDGMENT TASK

The Acceptability Judgment Task (AJT) consists in an evaluation of three types of sentences expressing 1) semantic extension 2) canonical meaning and 3) corrected meaning presented in (1). The evaluation of sentences was made using a scale from 1 (unacceptable) to 7 (acceptable) that corresponds to levels of acceptability rates.

The sentences included target words based on the same stimuli as those presented in the LDT (see description 3.6.2.2). As it is showed in **Figure 27**, in the Acceptability Judgment Task, the target words (presented in the form sentences) had two main dominances: Neighbor Density (Dom_N) in either French or Spanish and Morphological Family Size (Dom_M) in either French or Spanish. In the AJT, target words were integrated in different types of sentences corresponding to 3 conditions.

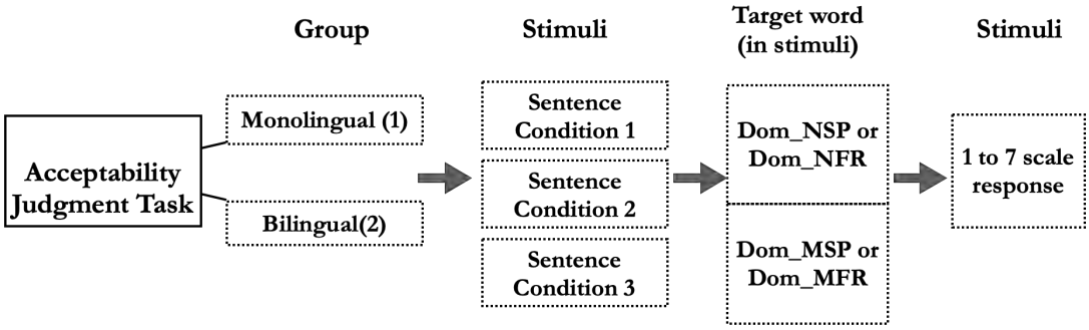

Figure 27 Experimental design of the Acceptability Judgement Task

- Condition 1: the sentence expresses semantic extensions from L2 to L1 i.e., the meaning of a given word in L2 is transferred to a word form in L1 (e.g., *'me ENERVA muy fácilmente Jesús con su manera tan prepotente de hablar'/ Jesús ENERVATES me very easily with his arrogant way to talk*)
- Condition 2: the sentence expresses the canonical meaning of the target word in L1 (e.g., '*me MOLESTA mucho María con su falta de tolerancia hacia los demás*'/ María ANNOYS me a lot with her lack of tolerance towards others)
- Condition 3: the sentence expresses the correct form and meaning of stimuli presented in condition 1, in this case, the corrected version of '*enervar*' in L1 (e.g., '*Me ENERVO al tomar éste medicamento y pienso no tomármelo más*'/ '*Taking this medicament ENERVATES me and I don't want to take it anymore*')

Special attention was given to the linguistic context in which target words were presented. The objective was to constrain meaning depending on the sentence condition. In opposition to the LDT in which word recognition concerned isolated words (non-contextualized), in the AJT, context was detailed through the sentence and target words were explicitly presented in capital letters.

The meaning of ambiguous words that are homographs is supposed to be activated during initial stage processing and there is not until later stages in which inappropriate meanings can be inhibited. This suggests that the inhibition of multiple meanings is made using the context that will help to constrain meaning (Lalor & Kirsner, 2001). This explains why this task is not a chronometric task as was the case for the Gap Completion Task and the LDT.

4.3.1. Statistical Procedure

The design of the AJT involves multiple explicative variables: group, condition and dominance. **Group** is divided into two levels (**Bilinguals and Monolinguals**) and **Sentence Condition** in three levels as described in section 3.6.3.2.1. Dominance included two variables: **Dom_N** for Neighborhood Density of words and **Dom_M** for Morphology Size of words whose levels are subcategorized in SP and FR (see **Figure 27**).

The nature of our dependent variable is ordinal, ordered, categorical and also multi-leveled. Taking into consideration the characteristics of our variables, the statistical method that we have chosen to analyze the AJT is an Ordinal Logistic Regression Model. This type of regression allows the isolation of the effects of each variable and the identification of the residual effects of the explicative variables (Joseph et al., 2019), in our case, the variable of study is the evaluation (1 to 7 scale) attributed to each sentence. The Ordinal Logistic Regression Model allows a qualitative analysis of variables that are beyond three levels and allows the preservation of the nature of the hierarchical ordered data without transforming it. We have performed the Ordinal Logistic Regression Model using R (R Core Team, 2017).

4.3.2. Results of the Ordinal Logistic Regression Model

For the Ordinal Logistic Regression analysis we have used the package Ordinal (Christensen, 2019), since our data was normally distributed. We applied the function CLMM (Cumulative Link Mixed Models) which allowed us to analyze ordinal observations into multiple cumulative possibilities. We have used the functions *lsmeans* (Lenth, 2016) and *CLD compact letter* display that complements the visualization of pairwise comparisons using the package *MulticompView* (Graves et al., 2019). These functions are used commonly for Linear Mixed Effects Model, but their specificities were applicable for the analysis of our multilevel Ordinal Logistic Regression Model. To make sure that both analyses were equivalent, we have reproduced the same Ordinal Logistic Regression Model using a Linear Mixed Effects Models and the results obtained were equivalents, indicating in both analysis the same model as significant in comparison to the others, which legitimates the use of the Linear Mixed Effects functions such as *lsmeans* (Lenth, 2016) and *MulticompView* (Graves et al., 2019) for post-hoc analysis in our Ordinal Logistic Regression Modeling. Even though, both types of analysis were similar, for the final

modeling, we have opted for an Ordinal Logistic Regression Model since it was better adapted to our ordinal data.

The procedure used for selecting the final Ordinal Logistic Regression Model (see **Table 17**) was identical to the LMER model used in the LDT, we have reduced progressively the model until the non-significant interaction and factors were excluded from the model, beginning by the suppression of the more complex and then the simpler interactions.

Table 17 Ordinal Logistic Regression Model explaining Ordinal data of the Judgment Task

Coefficients	Estimate	Std. Error	Z value	Pr(>|z|)
Bilinguals	-0.05220	0.06522	-0.800	0.42349
Dom_NFR	-0.05008	0.17687	-0.283	0.77705
Condition1	-0.01272	0.18475	-0.069	0.94511
Condition2	0.09452	0.07765	1.217	0.22350
Bilinguals :Dom_NFR	0.04738	0.07657	0.619	0.53608
Bilinguals :Dom_NSP	0.25656	0.08113	3.163	0.00156 **
Dom_NFR:condition1	-0.08959	0.09157	-0.978	0.32787
Dom_NSP:condition1	-0.22440	0.09681	-2.318	0.02045 *
Dom_NFR:condition2	-0.22015	0.09244	-2.381	0.01724 *

The results of the model presented in **Table 17** are completed with the Analysis of Deviance (Type II tests) in Table 18 which calculates the Likelihood-ratio tests, functioning as an ANOVA. We have used the package *Aidememoire* (Hervé, 2020). As it can be seen in Table 18 ANOVA CLMMs has confirmed the output of the Ordinal Logistic Regression Model, indicating a main effect of *Group* and an interaction between the factors *Group*Dom_N and Dom_N*Condition*.

This analysis has explained the ordinal variable target response (divided into multiple probabilities of response, from 1 to 7) and the factors that affect the levels of acceptability attributed by the two groups studied.

Table 18 Result of CLMM ANOVA type II of the Ordinal Logistic Regression Model

	LR Chisq	Df	Pr(>Chisq)
Group	4.1224	1	0.0423181 *
Dom_N	1.3834	2	0.5007116
Condition	0.4920	2	0.7819319
Group:Dom_N	14.5236	2	0.0007018 ***
Dom_N:condition	23.6473	4	9.398e-05 ***

4.3.3. Main Effects

The Ordinal Logistic Regression Model in Table 18 indicates a group effect (LR Chisq= 4.1224, df= 1, $p<0.0423181$, here bilinguals and monolinguals behave differently when judging sentences. The direction of this difference is going to be treated later when exploring interaction effects.

4.3.4. Significant Interactions

Analysis has revealed two interactions detailed in and 4.3.4.2. For a better understanding of the direction of these interactions, post-hoc analyses were made through the observation of contrasts that allowed to detect which factor (e.g., which group) had statistically different scores than the others factors (see 4.3.4.1 and 4.3.4.3). This was possible calculating the Least Square Means for Multiple Comparisons using the function *lsmeans* (Lenth, 2016).This formula makes pairwise comparisons which will indicate differences between the factors, e.g., bilinguals and monolinguals (**Appendix 19**).

4.3.4.1. Interaction between Groups and Neighborhood Density Dominance (Group*Dom_N)

There is an interaction effect between Group and Neighborhood Density Dominance (LR Chisq= 14.5236, df= 2, p<0.0007018) see Table 18. Post-hoc analyses (Least Square Means and Contrasts) showed an interaction between groups and L1 Dominant Neighborhood (**Appendix 19**). This means that bilinguals judged the sentence as more 'acceptable' when they were presented a sentence with a L1 dominant neighborhood density in comparison to monolingual group, and this, independently of the condition presented (see **Figure 28**).

In this task, the sentences were presented in the L1 of all participants, the sensibility that was shown by the bilingual group to high neighbored targets stimuli can be interpreted by bilinguals' advantage in terms of metalinguistic competence (Bialystok, 2001). Metalinguistic competence includes both linguistic knowledge and control of linguistic processing. Further analysis concerning the specific condition in which in which sentences are more or less acceptable are needed (see section 4.3.4.3).

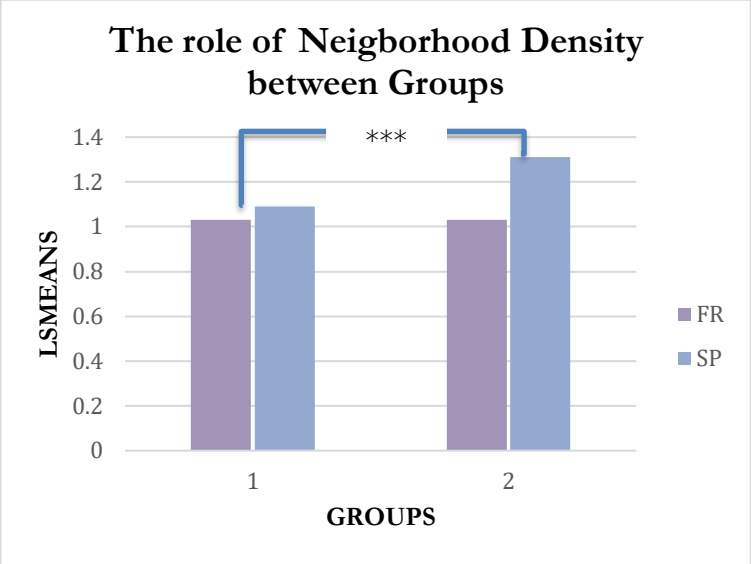

Figure 28 Least Square Means for Multiple Comparisons between Group*Dom_N

Group 1 refers to monolinguals and Group 2 refers to bilinguals.

*4.3.4.2. Interaction between Condition and Neighborhood Dominance (Condition*Dom_N)*

There is an interaction effect between Condition and Neighborhood dominance (LR Chisq= 23.6473, df= 4, p<0.0000001)[33] see **Table 18**.

Least Square Analysis shows an interaction effect between condition and the Neighborhood Density for all groups merged (**Figure 29**).This interaction was significant when the dominant neighborhood language Spanish was presented in condition 1 and 2 (**Appendix 20**).

Condition 1 involves the use of semantic extensions in the sentence *'me ENERVA muy fácilmente Jesús con su manera tan prepotente de hablar'/ 'Jesús ENERVATES me very easily with his arrogant way to talk*) and condition 2 involves the canonical production (here '*me MOLESTA mucho María con su falta de tolerancia hacia los demás*'/ María ANNOYS me a lot with her lack of tolerance towards others). This outcome is very interesting since it opposes contradictory interpretations of meanings and shows that the semantic extensions are less accepted than the canonical propositions.

As it was discussed in section 2.5. if we consider the role of the Reduced Redundancy Principle proposed by Seliger (1991),we can interpret that the effect observed mostly in L1 dominant neighborhood density will be related to the idea that L1 dominant neighborhood functions as a positive evidence for all participants. In this case,

[33] This corresponds to equivalent de 3.09e-5 as shown in the results of CLMM ANOVA type II

L2 resources (L2 Dominant Neighborhood) would not intervene during evaluation sentences in condition 1.

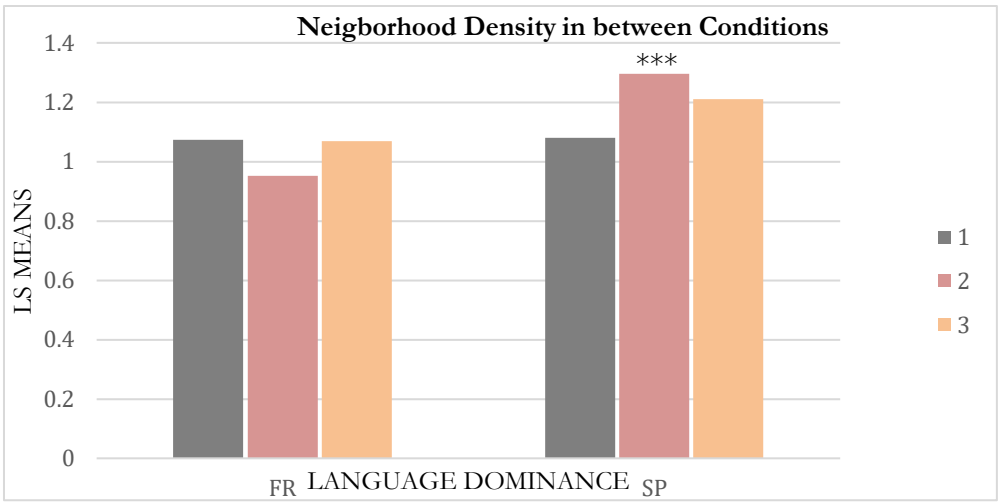

Figure 29 Least Square Means for Multiple Comparisons between Condition*Dom_N all groups merged

In the case of L2 dominant neighborhood no effect was observed for any group (see **Table 17**) following the same perspective, the negative influence is a plausible explanation that would make participants (Group 2 or bilinguals) to accept equally L2 Dominant Neighborhood Density targets in the sentences presented. In this possible case, L1 features that would be partially governed by the L2 dominant neighborhood. However, this particular interaction effect included all groups mingled, furthers analyses are needed to focus on this last alternative interpretation.

*4.3.4.3. Interaction between Group*Condition*

As mentioned above, the procedure of construction of an Ordinal Logistic Regression Model (Christensen, 2019), implies that nonsignificant factors are suppressed progressively in order to obtain a final model representing the data, reducing this way data noise and increasing the statistical power of interactions and main effects, facilitating analysis.

Table 19 Comparisons between Ordinal Logistic Regression Models

	AIC	logLik	LR.stat	df	Pr(>Chisq)
Mordinal6	19	57640	-28801		
Mordinal5	21	57641	-28800	2	0.2462
Mordinal4	23	57645	-28800	2	0.8784
Mordinal3	25	57648	-28799	2	0.4577
Mordinal2	29	57650	-28796	4	0.1988
Mordinal1	31	57653	0.9765	2	0.6137

We aimed to take into account differences of acceptance levels in the three conditions of the bilingual group. However, the interaction between Group and Condition was not significant (LR Chisq= 0.2593, df= 2, p=0.87840) see **Appendix 21**. Thus, this interaction was suppressed and excluded from the final model presented in **Table 17**.

Regarding the effect of Morphological Family Size on the AJT, we had predicted the bilingual group to be more sensitive to condition 1 (expressing a semantic extension) than the monolingual group. Results suggest that bilinguals and monolinguals behave similarly and that there was no effect of condition by itself but exclusively in interactions with Condition and Dominance Neighborhood.

*4.3.4.4. Interaction between Group*Dom_M*

The effect the Morphological Family size was not significant as main effect (LR Chisq= 0.0037, df= 2, p=0.99815) **Error! Reference source not found.** or in interaction with other factors, as Group* Dom_M (LR Chisq= 2.8076, df= 2, p<0.24566), see also **Appendix 21**, or Condition* Dom_M (LR Chisq= 1.5630, df= 4, p=0.8154330) in **Appendix 22**. Accordingly, we have excluded the factor Morphological Family Size (Dom_M) because it did not successfully predict the responses of participants.

4.3.5. Discussion

The objective of this task was to compare bilinguals and monolinguals in their assessment rates of the sentences presented in the three conditions. As we expected, bilinguals tended to evaluate sentences as more acceptable. However, this was the case for all conditions mingled, and specifically when target stimuli Neighborhood Density was dominant in the L1.

We expected differences in the evaluation of sentences that expressed semantic extensions in the L1 (Condition1) in comparison to canonical meanings (Condition 2). However, such effect was observed in both groups (bilinguals and monolinguals). This means that when participants evaluated sentences with uncanonical meaning (having target stimuli which were L1 dominant in neighborhood), these evaluations were less accepted than canonical meaning sentences' evaluations. In contrast, no significant effect was observed when the targets words within the sentences were L2 dominant in neighborhood.

4.4. GAP COMPLETION TASK

The Gap Completion Task (see **Figure 30**) is a controlled task that allows us to elicit language production of one or two word responses to an audive stimuli. The task is presented in L1 (Spanish) to all participants. Stimuli presented consisted on recorded sentences are detailled in **Appendix 11**.

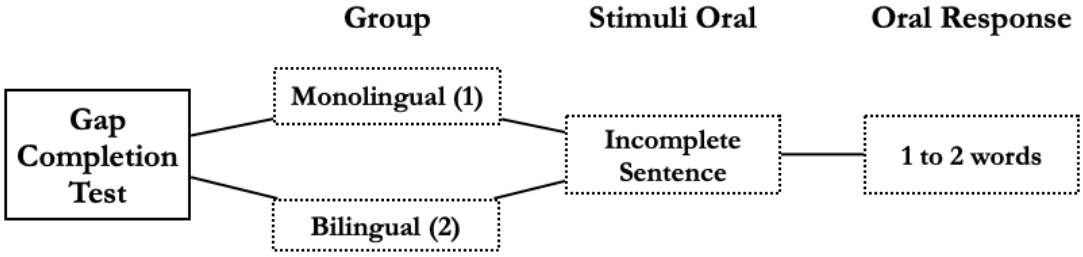

Figure 30 Experimental design of the Gap Completion Task

The main objective of the Gap Completion Task is to observe the oral of production both groups, in terms of canonical and uncanonical (L2 influence on the L1) responses during controlled L1 speech production.

4.4.1. Data Transcription

For the group (1) we collected 2226 responses and for the group (2) we gathered 1680 responses. Oral productions gathered in this task were orthographically transcribed (for a total of 42 responses of 1 to 3 words for each participant) by the researcher. These transcriptions were then checked by two native speakers to assure the accuracy of the first transcription. After comparison, the different types of responses were classified.

The orthographic transcription paid special attention to the phonological production of the bilingual group when it was necessary. Hesitations were noted, and inaudible productions coded as NA. When participants hesitated the first uttered response was the one that was included as response. Once all responses were transcribed, the RTS (see section 4.4.5) was calculated and the quality of the responses collected coded simultaneously (see section 4.4.2). Both analyses, quantitative — involving RT's and the percentage of repeated productions per group — and qualitative — involving uncanonical productions — (see section 4.4.8) were applied at comparing the responses obtained between groups.

4.4.2. Data Coding

The Gap Completion Task was conceived in order to observe the differences of lexical choice of bilinguals and monolinguals in L1 while completing a sentence. Different

targets responses were expected for each group, in bilinguals' semantic transfers from the L2 were expected as the evidence of crosslinguistic influence in a task performed in L1. For the monolingual group, more canonical responses (in accordance with the standard use of the language) were expected than in the bilingual group.

Depending on the types of phenomenon found and in order to take into account different types of responses, 6 responses categories were used and coded s 0,1,2,3,4,5, as noted in **Table 20**.

Table 20 Coding the type of responses of the Gap Completion Task

Code	Type of responses
0	NA (no answer)
1	Canonical
2	Plausible
3	Error
4	Semantic Extension
5	Other types of CLI

0) We coded as 0 the NA including any hesitation or productions such as 'no, I don't know', or productions having exceeded the time set of 4000 milliseconds. These productions may consist only of a vowel or any inaudible string.

1) 1 refers to the canonical or the expected target response, i.e., the exact predicted word or a close synonym. For example, for some stimuli the responses were very homogeneous such as stimuli R18 (results showed in **Table 21**) consisting in the following sentence:

Stimulus R18: Juan ha perdido su billetera, y ni puede efectuar ningún pago, puesto que ha bloqueado todas sus _____ (Tarjetas).

Translation: Juan has lost his wallet, and he cannot make any payments because he has blocked all his_____(Cards).

Table 21 Extract Sample of Canonical Scores of Stimulus R18

Tag	Group	Canonical response 1	%	Canonical response 2	%	Other responses %
R18	Monolinguals	Tarjetas	88,68%	Cuentas	11,32%	0%
R18	Bilinguals	Tarjetas	82,50%	Cuentas	15,00%	2,50%

2) Plausible responses were coded as 2. This category includes the use of antonyms or semantic related words that approximate the canonical responses and are plausible responses to the stimulus (context) presented. Plausible responses were not entirely satisfactory following our judgment. We have used as indicator of semantically related words using RAE reference)*Real Academia Española* (RAE,

2020) for Spanish, and for French, the reference of *CNRTL* (*Centre National des Ressources Textuelles et lexicales* (CNRTL, 2020).

The example presented in **Table 22** illustrates the variations obtained by a plausible response (1) which was most frequently produced by the participants and, other plausible responses (2) that are less representative productions, but nevertheless considered as accurate responses for stimuli R20.

Stimulus R20: El maestro de matemáticas explica bien, pero habla con una voz muy baja; me cuesta trabajo _____ (Escucharle).

Translation: the math's teacher explains well, but he speaks with a very low voice, I have a hard time to _____(Listen).

Table 22 Extract Sample of Plausible Scores of stimulus R20

Tag	Group	Plausible response 1	%	Plausible response 2	%	Other responses %
R20	Monolinguals	Entenderle	30,19%	Aprender	1,89%	49,05%
R20	Bilinguals	Entenderle	25,00%	Oírle	2,50%	55,00%
R20	Monolinguals	Entenderlo	5,66%			
R20	Bilinguals	Entenderlo	17,50%			
R20	Monolingual	Entender	13,21			

3) Responses coded as 3 relate to erroneous or implausible responses for the given context. This category includes productions that have no semantic relation with the sentence. Errors include also productions that the different judges evaluated as unintelligible, mispronounced words, and ungrammatical complements to the sentences. These scores are very useful in order to point out the stimuli that were not context specific enough. **Table 23** illustrates the coding of stimulus R25 for the sentence:

Stimulus R25: ¡Lamentable! Bebió demasiado y condujo, sin prever las consecuencias que eso le pudo' _____ (Ocasionar).

Translation: 'Unfortunate, he drank too much and drove, without foreseeing the consequences that it would' _____(Cause).

The coding of this category refined data cleansing procedure (described in section 4.4.2) as it was used as an indicator for the stimuli that were excluded for the analysis of RT in order to avoid bias of general RTs patterns.

Table 23 Extract Sample of Errors scores obtained of stimulus R25

Tag	Group	Error	%	Error	%	Other responses
R25	Monolinguals	Ocurrir	13,32%	Acaerrear	1,89%	75,35%
R25	Bilinguals		2,50%	Dar	2,50%	90,00%
R25	Monolinguals	Pasar	3,77%	Notecer	1,89%	
R25	Bilinguals		2,50%	Traera	2,50%	
R25	Monolinguals	Contraer	1,89%	Suceder	1,89%	

4) Category 4 refers to Semantic Extensions, an uncanonical production that can be explained by the influence of semantic features from the L2 on the L1. For this experiment, stimuli were chosen according to their potential to give rise to Semantic Extensions that are noncanonical productions. Table 24 shows the code corresponding to Semantic Extension for stimulus R30:

Stimulus R30: Lucia estaba muy conmovida porque su familia para su cumpleaños, un regalo muy especial le han _____ (Dado vs Ofrecido).

Translation: Lucia was very moved because her family for her birthday a very special gift to her was _____(Given vs Provided).

Table 24 Extract Sample of Semantic Extensions of stimulus R30

Tag	Group	Semantic Extension	%	Other responses %
R30	Monolinguals	Ofrecido	0%	100%
R30	Bilinguals	Ofrecido	12,50%	87,50%

The semantic extensions that were produced by participants in this experiment are described and discussed individually in section 4.4.7.

5) The productions coded in category 5 include different types of CLI other than semantic extensions such as loan translations (literal translation), phonetic transfer (changes in the pronunciation of phonemes in L1), borrowings (words produced in L2 French) and blended words. These different manifestations of CLI were compiled into one single category since the overall number of productions were not very important. However, in the qualitative analysis in section0 a description of each subcategory is proposed (**see Table 35**).

As an example of the coding that was made for the 42 stimuli tested, we provide an example in **Table 25** describing the responses obtained in stimulus R7 by the two groups, this illustrates the complexity of this type of experimental task in terms of diversity.

Stimulus R7: Para cargar todos sus cosméticos, sus llaves y su teléfono; Anna se compró un pequeño _____ (Bolso vs Saco).

Translation: To pack all of her cosmetics, her keys and her phone; Anna has bought a little _____(Bag vs Basket).

Table 25 Complete coding for Stimulus R7

Tag R7	Code	Monolingual %	Bilinguals %
NA	0	1,89%	0%
Bolso	1	79,25%	55,00%
Maletín	1	5,66%	2,50%
Moral	1	1,89%	0%
Estuche	2	1,89%	15,00%
Maletero	2	1,89%	0%
Cosmetiquero	2	0%	2,50%
Monedero	2	0%	5,00%
Billetera	3	1,89%	0%
Corazón	3	1,89%	0%
llavero	3	1,89%	0%
Pantalón	3	1,89%	0%
Celular	3	0%	2,50%
Cajón	3	0%	2,50%
Bolsillo	3	0%	2,50
Saco	4	0%	7,50%
Necessaire	5	0%	5,00%
		100%	100%

A total of 3906 responses produced by 40 Hispanic living in France and 53 Hispanic tested in Colombia for 42 stimuli were coded in 6 categories described above.

The whole data was independently coded by the experimenter (a Spanish native speaker from Colombia), and by a second native speaker from Latin America who had been very recently immersed in France (LoR: 2 months).

The two codings were crossed-evaluated until a common decision was adopted for each lexical response. When this was impossible, a third and final evaluation exchange was organized to come to an agreement.

Section 4.4.3 reports the percentages of responses in each category obtained by each group (all stimuli merged). In section 4.4.7 the responses classified as semantic extensions will be presented more in detail.

4.4.3. Global Results

In this section we describe the global results obtained per group following the coding procedure described previously. This description comprises the percentages obtained of the categories (0 to 5) defining the quality of the responses per group and accordingly to the context presented (incomplete sentences). These first global results allowed to complete data cleansing in which stimuli and responses were excluded (see section 4.4.7) for further analysis. It should be recalled that data analysis of the Gap Completion Task exclude the factor Dominant Neighborhood and Morphological Family Size, because oral input/stimuli presented were expected to trigger target words (involving or not Dom_N or Dom_M) so this variable cannot be controlled by the experimenter, besides sentences construction (see 3.6.1.2.),were no equated enough regarding both categories.

As can be seen in **Table 26**, the percentages of responses in each category shows that the responses produced by bilinguals and monolinguals are very similar. NA for both groups are very close, 4,72% for monolinguals against 5,18% for bilingual group.

Table 26. Percentage of the Responses in the GCT

Response	Code	Monolinguals	Bilinguals
NA	0	4,72%	5,18%
Canonic	1	63,99%	64,29%
Plausible	2	17,83%	18,33%
Errors	3	13,43%	6,73%
SE	4	0,91%	2,44%
CLI	5	0,13%	3,04%
		99,99%	100,00%

For canonical responses category, the bilingual group obtained 64,29% against 62,98% for the monolingual group. Regarding plausible responses differences are very marginal for the two groups with 17,83% for monolinguals and 18,33% for bilinguals. Briefly, it can be said that the two group of Hispanics have similar production in terms of the quality of global responses the percentage of accuracy defined here by means of NA, canonical and plausible scores obtained by each stimulus, which were used as indicators for the inclusion or exclusion of the stimulus. **Table 27** indicates that the percentages obtained per category were not significant across group. $t(5)= 1.199^{e}$,p=<0 .5.

Table 27 Paired-T-Test between groups

Paired Samples T-Test				
Measure 1	Measure 2	t	df	p
Group2	- Group1	1.199^{e}	-8 5	1.000

Note. Student's t-test.

Nevertheless, slightly differences across groups can be noticed, as showed in **Table 26** error rates obtained by bilinguals (6,73%) are lower than those of monolinguals (13,43%) suggesting that bilinguals are more accurate than monolinguals in their L1, however this output can be the result of lack of motivation observed in the monolingual group towards linguistic task, which was not the case for bilinguals who had studied foreign language (s) over the years.

But what is most interesting for us are the percentages of CLI and SE that are slightly higher for the bilingual group. These are not considered as errors, but intrusions of the other languages of the participant. Bilinguals produced, as expected, more semantic extension than the monolingual group with 3,04% against 0,13%, and more CLI for 3,04% against 0,13%. These two categories, which are the focus of our study, are analyzed individually in section 0 taking into consideration sociolinguistic information of the participants detailed.

4.4.4. Data Cleansing

It should be recalled that for the construction of the Gap Completion task, each sentence was supposed to provide enough contextual information to allow the participants to produce the target responses. Therefore, each sentence was pre-tested in an online test by 6 native speakers of Spanish in an L1. Only stimuli that were accurately completed by all the participants of the pre-test were included in the final experiment, resulting in 42 revised or corrected sentences. Nevertheless, the global responses show that some stimuli used were not sufficiently contextualized and gave rise to very heterogenous responses. Therefore, these stimuli were excluded before the statistical model was applied to the data.

Data cleansing procedure was made depending on the percentage obtained per category. The type of errors found have allowed to suppress data of stimulus that had multiple responses that deviated from the expected ones, and those who had a combination of more than 10,00% of NA and less than 50,00% of canonical responses per group. The stimuli were discarded when they had high scores of absences of responses (hesitations, and NA) and when very few common patterns according to all the responses obtained (compared by stimulus and by group). In these cases, when the responses obtained were spread out into multiple propositions (including errors or multiple plausive responses) with low occurrence (percentages) and with high NA values; we have deduced that the sentences or stimulus concerned were not targeted or specific enough, being too ambiguous and giving as a result inconsistent and heterogenous responses. Thus, data cleansing procedure is then articulated through the percentages obtained in NA by each stimulus and the consistency in responses (errors, canonical and plausible responses) obtained by both groups.

Data cleansing was made by stimulus when both groups sparse responses agreed with the fact that the sentence was too ambiguous or unspecific. 9,90% NA of answers

corresponded on 4,72 % for the monolingual group and 5,18% for bilinguals. First, we have used NA% indicated by stimulus and by group. NA comprised three types of responses: hesitation marks such as '*um*', '*no*', and '*I don't know*' and inaudible or mispronounced productions. When the percentage obtained for the canonical responses were inferior to 50,00% and plausible responses were inferior to 42,00% per group, this worked as an indicator showing that the targeted production was not successfully achieved. For example, in R37 and R23 stimuli the scores of canonical responses are 0% in both populations as described in Table 28.

Table 28 NA and Errors Percentages of Stimuli Suppressed for RT analysis

Group	Tag Target	NA%	Error %	Canonical %	Plausible%
Bilingual	R21 Cazados	12,50%	22,50%	32,50%	32,50%
Monolinguals	R21 Cazados	20,75%	24,55%	26,42%	28,31%
Bilingual	R31 Concedida	0%	5,00%	57,50%	37,50%
Monolingual	R31 Concedida	5,66%	24,55%	37,74%	32,09%
Bilinguals	R37 Mientras	65,00%	27,50%	2,50%	5,00%
Monolinguals	R37 Mientras	39,62%	58,54%	0%	1,89%
Bilinguals	R23 Extraídas	25,00%	25,00%	0%	47,50%
Monolinguals	R23 Extraídas	16,98%	37,76%	0%	45,32%
Bilinguals	R26Encargada	10,00%	17,50%	28,00%	43,00%
Monolinguals	R26 Encargada	13,21%	35,88%	30,19%	20,75

The general objective of the quantitative analysis was to identify patterns in each or both groups following the categories observed in section 4.4.2. including all data collected. However, data cleansing procedure is necessary for the chronometric analysis of reaction time. The data that are included are Canonical (1) Plausible (2) Semantic Extensions (4) and CLI production.

From 3906 productions, the percentage of suppressed data is 465 responses for 5 stimuli described in Table 28, which correspond to 11,90% of the total number of responses. From the remaining data, Errors (3) and NA (0) are suppressed representing 2,70% of the remaining data for 92 responses and (3) errors representing 8,24% for 281 responses that were considered as inaccurate productions.

4.4.5. Calculating Reaction Time

The maximum threshold of response time allowed is 5.000 milliseconds, past this time limit, recording was stopped in E-rime. The annotation of recordings was made in PRAAT, where spectrograms were automatically generated.

We marked manually the beginning of each sentence and then we have semi-automatize PRAAT to gather multiple annotation files that allow to calculate multiple

reaction times simultaneously. We analyzed mostly the beginning of the uttered speech, which was the first sound detected by PRAAT in order to make the annotations. However, in some cases, additional information was required to mark the beginning of the production accurately (see **Figure 31**). We corrected the RT in some particular cases, when the response began with occlusive consonants. Segmentation was different for initials occlusive consonants, for which we have counted between 40 and 70 milliseconds before the explosion, otherwise the voice onset time (VOT) following the indications of corrections of Vaissière (2010) and Nocaudie (2016).As is detailed in **Figure 31**.

For vowels, fricatives consonants, liquids and glides, segmentation was made taking into account the beginning of the spectrogram as well as pitch, intensity and formants.

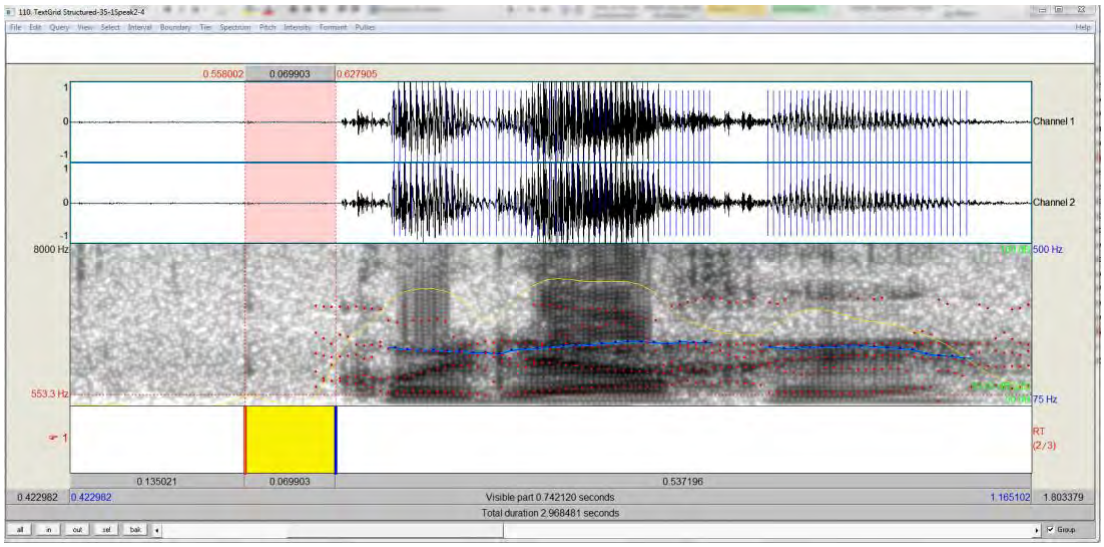

Figure 31 Example of the Annotation Procedure of a response made in Praat to calculate RT

In this production '*Coraje*' was corrected adding to the initial RT of 0,69903 milliseconds

4.4.5.1. Final Data

Remaining data includes correct responses (canonical and plausible) and uncanonical responses (semantic extension and CLI productions) produced by both groups. As a result of the cleansing data procedure and before the stage of statistical model construction we have reduced data to 3068 productions. The distribution of the three types of responses is reported in **Table 29**. As can it be seen, differences between groups are marginal for canonical and plausible responses. Monolinguals' scores are slightly higher for bilinguals than monolinguals, with 79,27% against 75,31% for canonical responses and with 19,38% against 18,02% for plausible responses.

On the other side, the number of uncanonical productions is higher for the bilingual group than the monolingual group. Semantic extensions represent 3,00% of the responses for the bilingual group against 1,17% for the monolingual group.

Table 29 Final Percentage of the of responses of the Gap Completion task

Response	Code	Monolinguals	Bilinguals
Canonic	1	79,27%	75,31%
Plausible	2	19,38%	18,02%
SE	4	1,17%	3,00%
CLI	5	0,18%	3,66%
		100,00%	100,00%

However, the 1,17% of semantic extensions reported for the monolingual group was mistakenly produced by an inversion of the battery of tasks. In fact, 0,65% of the 1,17% of the semantic extensions matches with an inverted order of tasks during experimentation, especially when LDT preceded GAP (see **Appendix 23** describing the order of completion of each experimental task by each participant).

Regarding the CLI observed, it should be noted here that 0,18% of the productions produced by the monolingual group is due to influence from English, learned in second language classes at the university. Actually, all participants have studied English for a length of time comprised between 6 months and 4 years No CLI of French was observed in the monolingual group, confirming our interpretation that semantic extensions observed were not due to interference from French in this group, but to sensitivity to target stimuli with similar-word form in Spanish.

Bilingual group production of CLI is very rich (3,66%), it includes different phenomenon such as: loan translation, borrowing and phonetic transfer. An interesting link can be made between the LoR (Length Of Residence) of the participant that produce CLI. Further analyses are presented in section 0 in which we proposed very brief case studies of participant that produced the most CLI and their sociolinguistic background.

4.4.6. Statistical Procedure

The experimental design of the CGT includes one dependent variable and two independent variables. The dependent variables are the **Response Time (RTs)**. The independent variables are: **Group** (**Monolingual** vs. **Bilingual**) and the and the **Response Coding** of responses (including canonical, plausible, semantic extensions and CLI). To analyze the data, a Linear Mixed Effects Model was created to explain RT differences across groups, RT differences when the responses were either plausible, SE,

CLI in comparison to canonical responses (used as referent value) and how changes in the RT are linked to the responses produce by the groups (plausible, SE and CLI).

4.4.7. Results of the Linear Mixed Effects Model

The results of the Linear Mixed Effects Models are detailed in **Table 30**. RT (dependent variable) is analyzed following the production of two groups (bilinguals and monolinguals) of participants and their responses attributed (Plausible, SE and CLI) to each stimulus presented. The random effects included in the model were groups and the stimulus presented. The referent variable for group is *Group 1* (monolinguals) and the referent variable for the types of responses attributed is *code 1* or canonical response type. The referent variable operates as a comparison referent, or control variable to the others for analysis (e.g. *group 1*).

Table 30 Gap Completion task Linear Effects Mixed Model

	Estimate	Std. Error	df	t value	Pr(>|t|)
Intercept	0,8317	0,04408	1.03200	18.866	<2e-16 ***
Group2	0,003451	0,04.955	9.7810	-0.070	0.94462
Plausible	0,01979	0,03.061	2.996000	6.466	1.17e-10 ***
SE	0,04994	0,1094	2.990000	0.452	0.65140
CLI	1,324	0,2737	2.969000	4.835	1.40e-06 ***
Group2:Plausible	0,04172	0,04.419	2.965000	0.944	0.34518
Group2:SE	0,1716	0,1304	2.969000	1.316	0.18827
Group2:CLI	-1.033	0,2800	2.962000	3.688	0.00023 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 31 Correlations effects between RTs and independent variables

	(Intr)	Group2	code2	code4	code5	Grp2:2	Grp2:4
Group2	-0.484						
code2	-0.137	0.112					
code4	-0.035	0.028	0.052				
code5	0.015	0.011	0.027	0.000			
Group2:cod2	0.085	-0.174	-0.635	-0.035	-0.016		
Group2:cod4	0.027	-0.058	-0.034	-0.811	-0.001	0.067	
Group2:cod5	0.014	-0.027	-0.026	0.004	-0.969	0.033	0.007

4.4.7.1. Main effects: Group effect

Concerning the differences in RT between groups, data suggest that no group effect was observed: β *=0,003451, SE=0,04.9., t=-0.070, p= 0.94462* nor it did for semantic extensions β *=0,04994, SE=0,1094 t=0.452, p=0.65140* as it was expected.

4.4.7.2. Main effects: Plausible and CLI productions

We have found two main significant effects, first, a difference is observed in RT of participants when responses were plausible and when responses were the result of CLI, here all groups are merged.

Plausible responses took globally longer than the other five categories for all the participants of the study: $\beta = 0{,}01979$, $SE = 0{,}03.061$, $t = 6.466$, $p < 0{,}001$, (see **Table 30**).

This output is not surprising because plausible responses are considered to be an alternative response for the canonical options (that completed the best the sentences or stimuli presented). We assume that plausible responses were a second option when the first option (canonical) was not available right the way in the mental lexicon because it involved longer RTs. When observing mean differences in between groups for canonical and plausible responses (see **Table 32**) canonical responses are faster by both groups than plausible responses.

Table 32 Means of Canonical and Plausible responses

Group	Mean Canonical	Mean Plausible
1	0,81344	1,05848
2	0,80384	1,13694

On the other hand, CLI productions are significantly different from the other types of responses by 1,324 milliseconds, all groups merged (see **Table 30**) fixed slope estimate for CLI responses are: $\beta = 1{,}324$, $SE = 0{,}2737$, $t = 4.835$, $p < 0{,}001$.

These results should be nevertheless interpreted carefully because the total production of CLI correspond only for 3,84% of the data. Further analysis and more data should be collected to advance further conclusions for this category. It should be noted that the advantage of this type of modeling is the inclusion of the subject variability.

4.4.7.3. Interactions: CLI Production by the Bilingual Group

The Linear Mixed Effects Models displayed an interaction between the RTs produced by the bilingual group in the production of CLI in comparison to the monolingual group. A significantly facilitatory effect of -1.033 milliseconds is shown for bilinguals: $\beta = 1.033$, $SE = 0{,}2800$, $t = 3.688$, $p < 0.00023$.

This result suggests that bilinguals use the available linguistic resources (L1 and L2) to cope with missing word problem presented in the AGT, and this facilitating RT in comparison to the monolingual group.

4.4.8. Qualitative Analysis of Semantic Extensions

Qualitative analysis of the specific responses obtained by the bilingual group are valuable to understand the functioning of uncanonical production including semantic extensions and CLI in oral L1 production.

In this section we will describe Canonical, and SE and responses obtained across groups, that is 11/42 stimuli. **Appendix 25** describes the percentages of responses obtained for the 11 stimuli involving semantic extensions and canonical productions per group. **Appendix 24** for a detailed description for the 42 stimuli tested.

Qualitative analysis is focused on the semantic extensions and canonical responses produced by the bilingual group. This analysis is based on online resources for each language taking into consideration lexicographic definitions for French and Spanish. For French, we used the resources of the *Centre National des Ressources Textuelles et lexicales* (CNRTL, 2020) which constitute an important linguistic data resource and natural language processing tool created by the *Centre National de Recherche Scientifique* (CNRS) in 2005. It includes the survey, documentation, normalization, storage, sharing, enrichment and publication of data on the French language. For Spanish, the official online resource is RAE *Real Academia Española*, which is a juridical institution in charge of the evolution of the Spanish Language in accordance to the needs of Hispanic speakers. The online dictionary resource (RAE, 2020) includes a complete overview of language use by Hispanics around the globe considering different varieties of Spanish (Latin America such us: Colombian, Peruvian or Venezuelan). REA and CNRTL resources are highly representative of language use, definitions, and constant updating of language changes by language experts.

We analyzed each stimulus, the corresponding responses and the percentage obtained per group, individually. In most cases, semantic extensions were produced by the bilingual group and these productions are interpreted as the result of the influence of the L2 on the L1.

4.4.8.1. Gentil

Stimulus R42: *Un joven ayuda a una mujer mayor con sus compras, ella le responde muchas gracias es usted muy_____(Amable vs Gentil)*

Translation: *A young man helps an old woman with her groceries, she answers thank you very much, you are very_____(Kind vs Gentle)*

In this example, the target sentence expresses a semantic extension that is articulated by a very subtle difference between '*amable*' (kind) in Spanish and '*gentil*' (gentle) in French.

Both languages share the Latin origin '*gentilis*' for the word '*gentil*' (RAE, 2020; CNRT, 2020) which means these are true cognates defined as sharing the same etymological origin (Jarvis, 2009). The use of '*gentil*' by both groups itself is not an uncanonical production, but a low frequency word effect that will affect the speaker word-choice processes (Hohenstein et al., 2004; Jarvis & Pavlenko, 2010). The contrasting frequencies of '*gentil*' and '*amable*', (see **Figure 32**), low frequency of words will play a role during language accessibility.

Figure 32 Frequencies of use of Amable and Gentil in the data base resource Books Ngram Viewer

Regarding the target word meaning, RAE (2020) referenced '*gentil*' as: '*pagan, spirited, remarkable, polite,* and *kind*. Due to the multiple meanings related to this particular word, it is indeed surprising that monolinguals used it in 13,21 % of the production studied against 5,00% for the bilingual group as shown **Figure 32**. We interpreted this result as a sociolinguistic phenomenon of *hypercorrection* that occurs when the speaker privileges rare or complex words, structures or phonological realizations, in order to unconsciously or consciously prove to have a sophisticated knowledge of the language tested (Labov, 2006). This may be explained here by the fact that the testing of the monolingual group took place at university.

The use of '*gentil*' in the bilingual group cannot be considered as semantic transfer from L2 here since monolinguals produced the same pattern. Another explanation is that that true cognates stimulus makes the bilingual to avoid L1 words that are formally close to the L2, increasing this way interference awareness.

4.4.8.2. Coraje

Stimulus R10: *Lanzarme del paracaídas? ? No gracias, para esas cosas no tengo suficiente suficiente _____(Valentía vs Coraje)*

Translation: *Skydiving? no thanks, for those things I don't have enough _____(Bravery vs Courage).*

The target responses *'courage'* and *'coraje'* again share the same etymological origin, *'Coraje'* originates from old French '*courages*'(RAE, 2020; CNRT, 2020). *'Coraje'* is defined as 1) courage and 2) anger. In the target sentence we expected participants to produce more frequent forms, such as '*valor*', or '*valentía*' (accordingly to the sentence context) which was the case for 50% of the production in both groups (see **Appendix 25**). Regarding the use of the form *'coraje',* we have found that 30% of the bilinguals privileged this word choice, while 16,90 % of the monolingual group used this form. As illustrated in **Figure 32** *'coraje'* has also a low frequency of use in comparison to '*valor',* (see **Appendix 9** for stimuli details). However, in some varieties of Spanish *'coraje'* is very frequent, namely in Central Latin America, for South America, the use is relatively weak. However, data of frequencies classified by country has not been found support this interpretation.

The patterns observed here are the same as for *'gentil'*: monolinguals tend to produce more the uncanonical form instead of the canonical form '*amable*'. This is interpreted as a canonical albeit infrequent lexical choice related to either L2 influence on the L1 or to hypercorrection prompting participants to use more infrequent and complex words instead of frequent and simple ones.

4.4.8.3. Anciana

Stimulus R24: *Esta casa fue construida en 1930, su valor comercial disminuye, puesto que es una casa muy _____(Antigua vs Anciana)*

Translation: *This house was built in 1930, its commercial value decreases because it's very _____(Old vs Ancient)*

The semantic extension expressed in this stimulus comes from the extended meaning of '*ancien*' generalized into the word '*anciano*' in Spanish. In Spanish '*anciano*' refers to an elder person but it may not be used with objects (e.g. **casa anciana*, opposing Spanish to both French '*maison ancienne'* and English '*ancient house*'. This semantic extension is listed by Quilis et al (1982, p. 153) as "*anciano—viejo (< a n c i e n): casas ancianas (o).*"

Both groups of Hispanics used the canonical form '*antigua*' around 47% of their productions. We have also found a production of semantic extension for 2,5% of the responses obtained by the bilingual group no uncanonical responses were produced by the monolingual group.

4.4.8.4. Saco

Stimulus R7: *Para cargar todos sus cosméticos, sus llaves y su teléfono ; Anna se compró un pequeño_____(Bolso vs Saco)*

Translation: *To pack all of her cosmetics, her keys and her phone; Anna has bought a little_____'(Bag vs Basket)*

The words '*Sac' and 'Saco'* are true cognates since both originated from the Latin '*saccus*', which comes from Greek '*σάκκος*' '*sákkos*' (CNRT, 2020; RAE, 2020). The semantic extension expressed in stimulus R7 is the result of the extended meaning of '*Sac*' in French into the Spanish word '*Saco'* which is defined by the RAE as a '*receptacle made of leather, fabric or paper*'. However, to refer to the clothing accessory, in this context the most frequent word is '*bolso*' and not '*saco*'. This semantic extension was described by Quilis et al, (1982, p. 157) as *"saco—cesta (<s a c)"*, here '*basket*' is actual meaning of '*saco*' in Spanish. Besides, '*Saco*' also makes reference to 'jacket' in some varieties of Spanish (e.g. Colombian or Venezuelan).

Regarding the responses obtained, both groups seems to agree that '*bolso*' is an accurate response for 79,25 % in the control group and for 55,00% in the bilingual population. Instead, '*Saco*' is used by the bilingual group for 7,50 % while there was no production of '*Saco*' for the monolingual group.

4.4.8.5. Ofrecido

Stimulus R7: *Lucia estaba muy conmovida porque su familia para su cumpleaños, un regalo muy especial le han _____(Dado vs Ofrecido)*

Translation: *Lucia was very moved because her family for her birthday a very special gift to her was _____(Given vs Provided)*

The semantic extension in this sentence is produced by a misleading association between '*Ofrecer*' and '*Offrir*'. In Spanish '*Ofrecer*' is defined as a commitment to give, to say or to do something *e.g., 'ofrecer ayuda'- 'to provide help*'). In others spheres of meanings, we find also to manifest, to present, to involve (e.g., '*la universidad ofrece cursos intensivos' – 'the university offers intensives courses'*).

As in English, '*offrir*' in French refers to offer a gift to someone ('*offrir des fleurs*- in English give flowers) or to pay someone something (e.g. '*offrir un verre'-' to offer a drink'*), to propose something to someone such as remuneration for a service or work.

In some cases, the meaning of '*Ofrecer*' in Spanish is not equivalent to the meaning of '*Offrir*' in French and vice versa. For example in Spanish '*ofrecer un homenaje*' is not transposable to the French '*offrir ses hommages***(CNRT, 2020; RAE, 2020). Quilis et al., (1982, p.156) had recognized the broad meaning of *"offrir as ofrecer—regalar (< o f f r i r regalar y ofrecer)".*

Nonetheless, in Spanish it can be said '*ofrecer un trago*' ('*to invite a drink*') in a very similar way as in French '*offrir un verre'*, but to express the meaning of French '*offrir un*

cadeau' ('*to offer -give a gift*') in Spanish one can also use '*dar un regalo'* or use the verb '*regalar*' that implies the act of offering something as a gift.

From this perspective, the L2 seem to affect the bilingual word choice as evidenced by the percentages of 12,50% for '*Ofrecido*' and NA for the bilingual group. Regarding the canonical forms, in this particular context for 50,00% in the bilingual group and 50,94% for the monolingual group. The use of '*Regalar* ' considered as a plausible response is higher for monolinguals than bilinguals with 30,19% against 7,50% of the responses.

4.4.8.6. Informaciones

Stimulus R30: Para llevar a término la investigación es necesario recopilar un mayor número de _____(Datos vs Informaciones)

Translation: To finish the research it is necessary to compile a bigger number of _____(Data vs Information)

The uncanonical target expected here is the plural form of '*información*' in Spanish, ('*informaciones*') which is associated with the use of '*informations*' (plural) in French. '*Informations*' (CNRT, 2020) is used in French to define facts, new events that are for public interest in order to get informed. It can also refer to a set of knowledge gathered about a determined subject (e.g. '*accès aux informations gouvernementales*' - access to government information) (CNRT, 2020).

The use of plural '*informaciones*' in Spanish is an extended meaning of French instead of the canonical use of '*datos*' ('*data*'). Bilinguals produced '*datos*' in 47,50% against 28,30% for monolinguals. The semantic extension '*Informaciones*' is used for 2,50 % by the bilingual group and NA for monolinguals.

4.4.8.7. Disputa

Stimulus R17: Los padres de Juan se van a divorciar, anoche tuvieron una fuerte _____(Discusión vs Disputa)

Translation: Juan's parents are going to get divorced, last night they had a heavy _____(Discussion vs Argument)

The semantic extension here involved a very subtle difference between '*discutir*' and '*disputar*' in Spanish and their counterparts in French ('*discuter*' and '*disputer*') and the specificities of each language to express distinctly the concepts of discussion and argument.

In French dispute means to debate about a subject (e.g. '*disputer de*', '*sur*') or to be in conflict with somebody about a positioning or an idea (e.g. '*disputer avec qqn de*'), from this

perspective, there is not a clear-cut distinction between discussion and dispute_CNRT, 2020). In contrast, in Spanish the concept of '*disputa*' or '*disputar*' is more physically orientated considering its multiples definitions: '*to fight*'; '*to combat*', '*a heat discussion*', or '*to compete*'(RAE, 2020).

The concept of '*Discusion*' in French refers '*to examine*' or '*to debate*' about a given subject defending different point of views. This definition will match with '*dispute*' in French.The concept of '*discusión*' in Spanish is closer to the concept of '*dispute*' to express debate. Besides '*dispute*' comprises both concepts in Spanish '*discusión*' and '*disputa*' which is not the case in the opposite direction '*dispute*' do not include the concept of debate as it was pointed by Quilis et al.,(1982, p. 154) in *"disputa-discutir (<d i s p u t e r disputar y discutir) <débattre (e)".*

This production serves as illustration of the differences between semantic and conceptual transfer exposed by Pavlenko (2009) (discussed in 1.1.7.1.). This example refers to a conceptual transfer rather than a semantic transfer because the transfer of linguistic knowledge involved conceptual specific features. In this case, formal misleading association between interlingual word-pairs is less plausible.

Global percentages of stimulus R17 indicates that more attributed response was '*discusión*' with 88,68% for the monolingual and 72,50 % for bilinguals. The use of '*disputa*' was exclusively observed in the bilingual group for 7,50% against NA for the monolingual group.

4.4.8.8. Partir

Stimulus R7: *Está muy agradable la fiesta, lo siento, desafortunadamente mañana madrugo, por eso tengo que _____(Irme vs Partir)*

Translation: *This party is very pleasant, I'm sorry, unfortunately tomorrow I have to get up early, so I have to _____(Leave vs Split)*

In stimulus R27, the semantic extension is originated from the association between the identical form of the verb '*Partir*' in French and in Spanish. Both forms shared the same etymological Latin origin '*partīre*'. Following the CNRT (2020) and the RAE, (2020), the main difference of meaning between this true false cognate is that '*Partir*' in French (in its intransitive form) means '*to leave to a place'*, it is commonly used with prepositions to refer to different contexts (e.g. *pour, vers, à, en, dans, de*) which make it a high-frequency word in French, (see **Appendix 6**).

In Spanish the meaning of '*Partir*' means for '*to divide*', '*to distribute*', '*to break*', or '*to separate*'. The semantic extension in this context sentence is produced by 15;00% of the bilingual participants, this is interpreted as an influence of the L2 meaning '*to leave'* instead of the L1 meaning '*to separate'*. no production of this kind was provided by the

monolingual group, instead both groups produced '*ir*' or '*irme*' as canonical response with 67,92% (monolinguals) and 80,00% (bilinguals).

4.4.8.9. Talón

Stimulus R28: *Dícese de un calzado femenino, generalmente son altos, puntiagudos y se utilizan en fiestas _____(Tacón vs Talón)*

Translation: *Let's say woman footwear, generally they're tall, pointy and used in parties _____(Heel vs Heals)*

In stimulus R38, this semantic extension can be considered as the result of both; formal and semantic overlap between French '*talón*' and '*tacón*' in Spanish.

In French, to refer to high heels, the composed word used is '*Chaussures à talon';* in Spanish the same idea is expressed by the word '*tacón*' or '*tacones*'. Besides, '*talón*' in Spanish refers to the anatomic part of the body, '*tacón*', instead refers to '*high heeled shoes'*. This semantic extension was listed by Quilis et al., (1982, p. 158) as "*talón—tacón (t a l o n talón y tacón) (e)*".

The same example could be associated to a literal translation or to translation loan. The percentages of responses match in both groups with 72,50% for '*tacones*' (bilingual group) and 67,92% (monolingual group) as it is shown in **Appendix 24**. However, '*talón*' is produced by 5,00% of the bilingual productions and NA for the monolingual group.

4.4.8.10. Sujeto

Stimulus R19: *No puedo mantener el hilo de la conversación, siempre cambió fácilmente de _____(Tema vs Sujeto)*

Translation: *I cannot keep the thread of the conversation; I always change readily of _____(Topic vs Subject)*

The semantic extension in stimulus R19 is triggered by the similarity of word form between '*Sujeto*' and '*Sujet*'. In this particular context, '*Sujeto*' in Spanish refers to a person or individual. In contrast, in French '*Sujet*" refers to '*a topic'*, '*a field of intellectual or artistic activity*' (CNRT, 2020). This semantic extension was defined as "*sujeto—asunto, tema (< s u j e t)*".

The canonical form '*tema*','*topic*' in English corresponded to 50,00% of the productions by the bilingual group and 43,40 % by the monolingual group as shown in **Appendix 24**. The misleading association between '*Sujeto*' and '*Sujet*' is produced by the bilingual group for 5,00% of the oral production and NA in the monolingual group.

4.4.8.11. Habitar

Stimulus R15: Durante la noche un robo ha sido notificado al lado de la casa en dónde Sara _____(Vive vs Habita)

Translation: During the night a robbery has been notified next to the house where Sara _____(Lives vs Inhabits)

In stimulus R15, the semantic extension is explained by the use of '*habitar*' in Spanish in the same manner of the French verb '*habiter'* instead of its Spanish equivalent '*vivir*' ('*to live'*, '*to stay*' or '*to remain*'). In Spanish the definition of '*habitar*' is to dwell or to inhabit something (RAE, 2020). The French counterpart '*habiter'* refers to the habitual residence of a place or location, '*to live*'. The use of '*habitar*' in Spanish is ambiguous and less frequent than '*habiter'* in French (see **Appendix 6**).

Canonical responses such as '*vivir*' has attained higher percentages by bilinguals with 67,08% and 67,50% by monolinguals. '*habitar*' is used for 2,50% by the bilingual group.

4.4.9. Extralinguistic Factors and the Production of Semantic Extensions

Table 34 describes the distribution of semantic extensions produced by the bilingual group. Additional data about the sample's linguistic background is reported as well: their LOR (length of residence), the percentages of L1 and L2 language use and L2 level following CEFR test or self-assessments. Some participants have tendency to produce more semantic extensions than others, we will overview briefly these participants for a better understanding on the production of semantic extensions.

The mean LOR of the participants is 4,69 years (minimum 1.0, maximum 11.0. As observed in the qualitative analysis 4.4.8, the production of semantic extension by the bilingual group may suggest that word choice is affected by the L2, after a mean of 4,69 year of residence in a L2 dominant context.

To evaluate language use, participants were told to complete a questionnaire (See questionnaire question 4A in **Appendix 1**) reporting for their L2 use in different domains and contexts using a Linkert scale (1 corresponding to rarely and 5 to frequent). The total of score of the questions were 70 (see **Appendix 26**). The distribution of L2 language use score for the subgroup of participants that have produce semantic extensions. The minimal value of L2 use is 0,37 and the maximum value of L2 use is 0,79.

The global L2 language use of this subgroup of participants have shown an L2 score mean of 44,07/70 giving a percentage of mean L2 use 0,63/1,0 (see **Table 34**).

Multiple information about the L2 level of the participant were compiled, although, in this section we will describe:

1) the CEFR (Common European Framework of Reference for languages) is an international standardized test includes all linguistic competences and abilities for second language learning and teaching levels, from beginners (A1, A2) intermediate (B1, B2) to advanced learners (C1, C2).

2) A self-assessment of L2 linguistic competence, including oral comprehension, oral production, written comprehension, and written production. Participant evaluated themselves as insufficient, regular, good, very good, native, or almost for each category. Each category. We have attributed values to those evaluation intervals from 0,25 to 1.

Global outcome of 1) has shown (see **Figure 33**) that the distribution of the CEFR levels comprises mostly intermediate (B2) for 15 participants and advance level (C1) for 5 participants. Basics levels are represented by 2 participants each (B1 and A2). 5 participants had no official CEFR test.

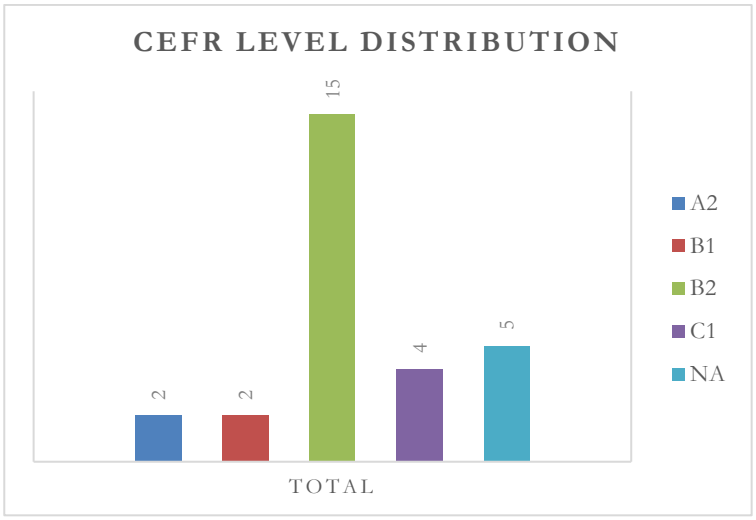

Figure 33 Distribution of CEFR levels of the participants that produced SE

The global outcome of 2) have shown a mean response of is 0,74 of a total score of 1 in **Table 34**. The minimal value is 0,60 and the maximum. 0,74 will be the equivalent of B2 level.

The objective of integrating 1) and 2) is that L2 official assessment test are rarely representative of the actual progress a SL learner has attained during an L2 dominant context or residency. Languages tests are commonly a requirement before entering the country or to the registration for study aboard programs. As learning is progressive and language test is a representation of the level attained at a certain period of time, it is interesting to compare both. We have analyzed 41 productions of semantic extension (see section 4.4.8) produced by 28 participants.

Following a qualitative approach, we propose to observe the participants who have tendency to produce more semantic extensions than others, we will briefly overview their bilingual profile and find common elements between them.

B2 level and self-assessment of 0,75/1 are common to participant 35, participant 6 and participant 13. Regarding the score of L2 language use which includes contexts of language and daily basis activities, participant 35 and participant 6 reported for a score of 0,64/1, contrary to participant 13 that reported for a score of 0,44.

Language level will function as an indicator of production of semantic extensions, however some participants with A2 (participant 38) and B1(participant 37) levels also produced semantic extensions. Nevertheless, their self-assessment indicated a high score of language use of 0,80/1 (participant 37) and 0,75/1 (participant 38). Their score of L2 use is 0,69 (participant 37) and 0,54 (participant 38) as shown in **Table 34**.

Now, following a statistical procedure including the total of the participants (40 participants) and their production (or not) of Semantic Extensions (SE) Pearson's correlations showed a significant correlation between the L2 level (CEFR) and the production of SE, r=-0,375, *, p<0,029* as showed in **Table 33**. The LoR in a L2 dominant context is not correlated to the production of semantic extension. r=-0,002 *, p=0,992*.

Table 33 Pearson's Correlation of the Bilingual Group and the Extralinguistic factors affecting the Production of SE

Pearson's Correlations

Variable		Participant	production of SE	L2 level test_7	L2 autoevaluation 1A	L2 language use 4A	LoR
1. Participant	Pearson's r	—					
	p-value	—					
2. production of SE	Pearson's r	0.260	—				
	p-value	0.105	—				
3. L2 level test_7	Pearson's r	-0.049	-0.375	—			
	p-value	0.785	0.029	—			
4. L2 autoevaluation 1A	Pearson's r	0.174	-0.126	0.296	—		
	p-value	0.282	0.440	0.090	—		
5. L2 language use 4A	Pearson's r	-0.011	-0.017	0.094	0.112	—	
	p-value	0.947	0.916	0.596	0.492	—	
6. LoR	Pearson's r	0.127	0.002	0.127	0.212	0.065	—
	p-value	0.435	0.992	0.475	0.188	0.690	—

Table 34 Productions of semantic extensions by the bilingual group

Participant	Stimuli	Tag	Production	Code	LOR	L2 use score	L2 use transformation	L2 CEFR Level	L2 self-assessment
1	R10.wav	Valor	coraje	4	3,5	26	0,37	B2	0,60
3	R30.wav	Dado	ofrecido	4	5	50	0,71	C1	0,80
5	R15.wav	Vive	habita	4	4	38	0,54	B2	0,70
6	R20.wav	Escucharle	comprender	*4*	3	45	0,64	B2	0,75
6	R42.wav	Amable	gentil	4	3	45	0,64	B2	0,75
7	R10.wav	Valor	coraje	4	4	58	0,83	B1	0,75
8	R17.wav	Discusión	disputa	4	6	49	0,70	NA	0,75
9	R19.wav	Tema	sujeto	4	5	44	0,63	NA	0,55
12	R30.wav	Dado	ofrecido	4	1,8	48	0,69	A2	0,65
13	R30.wav	Dado	ofrecido	4	6	31	0,44	B2	0,75
13	R10.wav	Valor	coraje	4	6	31	0,44	B2	0,75
14	R7.wav	Bolso	saco	4	5	51	0,73	B2	0,70
15	R27.wav	Irme	partir	4	3	29	0,41	NA	0,75
17	R10.wav	Valor	coraje	4	1	50	0,71	B2	0,70
18	R38.wav	Tacones	talón	4	5	40	0,57	NA	0,65
18	R10.wav	Valor	coraje	4	5	40	0,57	NA	0,65
19	R30.wav	Dado	ofrecido	4	6	47	0,67	B2	0,95

20	R27.wav	Irme	partir	4	1	51	0,73	B2	0,75
23	R27.wav	Irme	partir	4	5	39	0,56	B2	0,70
23	R10.wav	Valor	coraje	4	5	39	0,56	B2	0,70
25	R10.wav	Valor	coraje	4	2	51	0,73	B2	0,65
26	R17.wav	Discusión	disputa	4	7	46	0,66	B2	0,75
26	R10.wav	Valor	coraje	4	7	46	0,66	B2	0,75
29	R30.wav	Dado	ofrecido	4	1	53	0,76	C1	0,90
30	R42.wav	Amable	gentil	4	5	35	0,50	B2	0,55
30	R27.wav	Irme	partir	4	5	35	0,50	B2	0,55
33	R4.wav	Largo	carta	4	11	55	0,79	NA	0,80
33	R10.wav	Valor	coraje	4	11	55	0,79	NA	0,80
34	R7.wav	Bolso	saco	4	8	45	0,64	C1	0,70
35	R7.wav	Bolso	saco	4	3	46	0,66	B2	0,75
35	R17.wav	Discusion	disputa	4	3	46	0,66	B2	0,75
35	R41.wav	Grave	severas	4	3	46	0,66	B2	0,75
35	R27.wav	Irme	partir	4	3	46	0,66	B2	0,75
35	R10.wav	Valor	coraje	4	3	46	0,66	B2	0,75
36	R40.wav	Datos	informaciones	4	8	47	0,67	B2	0,80
37	R24.wav	Antigua	anciana	4	4	48	0,69	B1	0,80
37	R19.wav	Tema	sujeto	4	4	48	0,69	B1	0,80
38	R27.wav	Irme	partir	4	3	38	0,54	A2	0,75

38	R38.wav	Tacones	talon	4	3	38	0,54	A2	0,75
39	R10.wav	Valor	coraje	4	9	43	0,61	B2	0,70
40	R10.wav	coraje	coraje	4	4	43	0,61	C1	1
					4,64	44,07	0,63		0,74

4.4.10. CLI Analysis

CLI productions were coded as a single category (Code 5) however, this category refers to a multitude of CLI phenomena including loan translation or calques (described earlier), L2 borrowing and phonological transfer. 50 productions were coded as CLI, the percentages of each type of CLI are detailed in **Table 35** and show that more than half (56,00%) of the production CLI of the bilingual group corresponded to phonological transfer. For each case we describe briefly examples to illustrate the nature of the productions studied.

Table 35 Distribution of CLI in percentages

CLI	**%**
Blend	6,00%
Borrowing	10,00%
Loan translation	28,00%
Phonological transfer	56,00%
	100,00%

4.4.10.1. Word Blending

For word blending, we have found 2/3 productions which are produced by the same participant (participant 14), in this section we will describe the nature of these two productions and overview the bilingual profile of this participant.

The first production is '*escarpines'* as shown in Table 36, a word resulting from the blend of French '*escarpin*' and the corresponding Spanish plural morpheme '*es*'. This blend was produced instead of the target '*tacón'* (high heeled shoes) described in 4.4.8.9. This production is interpreted as the influence of the L2 over the L1 due to an effect of language use and to the availability of word frequency of French in the bilingual mental lexicon. However, this particular production may also be interpreted as a L2 borrowing with morphological adaptation.

The second production is '*siñalada*' instead of '*señalada*' which resulted from the blending of verbs '*signaler*' in French and '*señalar*' in Spanish (*to point out* in both languages). In stimulus:

Stimulus R33: *La grabación dice que la llamada no puede ser transferida, la tecla correcta no ha sido _____(Siñalada ** vs Señalada)*

Translation: *the recording says that the call phone can't be transferred, the right key wasn't_____(Pressed)*

Recall that this stimulus was included in the statistical model, despite the fact that the target production 'pressed' was rarely produced by only 7,50% by the bilingual group

and 0,00% by the monolingual group (see **Appendix 24**). However, the stimulus gave rise to interesting CLI productions representing 5,00% of the answers, which justified its inclusion in the model. The semantic extension that we expected for this stimulus is '*tapar*' ('*to cover*') for which the extended French meaning *taper* is *typing*. However, no production of this type was observed.

The third production '*enregistrar*' is a blending between the Spanish '*registrar*' and the French '*enregistrer*' whose semantic reference is *to record*.

Table 36 Detailed list of blending productions

Tag	Stimulus	CLI	Participant	G2	Other responses%
R7	Dícese de un calzado femenino, generalmente son altos, puntiagudos y se utilizan en fiestas	escarpines	14	2,50%	95,00%
R33	La grabación dice que la llamada no puede ser transferida, la tecla correcta no ha sido _____	siñalada	14	2,50%	95,00%
		enregistrada	25	2,50%	95,00%

The bilingual profile of participant 14 who produced 2/3 of blending had 5 years of LOR, B2 level, and her self-assessment was 0,70/1 at the moment of language test. The Questionnaire total score of language use was 0,75/01 for French against 0,25 for Spanish.

In the sociolinguist questionnaire section describing L1 and L2 activities participant 14 affirmed that French was the language used in daily activities such as reading, writing, and having formal and informal discussion and hobbies. In contrast, Spanish was only used during informal discussions. (**Appendix 27**).

Additionally, participants were asked to evaluate themselves using percentages describing the amount of French and Spanish they used during a regular day. Participant 14 described an amount of Spanish of 25% against 75% for French when we asked whether participant 14 would consider that her/his Spanish had suffered changes she responded positively. She described it as *'Olvido de palabras, conjugaciones, confundir las dos lenguas'* otherwise: '*the loss of words, conjugation and the confusion of the two languages*. These responses are consistent with the productions found: the blended words in L1.

The fixation of linguistic knowledge such as reading and writing are highly important to prevent language change and maintaining the L1, the permanent use of L2 as reported by participant 14 is certainly linked to the fact that her L1 production is highly influenced by the L2 (**Appendix 28)**.

4.4.10.2. L2 Borrowing

In our data we have coded 10,00% of the productions in L1 as L2 borrowing. In some cases, the productions were phonological close to the L2 and in other cases they were adapted into the phonological rules of the L1. **Table 37** describes the total occurrences of L2 borrowing, we will discuss some of these productions to illustrate the focus of the analysis that we have made.

For R38, the 3 productions of '*nécessaire*' instead of '*bolso'* are an L2 borrowing of the word '*nécessaire*' in French. Following the CNRTL (2020) '*nécessaire*' refers to a box, case, or a kit, that contains a set of objects that are '*necessary*' for an activity or an occupation, for example: '*nécessaire de toilette'* (toilet kit) and '*nécessaire de couture*' (sewing kit). The L2 borrowing of this word is produced by 3 participants with different degrees of approximative realizations with respect to the L2 phonetic system.

Stimulus R7 was already analyzed in previous sections, whose responses were described in Table 25, the semantic extensions were analyzed in section 4.4.8.4. Besides, in Table 36 we have described a production of R7 exemplifying word blending. In this case, another production of '*Esparpin'* is analyzed as an L2 lexical borrowing, in which the word form and the phonological realization is entirely "borrowed" from the L2.

Table 37 Detailed list of L2 borrowing Productions

Tag	Stimulus	CLI	Participant	G2	Other responses%
R38	Para cargar todos sus cosméticos, sus llaves y su teléfono ; Anna se compró un pequeño_____	nécessaire	8	2,50%	95,00%
		nécessaire	19	2,50%	95,00%
		nécessaire	37	2,50%	95,00%
R7	Dícese de un calzado femenino, generalmente son altos, puntiagudos y se utilizan en fiestas	escarpin	13	2,50%	95,00%
R18	Juan ha perdido su billetera, y no puede efectuar ningún pago, puesto que ha bloqueado todas sus _____.	carte	4	2,50%	95,00%

For stimulus R18 the use of '*carte*' (card) instead of the Spanish equivalent '*tarjeta*' is also an entire phonological L2 lexical borrowing, i.e., the word '*carte*' was realized in French.

4.4.10.3. Loan Translation

*Stimulus R14: Fumar es un vicio muy malo y ya no quiero hacerlo; por tres días voy a intentar _____pararlo***

Translation: Smoking is a very bad vice and I don't want to do it anymore, for three days I'm going to try to_____ quit

In this production we have found a loan translation emerging from the L2 (French) into the L1 (Spanish). This semantic transfer is due to the uncanonical production '*parar*', a literal translation of French '*arrêter*' or to quit (this particular linguistic context is related to quit smoking). '*Parar* means '*to prevent, block'* or *'stop a movement or somebody's action', 'to prevent', 'to prepare*,' and other referents following RAE (2020), however in any case '*Parar'* matches with the concept of '*arrêter*' on French (or to quit).

'*Arrêter*' has multiples meanings in common with its Spanish counterpart, such as '*to stop'* and *'to block'*. Nonetheless, the sense applied in this context is '*to restrict'* or *'limit the consuming of cigarettes'*. Following the *Modified Hierarchical Model* (Pavlenko, 2009), we can assume that the shared meaning features between languages will trigger a negative transfer of '*Parar*'. This negative influence is mostly observed by the bilingual groups with 22,50% productions of '*Parar*', this suggests strongly linguistic traffic from the L2 to the L1.

The control group produced nonetheless also 3,77% of productions of this kind, that we may relate to the influence of their L2 English (this interference is originated from the structure: '*stop smoking*'). Sociolinguistic data have shown that all of these participants have used and studied English as an L2.

The canonic use of '*dejar*' matches for 50,00% of the responses of the bilingual group and 52,83% in the monolingual group. This confirms '*dejar*' in this particular linguistic context, an accurate response candidate to in stimulus R14 (see **Table 38**).

Table 38 Responses for stimulus R14

Tag	Stimulus	Response	G1	G2	Other responses G1	Other responses G2
R14	Fumar es un vicio muy malo y ya no quiero hacerlo; por tres días voy a intentar _____	Dejar (lo)	52,83%	50,00 %	43,40%	27,50%
		Parar (lo)	3,77%	22,50 %		

4.4.10.4. Phonological Transfer

Phonological transfer represents half of the CLI production (28 productions) in the bilingual group. The coding of this category was confirmed by different judges that were experts of the L1 and L2 phonetic features. In some cases, it was not possible to decide whether the pronunciation was conformed to L1 or L2 since there are not always clear-cut boundaries between phonemes realizations. However, this information is not going to be reported here because acoustic analyses must be made in future research. In **Table 39** we propose to overview the phonological transfer (see section **Error! Reference source not found.**) that we found in our bilingual sample (included in CLI category productions).

Additional participant information is associated to the productions when participants have tendency to produce L1 phonological transfer more than once.

Table 39 Percentages of L2 phonological influence on the L1

Stimuli	Target word	Productions	% of 28
R10	Valor	[v]alor-[v]alentia	21,43%
R15	Vive	[v]ibe-[v]i[v]e	21,43%
R17	Discusion	Discu [sjõ]	14,29%
R6	Regresar	[v]ol[v]er-[v]olber	10,71%
R29	Conocer	Descu[v]rir /[v]isitar	7,14%
R33	Caminar	Camina[ʁ]	3,57%
R35	Suceder	Ocurri[ʁ]	3,57%
R36	Probarlo	[v]isitarla	3,57%
R27	Irme	I[ʁ]me	3,57%
R40	Datos	informa[sjõ]	3,57%
R5	Monedas	Tar-je-tás	3,57%
R3	Presionada	Acti[v]ada	3,57%
			99,99%

As it can be noted in Table 39, phonological transfer includes in most cases the realization of the phoneme [v] in Spanish which is never realized (in the European and South American varieties) as a voiced bilabial fricative but as a voiced explosive bilabial [b](Antonio Quilis, 1981). The second most produced realization is [sjõ] which in Spanish does not correspond to the standard pronunciation of the suffix *sion*. Here, the explosive nasal [n] in post nuclear position (after the vowel) is realized as voiced alveolar consonant, which is not the case for French. In French oppositions between nasals vowels and nasal consonants can be made, e.g., banc [bɑ̃], bon [bɔ̃]). Such distinctions between nasals and oral vowels are not pertinent in Spanish, unless the vowel is placed between two consonants [mano] in '*mano*' or in an initial position [ĩn.sa.ˈsja.βle] in '*insaciable*' (Antonio Quilis, 1981, p. 186).

The realization of [r] or [ɾ] by some bilinguals is not canonical and it is closer to realization of the phoneme [ʁ] in L2, especially in a final position and rarely in an initial position. The Spanish r involves two variants: simple (e.g., [ka.ɾo] for '*caro*' and multiple vibrating, e.g., [ka.ro] for '*carro*'. In French [ʁ] is fricative, vibrating and occlusive. In Spanish [r] is liquid, occlusive, and vibrating (Quilis, 1981).

Regarding stress, Spanish words are accentuated depending on orthographic rules and can be classified into their classes (e.g. substantives, adjectives, pronouns) or unaccentuated (e.g. prepositions, determined articles, conjunctions). This stress in some cases is indicated or not by a tilde (e.g, '*número*', '*numero*' and '*numeró*'), lexical stress, the changes the meaning of these words (Spanish is traditionally called a *free stress language*). Inversely, in French, lexical stress is supposed to be placed at last syllable (traditionally called a *fixed stress language)*, however French stress system is much more complex (for further discussion see Astésano (2001).

In Spanish words are categorized differently depending on the stress syllabic order and classified into (Oxytones, Paroxytones, Proparoxytone). In our data, we have found

the production of '*tar.**je**.ta*' [taʁ.ˈxɛ.ta] in which the stress syllable should be the penultimate syllable (i.e., *.je.*), also called Paroxytone. In this production the participant had displaced the stress into the ultimate syllable given as result an uncanonical accentuation of the L1 as in French, in '*tar.je.**ta***'. [taʁ. xɛ. ˈta].

Regarding closer, the participants having produced the most phonological transfers (see **Appendix 28**), common variables of this subgroup can be found. The levels attested in their L2 and their L2 self-assessment is between B2-C1 (for the L2 CEFR test) and self-assessment scores of 0,75 - 0,90/1 (see **Appendix 28**).

4.4.11. Discussion

As detailed in **Table 29**, through the Gap Completion Task we analyzed 3,00% of sematic extensions produced by the bilingual group (41 productions of sematic extensions spread in 11/41 stimuli, which were described in section 4.4.8). Regarding CLI productions, and CLI represented 3,66% of the bilinguals' productions (50 productions) spread across different phenomena) as detailed in **Table 35**. Regarding the Extralinguistic variables studied along with the production of SE, it seems that L2 CEFR levels seem to be correlated as observed in section 4.4.9.

We have hypothesized that in the Gap Completion Task, bilinguals will produce a certain amount semantic extension in L1 since the L2 is activated while the L1 is processed and produced. This was partially the case (**Table 24**). However, a more predominant phenomenon was CLI along with 50 occurrences. Half of these production (56,00%) involved phonological influence of L2 on production of L1 words.

As discussed by (Roelofs, 2003) the access of words during production has been neglected in the literature, he highlights that even though morphemes differ between languages, phonemic segments (such as vowels and consonants) are in some cases shared between languages. This is the case in our study with the L1 and the L2 sharing multiple phonetic features. Roelofs' studies suggest that during oral production and discourse planning of bilinguals, phonological shape of utterances build on shared representations for both languages. In our study, the observation of L2 phonological productions suggests that shared phonemic segments have been integrated or extended in the L1. Common patterns are then observed such as the realizations of [v] and [b], the realization of [ʁ] [] instead of [ɾ] nor [ɾ] , some nasals in the suffix [sjõ] such as in '*discussion*'.

Concerning the RTs analysis developed in 4.4.5.1, the results of the Linear Mixed Effects Models show that plausible responses took longer than canonical responses for both samples. Regarding the quality of the responses obtained, bilinguals and monolinguals behave similarly, however, CLI productions were more frequently observed in the bilingual group. CLI seems to facilitate (as shown by faster responses) in bilinguals compared to monolinguals. The interference of the L2 during L1 production indicates that the L2 is highly activated during L1 production. Hence, the advantages of being bilinguals involve being able to use more languages as resources to resolve linguistic

constraint tasks such as the Gap Completion Task. The enrichment of the mental lexicon that is shared between two or more languages will allow the bilingual to respond faster using the most available resources which do not always match to the L1, in other words, L2 interference works as a 'gaining time' strategy to handle linguistic constraints.

4.4.12. Limits and Further analysis

The construction of this particular oral elicitation task was very challenging because of the multiple constraints linked to oral production, namely achieving to create a sufficiently constraint context to induce the spontaneous production of a very specific target response and avoid multiple interpretations of meanings by participants. The first analyses showed that not all sentences contained enough contextual information in order to limit the responses, this is why we had to reduce the number of stimuli before applying the statistical model. Despite the difficulties considered, the analysis of the oral responses in terms of their quality and RT provided valuable information about bilingual lexical access and word selection. Since oral production is an online process, it will allow us at a later stage to overview the stages of planning oral output, such as hesitation marks and word search during production.

Among the CLI productions that we have found (see **Table 35**), the other two frequent phenomena reported were phonetic transfer and loan translation. These observations are an indicator of the direction of further studies: acoustic analyses of L1 targeted the realizations of the phonemes ([v], [ʁ],[sjõ]) by bilinguals and non-immersed L1 speakers.

The Gap Completion Task is an alternative proposition that differs from the traditional object naming, fluency / accuracy tasks, that have tendency to shape the production of the samples and minimize CLI phenomena that can be detected during less controlled oral production. CLI data from oral productions is certainly a valuable resource for a better understanding of bilingual lexical access. We were aware that the risk of this type of experimental design and data analysis, however, it was worthy of interest as showed by our qualitative analysis and our statistical modelling.

5. GENERAL DISCUSSION

5.1. OVERVIEW OF THE FINDINGS

We have implemented three experiments to study, from different perspectives, the case of Semantic Extension (SE) as the result of L2 influence on the L1 in late bilinguals, immersed in an L2 dominant context (France). In experiment 1, a *Gap Completion Task* (CGT) allowed us to analyze the *Production* of SE. In experiment 2, a *Lexical Decision Task* (LDT) was used to focus on bilinguals' early stages of language processing, specifically, *Word Recognition Processes* of SE. In experiment 3, using an *Acceptability Judgment Task* (AJT), we analyzed later stages of language processing involving *Comprehension* and *Metalinguistic Knowledge* that concerned not only the L1 but also the L2, i.e. through the evaluation of sentences involving SE.

We adapted the linguistic material to the constraints of each task and manipulated internal characteristics of the stimuli presented (Dominant Neighborhood Density and Dominant Morphological Family Size in either L1 or L2) testing and comparing bilinguals and monolinguals in their L1. Experiment 2 (LDT) and 3 (AJT) had in common the opposition of L2 or L1 Dominant Morphological Family Size and Dominant Neighborhood Density. However, the stimuli were presented differently: in experiment 2 (LDT) we presented stimuli in the form of isolated words associated with different primes while in experiment 3 (AJT) the stimuli were presented in context (i.e. within a sentence). We will discuss the results obtained from each experiment separately and then jointly.

The Analysis of semi-controlled *Oral Productions* of bilinguals and monolinguals in Experiment 1 (GCT) allowed us to conclude that the quality of the oral responses obtained by the participants has shown few differences between the bilingual and monolingual participants with respect to errors or canonical and plausible responses. However, some uncanonical productions including semantic extensions, and other CLI productions were observed in the bilingual group. We have hypothesized the production of a certain amount of Semantic Extensions in group 2 (Bilinguals) as the result of L2 activation during L1 processing and oral production:

H3: *In the Gap Completion task, bilinguals will produce L1 semantic extensions since the L2 is activated while the L1 is processed and produced.*

Results have shown that SE were produced by the bilingual group in some cases (41 productions) corresponding to 3% of the total productions detailed in section 4.4.8, which is consistent with studies in language attrition in which deviant productions rarely surpass 5% of the collected data (e.g. Schmid & Köpke, 2017). Surprisingly, other types of CLI were collected as well (51 productions) corresponding to 3,66%. The overall productions recorded and analyzed suggest that L2 is co-activated at the lexeme level since a number of coinage of words, loan translations and cases of phonological transfer were observed, but also at the lemma level, as shown through the production of some SE. We highlight that a reduced LoR of 4,63 years in a L2 context seems to be enough for late bilinguals (non-attriters) to produce uncanonical lexical choices that are less

restrictive (i.e., SE or phonological transfer) in relation to more targeted structures such as grammatical structures.

With respect to the RTs measured in (GCT) task, both groups (Bilinguals and Monolinguals) were slowed down to respond to plausible than canonical responses, but interestingly, for bilinguals only, RTs were faster for responses showing CLI implying the L2 was activated during the processing and the production in the L1. These results could be explained in at least two ways:

- The facilitatory effect of CLI productions might be explained by the view that co-activation is the most efficient (and therefore the fastest) mode of production for bilinguals, as advocated by Grosjean (2018) in the bilingual mode framework.
- On the contrary, the fact that bilinguals were slower than monolinguals to produce canonical and plausible responses might be due to the cost involved by preventing co-activation, here, two languages may compete to generate accurate responses.

Regarding the extralinguistic factors analyzed, only CERF levels seem to correlate to the number of productions of SE, in accordance to the position that high proficiency levels are necessary for semantic transfer to take place (e.g. Lindqvist, 2012), However, we cannot assume that lemmatical transfer is entirely due to misleading associations at a given moment in L2 learning. Recall that in the past, semantic transfer has been found in the production of L2 users regardless of their language level (e.g. Pavlenko & Jarvis (2002). We defend that competition levels between across languages are underlying the production of SE instead of fully meaning-based transfer.

Regarding *Word Recognition Processes* in Experiment 2, the Lexical Decision Task has shown that bilinguals were slower than monolinguals at processing words presenting L2 Dominant Neighborhood Density and L2 Dominant Morphological Family Size. This result suggests that bilinguals were sensitive to L2 dominant features in comparison to the non-immersed L1 group, probably due to:

- multiple linguistic resources available for the bilingual stemming from more than one language;
- differences in the way bilinguals use and process the L1 and L2 following the concept of multicompetence (Cook, 2003, 2013).

Bilinguals' RTs were more delayed by L2 Dominant Neighborhood Density than L2 Dominant Morphological Family Size, suggesting that formal features of the L2 were strongly coactivated during the task performed in L1 (Thierry & Sanoudaki, 2012).

We had predicted word processing of SE (Condition 3) to be affected by L2 Dominant categories:

H4: *In the Lexical Decision Task, bilinguals will process semantic extensions ~~d~~ (as shown by RTs) depending on accumulated sources of activation triggered by the Neighborhood Density and the Morphological Family Size (in the L1 and the L2).*

However, no interactions of the variation of latencies rates were detected by the bilingual group for condition 3 in which we presented SE associations (e.g. '*carta*' - '*tarjeta*'). In contrast, we have found a facilitatory effect for L1 Dominant Morphological Family Size for all groups merged, showing a significant interaction for condition 2 which represents L1 canonical meaning associations (e.g. '*exprimir*'- '*apretar*'). In our opinion, this shows that L1 lexical networks between lemmas and concepts are strongly rooted due to language use in both groups, as manifested through facilitatory effect across groups for L1 overlapping meaning condition. This finding points towards strong entrenchment (Steinkrauss & Schmid, 2017) of the L1 in monolinguals and late bilinguals as well.

The main finding of the LDT suggests that during word recognition processes, L2 Dominant Neighborhood produced a facilitatory effect for bilinguals when targets and prime associations were presented in condition 1 (e.g., '*exprimir*' - '*reprimir*'). This effect was shown through a three-way interaction between Dom_N in condition 1 which was exhibited exclusively by the bilingual group. As a result, L2 levels of activation of multiple neighbors will play an important role during early processing of L1 stimuli, allowing us to conclude that shared lexical features (i.e. phonological and orthographical features) between L1 and L2 create a '*resonance*' in the activation levels across languages (Kroll et al., 2010).

This finding confirms our theoretical positioning in which we propose to study semantic extensions as the result of formal competition between languages and not of crosslinguistic learned associations between languages (Jarvis, 2009). As we predicted, the facilitatory effect between language is triggered by word-form overlap instead of meaning overlap.

Regarding *Metalinguistic skills* tested in Experiment 3 (AJT), the analysis of the results points out that bilinguals show a tendency to evaluate sentences in the three conditions as more acceptable than the monolingual group. A closer look at the psycholinguistic factors at play, i.e. L2 Neighborhood Density and L2 Morphological Family Size, does not reveal any significant effect nor interaction. In contrast, L1 Dominant Neighborhood Density also affects bilinguals' performance, in this case through increasing the acceptance levels of the sentences in comparison to the monolingual group, and independently of the condition presented.

We had predicted a variation of the acceptance levels across group depending on the condition presented:

H5: *In the Acceptability Judgment Task, bilinguals will be more flexible in their evaluation of L1 sentences that express a semantic extension. We hypothesize that bilinguals may accept to a higher degree and maybe reject the corrected meaning of semantic extensions (Condition 3), the type of sentence accepted (condition 3 and 2) will indicate whether semantic or conceptual restructuring is rooted in the L1.*

However, no condition effect was reported involving exclusively the bilingual group, instead, both groups mingled (bilinguals and monolinguals) have shown differences in their evaluation rates between condition 1 (involving the use of semantic extensions) and condition 2 sentences (involving the use of L1 canonical meaning) when

Neighborhood Density was dominant in Spanish. This result suggests that, in this case, L1 functions a positive evidence (Seliger, 1991) intervening in the way the participants evaluated sentences. Thus, Spanish Dominant Neighborhood sentences in condition 1 were less accepted than condition 2 for all participants mingled. This may be due to the positive evidence that had reinforced Spanish Dominant target stimuli.

In the AJT, bilinguals behave as monolinguals and we can suppose that a monolingual mode was adopted, in which the bilinguals managed to deactivate one of their languages to avoid linguistic interference, despite the fact that this "*deactivation*" is rarely total in the bilingual mind (Grosjean, 2018).

In conclusion, the implementation of the experiments has shown that bilinguals produce semantic extensions and other types of CLI during L1 production. Moreover, Neighborhood Density and Morphological Family Size dominances plays distinct role in early and late language comprehension processes. For the LDT, we have found evidence that L2 Neighborhood Density plays a facilitatory role for word recognition processes in early recognition when L2 of word-form overlap conditions in the bilingual group, while for the AJT both groups were sensitive to L1 Neighborhood Density while evaluation sentences L1. In the LDT, L1 Morphological Family Size seems to play a facilitatory role during word recognition in both groups when canonical semantic overlap in the L1 is presented. We hypothesize that L2 interference during oral production and L2 facilitation for late bilinguals is consistent with the idea that lexical access is non-specific for language (La Heij, 2005), implying that the activation flow is spread to the lexical nodes of both languages, so L2 is highly activated during the processing and production of L1 linguistic material.

5.2. Word-Form overlap vs. Semantic Overlap

We recall that in this study the inclusion of Neighborhood Density and Morphological Family Size as psycholinguistic parameters was aimed at allowing us to analyze Semantic Extensions through two levels: word-form overlap (representing the lexeme level) and semantic overlap (representing the lemma level). Crosslinguistic networks in the bilingual lexicon were observed through the opposition of Dominant Stimuli in L1 or L2 representing semantic links and formal links across languages. The two psycholinguistic parameters included in our experimental design were hardly taken into account in previous modelling of bilingual lexical access, they are also viewed as an alternative to palliate difficulties that may be inherent to the study of the lexicon. The challenge of investigating the lexicon is related to the idea that it involves a complex combination of features (e.g. semantic constraints, specific context of language use, etc.).

The results of Experiment 2 (LDT) suggest that bilinguals were sensitive to L2 Neighborhood Density Stimuli. Accordingly, it seems that lexical access is facilitated by word-form overlap across languages when Neighborhood Density is dominant in L2, at least during early stages of word recognition (for condition 1, representing word-form

overlap). In experiment 3 (AJT), L2 Neighborhood Density did not display a significant role, while L1 Neighborhood Density did play a role during comprehension and later stages of word recognition, affecting the way participants evaluated the sentences presented (between condition 1 — expressing semantic extensions — and condition 2 — presenting canonical sentences). In other words, it seems that the role played by Neighborhood Density Dominance changes depending on the experimental task performed and the underlying processes involved. While in the L1, Neighborhood Density affected later stages of languages processing for all participants, in L2, Neighborhood Density would affect bilinguals only, facilitating lexical access in early stages of word-form recognition due to crosslinguistic overlap between the languages. This result is in line with Costa et al.'s (2006) proposition in which they propose that neighborhood density effects across languages facilitate lexical access in one of the languages whilst for words with few neighbors such benefits are smaller.

Regarding L1 Morphological Family Size in Experiment 1 (LDT), a facilitation effect was observed for all participants when semantic overlap was reinforced through condition 2 (i.e., expressing a L1 canonical semantic relationship) suggesting that the groups were highly sensitive to intralingual semantic links since participants shared the same L1. This confirmed the role of language use on language entrenchment (Steinkrauss & Schmid, 2017). Following De Jong et al., (2000) stronger semantic activation increases with high morphological family size of the targets, which was the case in our study exclusively in the L1. In contrast, L1 Morphological Family Size in Experiment 2 (AJT) did not display significant interactions in neither group.

We can conclude that there is not one psycholinguistic parameter to be more important than the other, instead a combination of the two factors affected the way bilinguals and monolinguals processed lexical information. In the case of semantic extensions, neither the conditions representing uncanonical priming associations representing semantic extensions (LDT), nor presenting uncanonical sentences representing semantic extensions (AJT) allowed us to observe significant differences between the groups. This suggest that Semantic Extension's conditions were not processed differently in neither group, however, L2 dominance influenced the bilingual groups' performance indicating that L2 formal features do intervene during early stages of word recognition in L1.

Taking into consideration previous modeling of the bilingual lexicon and according to our results, we agree that taking into account L2 transfer as a basis of a restructuring and learning processes (i.e. in Modified Hierarchical Model, Pavlenko 2009) is valuable for the understanding of semantic transfer. In our study, semantic specificities in the L2 appear to be "shaped" into the L1, which affects linking processes between lexical and conceptual representations when bilinguals produced semantic extensions.

5.2.1. SE as the result of formal competition

In previous research, Semantic Extensions were considered to be the result of learned crosslinguistic associations processes instead of the result of activation level processes (Jarvis, 2009). This assumption implies that during semantic transfer, the linguistic knowledge that is rallied during transfer concerns exclusively the lemma level (i.e. lemma to lemma associations) and excludes the lexeme level (i.e. orthographic and phonological features). Our study proposed to extent this narrowed perspective to one including lexemic features as an explanatory framework for semantic extension, in which also competition between lexemes plays a role, particularly when confronting perceived similarities across languages.

As shown by the LDT, L2 Dominant Neighborhood plays a facilitatory role during early stages of word recognition for bilinguals, in other words, co-activation of shared phonological and orthographical features will produce a priming effect when targets overlap in terms of perceived similarities (in condition 1 '*exprimir*' - '*reprimir*'). This effect is not observed for the monolingual group, suggesting that for bilinguals, L2 Dominant Neighborhood processing was mainly motivated at the lexemic level.

This effect was not observed for condition 3 (expressing uncanonical associations between primes and targets) which was not significant in comparison to the other conditions, nor across groups. However, we must highlight that word-form competition was affected during the processing of L2 Dominant Neighborhood Stimuli.

The sensitivity to L2 Dominant Neighborhood categories produced by the bilingual group supports our proposition in which formal overlap plays a role in bilingual processing due to the determinant role of competition levels between languages. Here L2 features indicate a strong activation in comparison to the L1, we interpret that lexeme levels are co-activated facilitating early word recognition processes, confirming our hypothesis according to which formal overlaps play determinant role at explaining SE. Additionally, we can imagine that the constant positive evidence (Seliger, 1991) provided in a L2 dominant context (Seliger, 1991)may also plays a role, resulting in formal features (L2 Dominant Neighborhood) that are probably more accessible in the L2 than L1.

Moreover, some of the production of SE can be interpreted, as a sign of cognitive restructuring, i.e. when the bilingual accommodates semantic representations into those of the dominant linguistic community (L2). In oral productions (e.g., stimuli in section 4.4.8.8), L1 word choice may be affected by the availability of the L2 lexicon or be triggered by formal crosslinguistic overlaps.

One of our contributions to bilingual research is the analysis of semantic extensions, which has been rarely studied in the context of late bilinguals. The originality of studying SE from different perspectives in production, in early and late stages of language processing through different experimental tasks, allowing us to oppose L1 and L2 dominances (representing lemmatic and lexemic level). Additionally, we presented our linguistic material using different associations (conditions) also opposing word-form and

semantic-relations. Sensibility to the L2 suggest SE not only as the result of crosslinguistic learned associations but also as the result of lexical competition levels.

5.3. THE L1 AS A FLEXIBLE LANGUAGE SYSTEM

Our perspective of bilingualism involved the view of a dynamic lexicon in which both languages interact constantly at different linguistic levels and this independently of the type of bilinguals under investigation, i.e., late or early bilinguals. Both will equally experience linguistic traffic and changes, these changes involve the way bilinguals use both languages L2 and L1 (Cook, 2003).

Extralinguistic factors such as L2 level, L1 and L2 contact and use, the linguistic environment and exposure to the language (e.g., immersion) will certainly influence the way languages interact, especially regarding L1 and L2 lexical accessibility. However, our data were supposed to explain such effects, instead we observed L2 interference during oral production and a sensibility to L2 Dominant Neighborhood during early word recognition in the bilingual group.

In Experiment 1 we studied bilingual production through the analysis of semi-controlled L1 oral production. Recall that word choice is a paradigmatic category intervening in CLI in which a lexical unit is selected among multiple lexical competitors (Jarvis, 2002). The interest of this experiment was to observe the role played by lexical accessibility of L2. Results suggest that the L1 lexicon is permeable to L2 intrusions manifested by diverse types of CLI (3,66% of the total production) including the production of phonological transfer, loan translation and borrowing. More surprisingly, it seems that when bilinguals produce a 'deviant' word choice, here considered as CLI, their RTs show a facilitation effect, suggesting that producing CLI makes them respond faster compared to monolinguals and other types of responses. Sociolinguistic data suggests that the production of CLI may not depend on a specific extralinguistic factor but on a combination of several factors implying that there are still a lot of unknown aspects regarding the role of extralinguistic factors in L2 interference.

Regarding semantic extension, we collected only a few productions (3,00% of the total production) which we analyzed one by one. The CEFR level in L2 of the bilinguals seems to be correlated to the number of productions of SE, suggesting that a certain language level is probably necessary for lemmatical transfer to occur. From this perspective, semantic extensions may be the result of the strengthening of semantic networks due to language mastery and use.

As it can be seen, bilingualism can be viewed as a genetic chain that is unique depending on multiple configurations and on the subject in question. We consider languages to be dynamic and evolving constantly rather than stationary. "Deviant productions" and L2 sensibility can be considered here as evidence supporting that the L1 is flexible enough to change (in terms of multicompetence and language accessibility)

even for late bilinguals whose L1 is traditionally viewed as fully acquired and a static end-state.

CONCLUSION

5.4. IMPLICATIONS FOR FUTURE RESEARCH AND PERSPECTIVES

An original configuration in this research was to focus on subtle changes that go unnoticed in current use, such as semantic extensions. This phenomenon is considered valuable for bilingual research because it allows us to study the lexicon following an analysis centered on individual patterns of L1 language change instead of changes experienced by second or third generation immigrants. The study of the very first signs of language change (i.e. CLI) during L2 immersion opens a research path for a better understanding of more pronounced manifestations of language change such as attrition. Similar research proposed by Lynch (2017), comparing heritage language speakers and L2 speakers, suggests that L2 acquisition and L1 attrition are closely related and considered as part of the same continuum.

From this perspective, L1 attrition and CLI (including semantic extensions, phonetic transfer, loanword and coinage of words) are probably triggered by the same external conditions such as reduced use of the L1 and changes in the linguistic environment or immersion. Here, L1 inaccessibility constituted an explanatory framework that would work for both cases. Future research is meant to support the relation between CLI and attrition as a result of SLA, an upcoming research target to compare bilinguals in their L1 (Spanish) to L2 learners of Spanish in order to find similarities of language use at the beginning of SL learning and as the result of bilingualism (see Schmid & Köpke 2017). A parallel can be made between the deviant production of L2 language users during early of learning in which formal associations coming from the L1 may result in L2 semantic extensions. In the same line, immersed bilinguals may produce the same associations in their L1, here both phenomena are motivated by the same underlying phenomenon, implying that language restructuring and language attrition are inherent to acquisition processes.

As similar case was observed during experiment GCT, when the order of tasks presented was inverted (LDT, CGT, AJT) in the control group. This change has triggered the production of SE by monolinguals resulting in a little number of SE in comparison to those monolinguals where experimental order was respected (GCT, LDT, AJT). Participants who began by recognizing SE in canonical and uncanonical associations were more prone to produce those target words in the oral production task, suggesting that uncanonical association can be easily triggered in monolinguals. As similar example was found by Dewaele (2018) regarding the instability observed between the variety of English used by Americans living in the UK, who had changed or restructured their use of emotion-laden words. From this perspective, the role of L2 immersion on the production of CLI cannot be neglected.

Another aspect that is worth mentioning is related to the challenges related to the analysis of sociolinguistic data in the bilingual group. In order to explain CLI productions,

we have taken into account sociolinguistic data to try to link common characteristics of the subject producing the same type of CLI (e.g. the case of word blend is viewed in 1.2.1.2.3.). However, exploring such data is challenging due the individual variability of the groups studied, for example, each participant may have used a different strategy for solving the same problem of lexical access. Nevertheless, we defend the idea that sociolinguistic data constitute a valuable key for a better understanding of CLI phenomena, for example an L2 learner who is conscious of his/her lack of linguistic knowledge may be particularly inclined to produce coinage of words. In the same way we can imagine, that L2 learners who mastered the best L2 phonetic system. may be more incline to produce CLI of phonological transfer.

Even though ordinal data are hard to code and to analyze, it is worth to create new methodological approaches will facilitate not only questionnaire construction but also coding and analysis. Sometimes the calculation of means or medians is not representative of the specific profiles within the sample studied, the transformation of ordinal data into numeric is highly questionable since it minimizes valuable information about the participants studied. From this perspective future research should certainly focus on new ways of collecting and analyzing sociolinguistic data that define the way bilinguals and monolinguals behave facing language experimentation.

5.5. FINAL REMARKS

One of the risks and at the same time a novelty of our study was to test bilinguals in their L1. We considered necessary to compare both groups in the same language instead of focusing on language switching tasks, which from our perspective would not be representative of crosslinguistic influence of the L2 on the L1, but a task triggered effect. In this study, the sensitivity to L2 dominant stimuli (which was presented through an L1 equivalent) and during L1 production, the production of CLI observed by the bilingual group are interpreted as the result of processing two languages trough the same mechanisms (Köpke & Keijzer, 2019).

Concerning the multicompetence framework, late bilinguals have displayed L1 changes evidenced under the influence of the L2 at different levels: oral production and word-processing. Languages can be seen as dynamic systems that are affected bidirectionally, in the same way a L2 users enriches language competence through language use, the L1 does not remains unaffected by the L2, confirming that the L1 is subject to change at any stage of life specially during L2 immersion in which positive feedback or evidence (Seliger, 1991) is limited, suggesting that language stabilization is a long-life process centered on language use.

The case of semantic extension in the L1 can be interpreted from different points of view. Regarding acquisition, the production of SE and sensitivity to L2 form as shown by the bilingual suggest that a possible cognitive restructuring takes place during L2 learning. Here, restructuring is seen as the result of bilingualism.

From a psycholinguistic point of view, this study provides us with a new path of research, concerning the factors that may play a role during processing and in the architecture of the bilingual lexicon, such as the across language proximity of words (Neighborhood Density) and the morphological links of languages (Morphological Family Size). Further studies including these psycholinguistic factors will be valuable for a better understanding of the underlying processes behind 'deviated productions.

Under the lexical approach, semantic extensions suggest that languages are changeable, 'a bilingualism effect' seems to be related to the enrichment the lexical levels and affecting ways information is structured creating new connections across semantic and formal levels in the bilingual mind.

An application in the context of L2 learning can be the use of pedagogical activities in which comparisons across languages are included, reinforcing metalinguistic analysis will be helpful in two different ways. In one hand, highlighting these inter-language associations may reinforce formal links across languages avoiding misleading meanings at the beginning of L2 learning. On the other hand, taking advantage of intra-language morphological links (within the L2) may reinforce L2 knowledge and thus facilitate learning and intralingual lemmatic association.

RESUME EN FRANÇAIS

Introduction

Cette thèse met l'accent sur les extensions sémantiques en L1 par le bilingue tardif espagnol (L1) français (L2). Le phénomène d'extension sémantique induit un transfert étendu de la signification d'un mot d'une langue vers un mot d'une autre langue. Il provoque ainsi une association erronée entre deux mots inter-langues (Jarvis, 2009). Un exemple, illustré par Grosjean & Py (1991, p. 58), concerne le sens du mot « *entendre* » en français, qui est étendu en L1 (espagnol) dans la structure : « *no entiendo el ruido del tren »,* « *je ne comprends pas le train* » résultant en une extension sémantique, distincte de la production canonique « *no oigo el ruido del tren* » « *je n'entends pas le train* ».

Les extensions sémantiques peuvent se manifester de L1 à L2 et inversement, ce type de transfert translinguistique présente un grand intérêt car il peut se situer à travers deux types de transferts : lemmatique ou lexémique. D'après Jarvis (2009), l'extension sémantique serait plutôt un transfert au niveau du lemme qui concerne l'ensemble des connaissances et des associations sémantiques et syntaxiques d'un mot avec d'autres mots. Nous partons de l'hypothèse que les pairs de mots inter-langues du type « *entender* » (espagnol) et « *entendre* » (français) produiraient de fausses associations inter-langues, qui induiraient aussi un transfert au niveau du lexème (et non exclusivement du lemme) puisqu'il concerne l'ensemble d'associations des propriétés formelles des mots.

Notre intérêt se porte sur le rôle de deux principaux facteurs : la densité du voisinage orthographique et la taille de la famille morphologique puisqu'ils produiraient une activation supérieure dans l'une des deux langues. Au niveau du lexème, nous tenons compte de la densité du voisinage orthographique (Grainger & Segui 1990), et au niveau du lemme, nous observons la taille de la famille morphologique (Mulder et al., 2013) en L1 et en L2. Nous nous intéressons plus particulièrement aux liens qui pourraient exister entre ces deux facteurs dans la compétition inter-langues (Costa et al., 2007), le but étant de déterminer le rôle du chevauchement formel et des liens sémantiques en production, en reconnaissance de mots, et en jugement d'acceptabilité de phrases en L1.

Cadre Théorique

Pour comprendre le fonctionnement du lexique mental, nous nous sommes intéressés (i) aux modèles psycholinguistiques afin de discerner les différents processus cognitifs liés aux interactions entre plusieurs langues mais également (ii) à la dynamique du lexique bilingue en lien avec les théories sur l'influence translinguistique autour d'une réflexion sur l'interprétation des extensions sémantiques.

Un modèle psycholinguistique permet de comprendre et de représenter de manière plausible les processus cognitifs expliquant la mémoire bilingue. Ces représentations de deux ou plusieurs systèmes linguistiques interagissent dans certains cas à différents niveaux linguistiques, par exemple entre les niveaux conceptuel et lexical. Elles

peuvent également représenter des unités linguistiques spécifiques telles que des phonèmes ou des caractéristiques formelles des mots. Le fonctionnement d'un modèle peut suivre une structure '*bottom up*', ce qui signifie que le processus d'activation commence au niveau (formel) le plus bas représenté pour arriver au niveau (conceptuel) le plus élevé de représentation ; en revanche, la structure '*top down*' va du niveau le plus bas au niveau le plus élevé de représentation (Green, 1986). Les modèles mentaux ont permis jusqu'à présent aux chercheurs de valider ou de remettre en question les hypothèses expliquant le traitement bilingue et nous permettant en même temps d'améliorer et de compléter les modèles précédents afin d'avoir une meilleure représentation des processus expliquant le lexique mental bilingue.

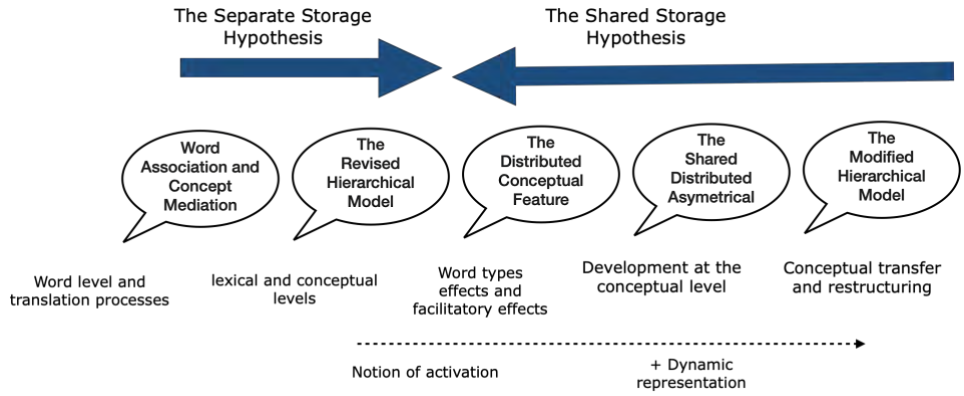

Figure 34 Comparaison des différents modèles et leurs apports à la compréhension du lexique mental bilingue

Comme nous pouvons le voir dans la **Figure 34**, au début de la représentation du lexique bilingue, les modèles tels que *Word Association* (Potter et al., 1984) limitent l'interaction lexicale entre deux systèmes (langues) exclusivement à travers des processus de traduction. Dans ce cas, l'accès au niveau conceptuel n'est possible que par la L1. D'après le *Concept Mediation Model* (Potter et al., 1984) ce n'est pas que lorsque la compétence en L2 s'améliore que des liens se renforcent, impliquant la construction de structures partagées au niveau conceptuel. Plus tard, *the Revised Hierarchical Model* Kroll & Stewart (1994) rajoutera une composante qui facilitera la représentation du lexique bilingue, celle de la modulation des liens (plus ou moins forts) existant entre les niveaux lexicaux et conceptuels entre les langues, néanmoins, le RHM indique que la L1 est censée être toujours dominante sur la L2, indépendamment du niveau de compétence des bilingues D'après ce modèle, les liens en L1 seraient plus forts qu'en L2.

Les modèles tels que *the Distributed Conceptual Feature Model* (de Groot 1992), proposent d'intégrer dans la modélisation les différents types de mots cognats et faux cognats qui affecteraient l'activation des traits linguistiques inter-langues en introduisant des effets facilitateurs. Dans cette optique, *the Shared Asymmetrical Distributed* (Dong et al., 2005) apporte une représentation plus dynamique des processus bilingues au fur et à

mesure que le niveau de maîtrise de la L2 augmente, les liens se renforcent progressivement entre les représentations formelles et les concepts spécifiques à la L2.

Enfin, le *Modified Hierarchical Model* (Pavlenko 2009) comprend des aspects de transfert et de restructuration conceptuels en accord avec, et en concevant le transfert comme un stade faisant partie de l'apprentissage, ce modèle suggère que les représentations conceptuelles translinguistiques ne sont pas complètement partagées mais partiellement partagées dans les catégories linguistiques.

En conclusion dans les modèles du lexique mental bilingue, nous avons examiné l'interaction des représentations conceptuelles et des représentations lexicales et leur place dans lesdits modèles. Cette interaction a évolué au fil du temps : les connexions indépendantes entre les deux types de représentations (conceptuelles et lexicales) ont ensuite été décrites comme des connexions distribuées/partagées. Ces variations, évolutions, apparaissent selon le positionnement théorique des auteurs et les approches méthodologiques qu'ils adoptent. Notre attention s'est particulièrement portée sur la question du transfert et de la restructuration des langues au niveau conceptuel. Ceci nous a conduit à examiner plus attentivement la reconnaissance de la dynamique et de la flexibilité des systèmes linguistiques chez le bilingue.

Nous proposons un aperçu théorique de la CLI (*Crosslinguictic Influence*) particulièrement axé sur le niveau lexical pour illustrer deux niveaux de transfert distincts : le niveau lexémique et le niveau lemmatique. Afin d'étudier leur rôle, nous proposons de prendre en compte deux caractéristiques psycholinguistiques : la densité du voisinage et les liens morphologiques inter-langues comme une perspective intéressante pour étudier les niveaux lexémiques et lemmatiques.

L"extension sémantique est « traditionnellement » étudiée suivant une approche lemmatique. Celle-ci attribue l'extension sémantique aux liens entre deux ou plusieurs lemmes dans différentes langues. Ce réseau entre les lemmes de chaque langue est censé être le résultat des associations inter-linguistiques apprises par les bilingues. Cependant, dans la présente étude, le rôle des relations lexémiques entre les langues est l'un des centres d'intérêt, étant donné que l'information lexémique implique des aspects formels des mots (par exemple, l'information phonologique et graphémique) qui sont censés être expliqués dans le cadre des niveaux d'activation et de compétition.

À cet effet, Jarvis (2009) considère que le transfert lexémique peut probablement être étendu au niveau du lemme (p. 106) mais qu'il est censé être déterminé par des facteurs purement lexémiques (par exemple, les faux cognats).

En gardant à l'esprit que les extensions sémantiques sont formées par des paires de mots interlinguistiques, qui comprennent des informations lemmatiques et lexémiques, on peut considérer que les deux approches sont complémentaires concernant la nature des extensions sémantiques.

Compte tenu des limites liées à l'étude des extensions sémantiques (par exemple, Grosjean & Py, 1991) et considérant la pertinence du principe de la redondance réduite

(Seliger, 1991), la présente étude propose une approche combinée impliquant des associations translinguistiques aux niveaux lexémique et lemmatique par le contrôle de deux caractéristiques internes des mots dont on sait qu'elles jouent un rôle dans leur accessibilité lors des processus d'accès lexical. Ces caractéristiques ont été contrôlées dans nos expérimentations avec les paires des mots inter linguistiques présentées et décrites dans la méthodologie.

La première de ces caractéristiques est la densité de voisinage interlinguistique qui joue un rôle dans la définition des liens lexémiques entre les mots. L'intérêt de cette dimension réside dans les chevauchements formels de mots existant entre les langues et leur influence sur les extensions sémantiques. Deuxièmement, la dimension de la morphologie interlinguistique implique des liens sémantiques et des relations étymologiques entre les mots qui correspondent aux niveaux lemmatiques des extensions sémantiques. L'aspect novateur de cette approche est qu'elle nous permettrait de prendre en compte des niveaux d'activation qui concurrent entre les langues, ayant pour but d'élargir le champ d'analyse des extensions sémantiques, dépassent la perspective de l'association lemmatique, précédemment prise en compte pour expliquer les extensions sémantiques.

L'effet du Voisinage

Dans cette section, nous décrirons les effets de voisinage chez les monolingues et les bilingues en nous concentrant sur les études relatives à la densité de voisinage interlinguistique. Il faut s'interroger dans un premier temps sur les processus mis en place lors du traitement visuel d'un mot par un locuteur donné. On suppose généralement qu'un mot active d'autres mots qui sont orthographiquement similaires au mot cible en mémoire. Pour illustrer ceci, lors de la reconnaissance du mot anglais « *sand* » chez le monolingue, des mots orthographiquement similaires de la même langue sont activés, comme « *hand* », « *land* » etc., (Grosjean et al., 2013, p. 82). Dans le lexique mental bilingue, les choses sont un peu différentes.

Le concept de densité du voisinage orthographique a été introduit par Coltheart et al. (1977). Grainger & Segui (1990) définissent le voisinage comme l'existence de similitudes physiques entre des mots au niveau de l'information orthographique et phonologique. Nous utiliserons le terme de « *form overlap* » (« chevauchement formel ») pour rendre compte du voisinage en général. Un voisin orthographique est un mot qui diffère du mot cible d'une seule lettre, tout en gardant la même position et la même longueur de mot (par exemple, pour « *passez* » : « *passer* », « *casser* »). Le nombre de voisins d'un mot dépend de la longueur et de la position des lettres du mot. Le voisinage peut être de nature orthographique mais également phonologique.

La densité de voisinage peut être calculée en comptant le nombre de mots orthographiquement ou phonologiquement similaires qui existent lors du changement d'un phonème ou d'une lettre du mot-cible. Trois types de phénomènes permettent de déterminer le nombre de voisins orthographiques ou phonologiques : la suppression, la

substitution ou l'ajout d'une lettre ou d'un phonème. Par exemple, le mot « *cat* » en anglais a un nombre élevé de voisins tels que : « *pat* », « *that* », « *mat* », « *sat* », « *cut* », « *cot* ». Par conséquent, le mot « *cat* » pourrait être considéré comme un mot avec un voisinage orthographique dense en anglais. En revanche, le mot « *cry* » ne compte que peu de mots similaires, tels que « *fry* », « *try* », « *dry* ». Nous pouvons donc dire que « *cry* » a un voisinage dispersé (Costa et al., 2006, p. 144). Nous pouvons remarquer dans ces exemples, comme dans d'autres, que les significations des voisins lexicaux ne sont pas du tout liées.

Coltheart et al. (1977) et Andrews (1989) suggèrent que les monolingues sont sensibles au nombre de voisins orthographiques des mots cibles et à leur fréquence lors de la reconnaissance et la dénomination des mots. Chez les locuteurs bilingues, les paires de mots inter-lexicaux ou interlinguistiques partagent des formes orthographiques et des caractéristiques phonologiques avec leur langue correspondante, en particulier lorsque les deux langues utilisent le même système alphabétique. Cela implique qu'une paire de mots interlinguistique peuvent être à la fois voisins phonologiques et voisins orthographiques (Groot, 2011).

Les effets de la densité du voisinage dans les études en CLI (Grainger & Dijkstra, 1992 ; van Heuven et al., 1998) suggèrent que la taille de la densité de voisinage de la L2 a un effet sur l'activation de l'autre langue (L1) et inversement. Pour illustrer ceci, un mot monosyllabique en espagnol comme « *mil* » (« *mille* ») a peu de voisins en espagnol (par exemple, « *vil* », « *mal* », « *mis* »), donc « *mil* » à un voisinage dispersé en espagnol mais il a un voisinage dense en anglais (par exemple, « *kill* », « *chill* », « *gill* », « *bill* », « *till* », « *miss* ») (exemple de Costa et al., 2006, p. 145). Par conséquent, les effets de densité de voisinage inter linguistique peuvent faciliter l'accès lexical dans l'une des langues seulement, tandis que pour les mots ayant peu de voisins ou une densité dispersée, ces avantages seront moindres (Costa et al., 2006).

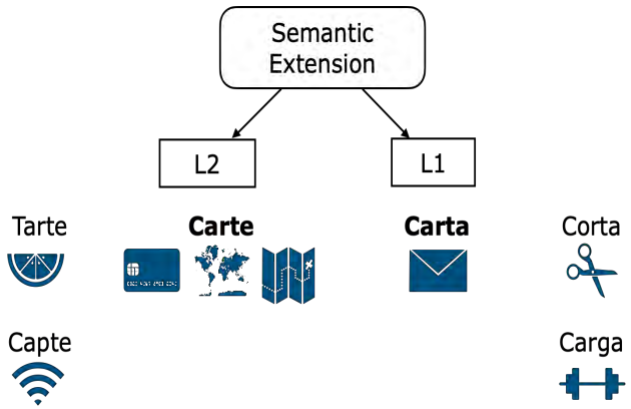

Figure 35 Illustration des mots intralinguistiques et de leurs voisins orthographiques en français et en espagnol

Dans la **Figure 35** « *Carte* » en français signifie « *carte de crédit* » contrairement à « *carta* » en espagnol qui signifie « *lettre* ». Une production contenant cet exemple a été observée sous la forme de « *se me olvidó la carta* del metro* » (« *j'ai oublié la carte de métro* ») au lieu de « *tarjeta del metro* ». Ce transfert sémantique provient de la « *carte de métro* » en L2.

Une approche tenant compte de la densité de voisinage est intéressante car elle permet aux chercheurs d'explorer les processus lors de la reconnaissance bilingue de mots, inhibée ou activée par des caractéristiques linguistiques dans une autre langue. À cet égard, l'étude de Grainger & Dijkstra (1992) conclut que les voisins orthographiques influencent la langue non-cible dans les performances bilingues, ce qui suggère que les processus de reconnaissance de mots bilingues impliquent l'activation des deux systèmes linguistiques, même lorsqu'une partie de l'information linguistique est traitée inconsciemment (comme dans les techniques d'amorçage masqué). Le paradigme de la reconnaissance visuelle des mots suggère que le voisinage orthographique est également influencé par la fréquence des mots qui font partie du voisinage et qui sont déclenchés par le mot cible (par exemple, lors d'une tâche de décision lexicale). Ainsi, si « *lame* » a 17 voisins orthographiques et que certains d'entre eux ont une fréquence élevée, l'effet de fréquence de « *lame* » sera encore plus important. La compétition translinguistique des représentations de mots entre la langue non-cible et la langue sélectionnée soutient l'idée d'un « *integrated lexicon of bilinguals* » (Brysbaert & Duyck, 2010).

Pour conclure, le lexique bilingue montre une sensibilité à la densité de voisinage interlangue, ce qui suggère que chez le bilingue les langues sont interconnectées au moins aux niveaux formels de représentation. On pourrait supposer que les liens lexicaux sont construits en intégrant les ressources lexicales provenant de plus d'une langue. Ainsi, la densité de voisinage est un facteur qui permettra d'étudier les extensions sémantiques en termes de transfert formel ou lexémique, ce qui suggère que lors de la recherche lexicale, les mots avec un voisinage dense sont plus faciles à retrouver que les mots avec un voisinage dispersé, en raison des multiples sources d'activation accumulées (Costa et al., 2006). Par la suite, l'extension sémantique peut être expliquée non seulement comme le résultat d'associations interlinguistiques apprises mais aussi comme étant un transfert lexémique dans lequel de multiples caractéristiques formelles sont activées, induisant la CLI.

La famille Morphologique :

Le rôle de la morphologie dans le traitement du langage est censé être déterminant pour définir la structure des mots, en influençant les niveaux sémantique, syntaxique, orthographique et phonologique (Sánchez-Casas & García-Albea, 2005).

L'étude des liens morphologiques est proposée ici pour illustrer les processus de co-activation au niveau du lemme lorsqu'un locuteur est confronté au traitement des extensions sémantiques dans une langue donnée. Contrairement au voisinage orthographique, les mots ayant des liens morphologiques sont généralement liés sémantiquement (par exemple, les dérivés morphologiques de « *s'énerver* » en français sont

« *énervement* », « *énervant* »). Comme mentionné ci-dessus, la structure morphologique est très complexe et suit des principes différents selon la langue concernée. Les études en CLI sont particulièrement intéressantes pour explorer la nature des représentations lexicales dans le lexique mental bilingue qui semblent être définies par la morphologie (Lalor & Kirsner, 2001). Le type de lien morphologique sur lequel se concentre la présente recherche est la taille de la famille morphologique, suivant l'hypothèse que le chevauchement de la taille de la famille morphologique contribuera à la compétition translinguistique (Mulder et al., 2013).

La taille de la famille morphologique est définie comme le nombre de mots complexes morphologiquement liés dans lesquels la racine d'un mot donné se retrouve en tant que constituant, ainsi, par exemple les mots faisant partie de grandes familles morphologiques sont traités plus rapidement dans une tâche de décision lexicale que les mots ayant une famille morphologique de plus petite taille (De Jong et al., 2000).

De plus, Frost & Grainger (2000) ont centré leur attention sur des études portant sur le traitement morphologique interlinguistique. Ils suggèrent qu'indépendamment de l'intersection orthographique et phonologique entre les langues, la morphologie joue un rôle déterminant dans le traitement. Dans le cas du finnois, des tâches de lecture impliquant l'analyse des mouvements oculaires ont montré que la taille de la famille morphologique affecte l'accès lexical dans les processus de reconnaissance de mots non fléchis (Bertram et al., 2000). Les résultats suggèrent que le niveau de familiarité et la fréquence du mot entier affectent les stades précoce et tardif de la perception des mots, ainsi que la fréquence du morphème racine jouant également un rôle dans le traitement visuel des mots.

Comme le soulignent Mulder et ses collaborateurs (2014, p. 60), « *l'effet de la taille de la famille est prédictif, en plus d'autres propriétés lexicales telles que la fréquence des mots, la fréquence des morphèmes, la longueur des mots, la taille du voisinage orthographique, la fréquence des bigrammes* ».

Forster & Azuma (2000) suggèrent que les effets morphologiques se produisent indépendamment de la similitude des formes. Ceci est le résultat de l'analyse des effets d'amorçage entre des mots morphologiquement liés et sémantiquement transparents (e.g., pli-déplié) partageant une racine commune (e.g., soumettre-permettre) avec une opacité sémantique, et comparés à une condition de contrôle orthographique. Après réduction du chevauchement orthographique, les effets d'amorçage des mots sémantiquement transparents ou opaques sont maintenus, ce qui suggère que les effets morphologiques purs n'incluent pas le chevauchement formel.

Les limites éventuelles des méthodes traditionnelles de recherche peuvent conduire à des résultats peu concluants en ce qui concerne la dichotomie entre les formes. De récentes études ERP corroborent la distinction entre voisinage orthographique et taille de la famille morphologique. Müller et ses collaborateurs (2010) suggèrent que l'effet de la taille de la famille dans les deux langues est un prédicteur plus fort que la densité du voisinage car les liens morphologiques comprennent également des représentations sémantiques des membres de la même famille.

Mulder et al. (2013) ont étudié les effets de la taille de la famille morphologique en L1 et en L2 chez des bilingues néerlandais-anglais en utilisant des ERP pour observer si les niveaux d'activation de la langue non sélectionnée (au niveau du lemme) sont limités ou s'ils s'étendent à la famille morphologique du lemme concerné. Les résultats ont montré que les signaux ERP sont plus sensibles aux lemmes de grande taille de la famille morphologique qu'aux lemmes de petite taille de la famille morphologique en L1 néerlandais. Par conséquent, un plus grand nombre de membres de la famille convergeant sémantiquement les uns avec les autres faciliterait le traitement lexical. En explorant les effets translinguistiques, ces chercheurs suggèrent que la taille de la famille L1 affecte la famille L2 en montrant des réponses plus rapides aux TRs pour des stimuli de grande taille morphologique que pour des familles de petite taille morphologique. Ainsi, l'activation du lemme se propagerait aux membres de la famille de la langue non sélectionnée.

Malheureusement, il existe peu d'études associant la recherche sur les ALS (Acquisition Langue Seconde) et la question de savoir comment les bilingues commencent à développer des relations morphologiques et sémantiques entre les mots de la L2. Cependant, nous savons que la construction de la famille morphologique primaire de la L2 du bilingue peut différer de la façon dont elle est constituée dans la L1, ce qui s'explique par la spécificité de chaque langue.

En suivant ces points de vue, on peut émettre l'hypothèse que l'apprentissage morphologique est un processus d'associations translinguistiques. Ce processus pourrait être déclenché par une sur généralisation ou une simplification des règles de deux systèmes morphologiques. Indépendamment des intersections orthographiques et phonologiques entre les langues, la morphologie semble jouer un rôle dans le traitement des mots interlinguistiques (Frost & Grainger, 2000).

L'influence translinguistique (CLI)

L'influence translinguistique (CLI) concerne la manière dont la connaissance linguistique d'une langue déjà apprise/acquise affecte ou influence l'apprentissage/acquisition d'une autre langue dans la production, la perception et la compréhension (Jarvis & Pavlenko, 2010), CLI concerne tous les niveaux linguistiques : phonologique, lexical, sémantique, syntaxique et morphologique. Cette interaction entre les systèmes a traditionnellement été étudiée de la L1 à la L2, principalement aux niveaux morphologique, syntaxique et lexical ; d'autres niveaux, comme la sémantique, ont été progressivement étudiés au fil des ans (Gathercole & Moawad, 2010). En revanche, des études en CLI de la L2 vers la L1 ont été moins nombreuses dans la littérature, malgré le fait que la compétence en L1 n'est pas un système stable et fixe comme on le supposait auparavant. La CLI au niveau lexical est d'un grand intérêt pour l'ALS car elle permet aux chercheurs d'explorer la question de l'interconnectivité entre les lexiques (Cenoz et al., 2007) dans laquelle la relation typologique entre les langues joue un rôle, ainsi que les interconnexions translinguistiques telles que la compétence linguistique.

Nous proposons de faire une distinction entre le transfert lemmatique et lexémique car cela nous permettra de mettre en évidence et d'analyser plus en détail la nature de la CLI au niveau lexical. Dans la plupart des études, la distinction entre le transfert lemmatique et lexémique n'est pas prise en considération. Les erreurs lexicales fournissent des informations précieuses sur les processus sous-jacents lors du transfert lexical. Trois types d'erreurs lexicales ont été observés et étudiés dans la littérature par Ringbom (1987, 2001) et Jarvis (2009). Ces trois types sont les faux amis, le code-switching et télescopages. Le transfert lemmatique, également appelé transfert sémantique par Ringbom, englobe les contraintes ou les restrictions collocatives, morphologiques et syntaxiques des mots dans lesquelles les correspondances sémantiques entre les lemmes sont faites, par exemple, l'utilisation restreinte des verbes à particule en anglais dans lesquels une combinaison de verbes, de prépositions, et/ou d'adverbes créent une nouvelle définition. 4 types de transferts lemmatiques sont ainsi différenciés : les extensions sémantiques, les calques, le transfert collocatif et le transfert de sous-catégorisation.

Modèle de multi-compétence

Schmid (2011, p. 13) interprète le concept de multi-compétence de Cook (2003) en y ajoutant une perspective intéressante où la multi-compétence correspond à la capacité des bilingues de changer et de fusionner leurs langues, qui peuvent en même temps être utilisées séparément et de manière sélective. Dans cette perspective, Pavlenko & Jarvis (2002) suggèrent que le cadre de la multi-compétence est en accord avec la perspective du CLI car il reconnaît que les utilisateurs de la L2 ont un état d'esprit composite qui diffère de celui de deux monolingues en un (Cook, 2003 ; Grosjean, 1989, 2010).

Ce modèle définit le bilinguisme en fonction des compétences linguistiques que les apprenants de la L2 ont acquises en tant qu'utilisateurs de la langue, en dissociant les normes de la L1 du processus d'apprentissage de la langue seconde.

En outre, il propose une vision dynamique et non linéaire de l'interaction linguistique qui a fait évoluer la recherche grâce à des perspectives plus riches allant au-delà des comparaisons traditionnelles entre la L1 et la L2 ou de l'analyse de l'apprentissage d'une langue seconde basée sur les modes d'utilisation de la langue par les locuteurs natifs de la L2.

La présente étude vise à contribuer à la recherche traitant les niveaux d'interaction des langues, tout en accordant une attention particulière au lexique, en se basant sur l'hypothèse que les systèmes sont des structures dynamiques qui se complètent et se développent en parallèle au cours de l'acquisition d'une langue, et sur les différents facteurs qui influencent le transfert, tels que l'utilisation de la langue, la dominance de la langue et l'immersion linguistique. Nous situons les extensions sémantiques dans d'autres paradigmes théoriques expliquant le changement de langue, notamment : la restructuration de la langue et l'attrition de la première langue. Enfin, nous proposons

d'interpréter le cas des extensions sémantiques comme étant le résultat d'une stratégie cognitive, l'interface entre l'utilisation des ALS et de la L1 et comme étant le résultat des processus d'inhibition de la L1.

Méthodologie

Concernant la méthodologie, trois tâches expérimentales ont été élaborées à partir de stimuli soigneusement sélectionnés en respectant : la densité du voisinage orthographique et la taille de la famille morphologique forte ou faible en L1 ou L2. Deux groupes d'hispanophones ont été testés dans leur L1 : (i) 53 monolingues en Colombie (ii) et 40 participants bilingues tardifs dont le temps moyen d'immersion en France est de 4,8 ans.

Ces indicateurs psycholinguistiques (la densité du voisinage orthographique et la taille de la famille morphologique) ne sont guère pris en compte dans les modèles d'accès lexical bilingue). C'est la raison pour laquelle nous proposons d'étudier le rôle de ces indicateurs à travers de 3 tâches expérimentales dans la L1 (espagnol) de nos participants bilingues et monolingues.

La production des extensions sémantiques dans une tâche à trous à l'oral

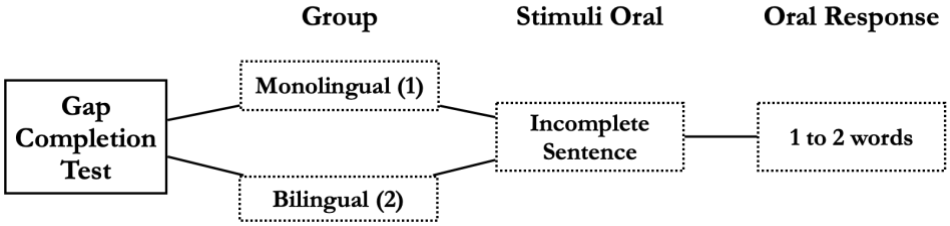

Figure 36 tâche expérimentale 1 :tâche à trous (TAT)

Le but de cette tâche a été d'observer l'influence de la L2 sur la L1 lors d'un production orale contrôlée. En fait, les extensions sémantiques apparaissent surtout dans la production orale puisque les processus impliqués pendant la parole sont en ligne (on-line).

Lors de cette tâche les participants ont été invités à écouter des stimuli oraux de phrases incomplètes (constitué par l'expérimentateur). Un signal sonore leur indiquait qu'ils devaient compléter les phrases oralement aussi vite que possible, en utilisant un ou deux mots au maximum.

Des instructions complémentaires étaient lues avant le début de la tâche. Le temps moyen investi pour cette tâche était de 8 minutes, chaque réponse étant chronométrée pendant 6.000 millisecondes.

La reconnaissance de mots dans une tâche de décision lexicale

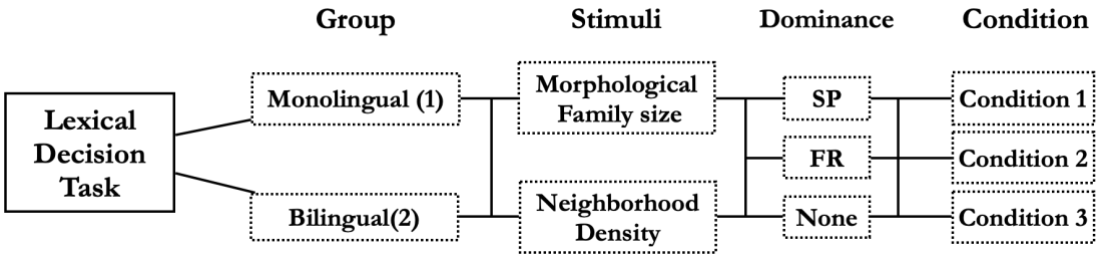

Figure 37 tâche expérimentale 2 : tâche de décision lexicale (TDL)

Cette TDL a été construite afin d'observer trois niveaux différents de traitement des extensions sémantiques. C'est pourquoi nous avons répété la même cible dans trois conditions :

- Dans la condition 1, nous avons cherché à observer si le chevauchement des formes de mots entre les langues facilitait l'activation de la L2 pendant le traitement en L1 (e.g., « *reprimir »*-«*exprimir*»);
- Dans la condition 3, nous avons testé si le chevauchement sémantique entre L1 et L2 aurait un effet sur le traitement des mots (e.g., « *expresar »*-« *exprimir »*);
- Enfin, dans la condition 2, nous avons demandé si les liens sémantiques entre les mots cibles et contrôles sont déterminés uniquement par la L1, contrairement à la condition précédente. (e.g., *apretar »,« exprimir)*;

Afin d'articuler la TDL, les stimuli cibles ont été présentés avec les pseudo-mots cibles et leurs amorces respectifs. Les participants ont décidé si les mots présentés à l'écran existent ou non en espagnol (L1). Si un mot existait en espagnol, les participants devaient taper la touche « 6 » sur le clavier (OUI), si le mot n'existe pas, les participants devaient taper la touche « 7 » (NON) sur le clavier.

La Compréhension dans la tâche de jugement de l'acceptabilité

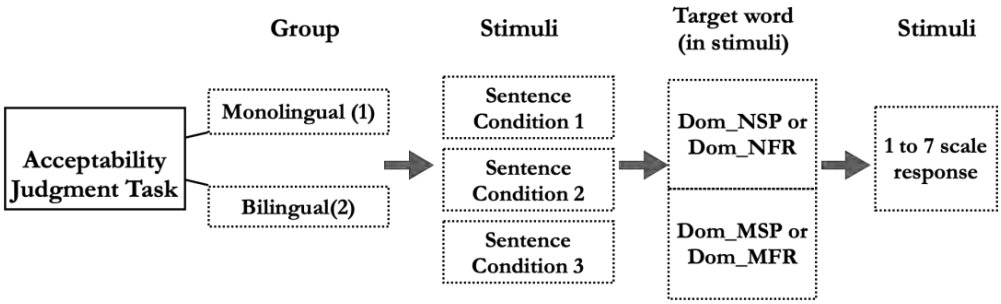

Figure 38 tâche expérimentale 3 : tâche de jugement d'acceptabilité (TJA)

L'un des principaux objectifs de la tâche de jugement d'acceptabilité est de reproduire l'étude de Grosjean & Py (1991) pour la comparer avec des résultats plus récents. Le deuxième objectif principal de tâche de jugement d'acceptabilité est de comparer les niveaux d'acceptation du groupe immergé dans un contexte dominant en L1 et ceux du groupe immergé dans un contexte dominant en L2 pour les trois types de conditions suivantes :

- La condition 1 exprime une extension sémantique en L1, qui est la signification transférée ou calquée sur la signification de la L2. *« Jesús me ENERVA muy fácilmente con su manera tan prepotente de hablar* » « *Jesús m'ENERVE très facilement avec sa façon arrogante de parler* »
- La condition 2 indique la signification canonique du mot cible en L1 :

  « *Me MOLESTA mucho María con su falta de tolerancia hacia los demás* »

  « *María m'ANNONCE beaucoup avec son manque de tolérance envers les autres* »
- La condition 3 révèle le sens approprié L1 du mot cible qui a été présenté dans la condition 1. En d'autres termes, elle se réfère à la signification "corrigée" de "*enervarse*" en espagnol.

  « *Me ENERVO al tomar este medicamento y pienso no tomármelo más*»

  « *Prendre ce médicament m'ENERVE et je ne veux plus le prendre* »

La consigne consistait à évaluer les phrases en tenant compte du contexte et en se concentrant particulièrement sur le mot présenté en majuscules. Dans chaque phrase, le mot cible était présenté en majuscules afin de permettre aux participants d'identifier le mot cible de la phrase. Les résultats attendus prédisaient que le groupe bilingue seraient plus flexible quant à son niveau d'acceptabilité des phrases que le groupe monolingue, surtout en condition 1.

Résultats

Les données obtenues ont été analysées à partir d'un modèle de régression linéaire à effets mixtes (Bates et al., 2015) et d'un modèle ordinal sur le logiciel statistique R. Nos résultats indiquent que la L2 a des effets sur la L1 en production et ce pendant le traitement et la reconnaissance de mots chez les bilingues. Cela implique donc une activation parallèle de la L1 et de la L2.

Nous avons adapté le matériel linguistique et les stimuli présentés en fonction des tâches expérimentales. Les caractéristiques internes des stimuli ont été manipulées en opposant deux types de dominances : Dominants en L1 ou L2 concernant la densité de voisinage des mots (Dom_N), et Dominants en L1 ou L2 concernant la taille de la famille morphologique des mots (Dom_M). Nous avons comparé deux groupes expérimentaux (bilingues et monolingues) partageant la même L1.

Nous avons constitué trois tâches expérimentales qui ont pour objectif l'étude des extensions sémantiques sous différentes perspectives, comme étant le résultat de l'influence de la L2 sur la L1 chez les bilingues tardifs en situation d'immersion (France).

Pour la première tâche expérimentale nous avons créé une tâche à trous (TAT) dans laquelle nous avons analysé la production des extensions sémantiques. Nous avons analysé 3,00% des extensions sémantiques produites par le groupe bilingue (41 productions réparties en 11/41 stimuli. En ce qui concerne les productions CLI qui représentaient 3,66% des productions des bilingues (50 productions) qui sont réparties sur différents phénomènes). En ce qui concerne les variables extralinguistiques étudiées en lien la production d'extension sémantique, il semblerait que les seuls niveaux de langue en L2 seraient corrélés avec le nombre d'extension sémantique produites par les participants.

Concernant la deuxième tâche expérimentale, à partir d'une tâche de décision lexicale (LDT), nous nous sommes focalisés sur les étapes initiales du traitement des langues, notamment sur la reconnaissance des mots présentant ou non une extension sémantique.

Une interaction à trois niveaux dans le groupe bilingue est observée en condition 1 lorsque la densité du voisinage est dominante dans la L2. Dans cette condition de chevauchement des formes de mots, des effets facilitateurs sont observés pendant la reconnaissance des mots pendant -41,092 millisecondes : β= -41,092, ES=17,877t=-2,299, p<0,021541.

Pour la troisième tâche expérimentale, nous avons utilisé une tâche de jugement d'acceptabilité (TJA), pour laquelle nous avons analysé les étapes tardives du traitement des langues, concernant la compréhension et les connaissances métalinguistiques à travers l'évaluation des productions déviantes en L1 (extensions sémantiques) et autres (productions canoniques et corrigées).

Il existe un effet d'interaction entre la dominance en voisinage (LR Chisq= 14,5236, df= 2, p<0,0007018). Des analyses post-hoc (moyennes et contrastes) ont montré une interaction entre les groupes lors que le stimuli est présente à un voisinage dominant en L1. Cela signifie que les bilingues ont jugé la phrase plus "acceptable" lorsqu'on leur a présenté une phrase de ce type en L1 par rapport au groupe monolingue, et ce, indépendamment de la condition présentée.

Conclusion des Résultats

Les différentes expérimentations ont montré que les bilingues produisent des extensions sémantiques et d'autres types d'influence translinguistique (CLI) lors de la production en L1. De plus, les dominances de densité de voisinage et de taille de famille morphologique jouent un rôle distinct dans les processus initiaux et finaux de compréhension. Pour la TDL, nous avons trouvé que pour le groupe bilingue, la densité de voisinage en L2 joue un rôle facilitateur dans les processus de reconnaissance des mots

en L1, dans la condition de chevauchement des formel de mots. En revanche, en ce qui concerne la TJA, les deux groupes étaient sensibles à la densité de voisinage en L1 lors de l'évaluation des phrases de la L1. Dans la TDL, la taille de la famille morphologique de la L1 semble jouer un rôle facilitateur lors de la reconnaissance des mots dans les deux groupes, lorsque le chevauchement sémantique canonique dans la L1 exprime une association canonique dans la L1. Nous émettons donc l'hypothèse que pour les bilingues tardifs, l'interférence et la facilitation de la L2 pendant la production orale sont cohérentes avec l'idée que l'accès lexical n'est pas spécifique à la langue (La Heij, 2005). Cela implique que le flux d'activation se propage aux nœuds lexicaux des deux langues, de sorte que la L2 soit fortement activée pendant le traitement et la production du matériel linguistique en L1.

Discussion

D'après les résultats, nous concluons qu'il n'y aurait pas un paramètre psycholinguistique plus important que l'autre. Ce serait plutôt une combinaison de deux facteurs qui affecte la façon dont les bilingues et les monolingues traitent l'information lexicale. Dans le cas des extensions sémantiques, ni les conditions représentant des associations d'amorçage non-canoniques représentant des extensions sémantiques (TDL), ni la présentation de phrases non-canoniques représentant des extensions sémantiques (TJA), ne permettent pas d'observer des différences significatives entre les groupes. Cela suggère que les conditions d'extensions sémantiques n'ont pas été traitées différemment dans les deux groupes. Cependant, la dominance de la L2 a influencé les performances des groupes bilingues, indiquant que les caractéristiques formelles de la L2 interviennent dans les premières étapes de reconnaissance des mots en L1.

En prenant en considération la modélisation précédente du lexique bilingue et selon nos résultats, nous convenons que la prise en compte du transfert de la L2 comme base de processus de restructuration et d'apprentissage (i.e. Modified Hierarchical Model de Pavlenko, 2009) est précieuse pour la compréhension du transfert sémantique. Dans notre étude, des spécificités sémantiques de la L2 semblent être « façonnées » dans la L1, ce qui affecte les processus de liaison entre représentations lexicales et conceptuelles lorsque les bilingues produisent des extensions sémantiques.

Dans les recherches antérieures, les extensions sémantiques sont considérées comme le résultat d'un « processus d'associations translinguistiques » ou « résultant de l'apprentissage », plutôt que le résultat de processus de niveau d'activation et compétition entre les langues (Jarvis, 2009). Cette hypothèse implique que lors du transfert sémantique, les connaissances linguistiques ralliées pendant le transfert concernent exclusivement le niveau du lemme (i.e. les associations de lemme à lemme) et excluent le niveau du lexème (i.e. les caractéristiques orthographiques et phonologiques). Notre étude propose d'élargir cette perspective en incluant les caractéristiques lexémiques comme cadre explicatif de l'extension sémantique, dans lequel la compétition entre les lexèmes joue également un rôle, en particulier lorsqu'il s'agit de confronter les similitudes perçues entre les langues.

Comme le montre la TDL, le voisinage dominant en L2 joue un rôle facilitateur au cours des premières étapes de la reconnaissance des mots chez les bilingues. En d'autres termes, la co-activation de traits phonologiques et orthographiques communes produira un effet d'amorçage lorsque les cibles se chevauchent en termes de similarités perçues (en condition 1 « *exprimir* » - « *reprimir* »). Cet effet n'est pas observé pour le groupe monolingue, ce qui suggère que pour les bilingues, le traitement du voisinage dominant en L2 est principalement motivé au niveau lexémique.

Comme on peut le constater, le bilinguisme peut être considéré comme un code génétique unique, dépendante de multiples configurations mais également du sujet en question. Nous considérons que les langues sont dynamiques et en constante évolution plutôt que stationnaires. Les « productions déviantes » et la sensibilité à la L2 peuvent être considérées ici comme une preuve que la L1 est suffisamment flexible pour changer (en termes de multicompétence et d'accessibilité linguistique), même pour les bilingues tardifs dont la L1 est traditionnellement considérée comme pleinement acquise et à un état stable définitif.

Conclusion

D'un point de vue psycholinguistique, cette étude nous ouvre une nouvelle voie de recherche, concernant les facteurs pouvant jouer un rôle lors du traitement et dans l'architecture du lexique bilingue, tels que la proximité translinguistique mots (densité du voisinage) et les liens morphologiques des langues (taille de la famille morphologique). D'autres études incluant ces facteurs psycholinguistiques seront utiles pour mieux comprendre les processus sous-jacents des « productions déviantes ».

6.
BIBLIOGRAPHY

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242–275. https://doi.org/10.1016/j.jneuroling.2006.10.003

Ahumada-Ebratt, L., Köpke, B., & Mytara, K. (2018). Les différentes langues du multilingue en interaction : Entre influence translinguistique et attrition. *Revue francaise de linguistique appliquee*, *Vol. XXIII*(2), 15–28.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Altarriba, J., & Bauer, L. M. (2004). The Distinctiveness of Emotion Concepts: A Comparison between Emotion, Abstract, and Concrete Words. *The American Journal of Psychology*, *117*(3), 389. https://doi.org/10.2307/4149007

Altarriba, J., & Heredia, R. R. (Eds.). (2008). *An introduction to bilingualism: Principles and processes*. Erlbaum.

Altenberg, E. P., & Vago, R. M. (2004). The role of grammaticality judgments in investigating first language attrition, A cross-disciplinary perspective. In M. S. Schmid, B. Köpke, M. Keijzer, & L. Weilemar (Eds.), *First language attrition: Interdisciplinary perspectives on methodological issues*. John Benjamins Pub.

Ameel, E., Malt, B. C., Storms, G., & Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language*, *60*(2), 270–290. https://doi.org/10.1016/j.jml.2008.10.001

Ameel, E., Storms, G., Malt, B. C., & Sloman, S. A. (2005). How bilinguals solve the naming problem☆. *Journal of Memory and Language*, *53*(1), 60–80. https://doi.org/10.1016/j.jml.2005.02.004

Ammerlaan, T., Hulsen, M., Strating, H., & Yağmur, K. (Eds.). (2001). *Sociolinguistic and psycholinguistic perspectives on maintenance and loss of minority languages*. Waxmann.

Andersen, G. (2011). *Corpora as lexicographical basis – The case of anglicisms in Norwegian*. Methodological and Historical Dimensions of Corpus Linguistics. http://www.helsinki.fi/varieng/series/volumes/06/andersen/

Andrews, s. (1989). Frequency and Neighborhood effects on Lexical access - activation or search. *journal of experimental psychology-learning memory and cognition*, *15*(5), 802–814. https://doi.org/10.1037/0278-7393.15.5.802

Aparicio, X., & Lavaur, J.-M. (2018). Lexical access in trilinguals: Evidence from a double masked translation priming paradigm. *Translation, Cognition & Behavior*, *1*(1), 42–73. https://doi.org/10.1075/tcb.00003.apa

Astésano, C. (2001). *Rythme et accentuation en français: Invariance et variabilité stylistique*. L'Harmattan.

Athanasopoulos, P., & Kasai, C. (2008). Language and thought in bilinguals: The case of grammatical number and nonverbal classification preferences. *Applied Psycholinguistics*, *29*(1), 105–123. https://doi.org/10.1017/S0142716408080053

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bambini, V. (2012). Neurolinguistics. In J.-O. Östman & J. Verschueren (Eds.), *Handbook of Pragmatics* (pp. 1–34). John Benjamins Publishing Company. https://doi.org/10.1075/hop.16.neu1

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bates, D., Maechler, M., Bolker, B., Walker, S. C., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., & Fox, J. (2019). *Linear Mixed-Effects Models using "Eigen" and S4* (version 1.1-21) [R package]. https://CRAN.R-project.org/package=lme4

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology*, *24*(4), 560–567. https://doi.org/10.1037/0012-1649.24.4.560

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.

Bice, K., & Kroll, J. F. (2015). Native language change during early stages of second language learning: *NeuroReport*, *26*(16), 966–971. https://doi.org/10.1097/WNR.0000000000000453

Birdsong, D. (2014). Dominance and Age in Bilingualism. *Applied Linguistics*, *35*(4), 374–392. https://doi.org/10.1093/applin/amu031

Boersma, P., & Weenink, D. (2020). *(2020). Praat: Doing phonetics by computer [Computer program]*. Version 6.1.16, retrieved 6 June 2020 from http://www.praat.org/

Boroditsky, L., Scmidt, L. A., & Phillips, W. (2003). Sex, syntax and semantics. In *Language in mind: Advances in the study of language and thought* (pp. 61–79). MIT Press.

Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in l1-l2 encoding of manner in speech and gesture: A Study of Japanese Speakers of English. *Studies in Second Language Acquisition*, *30*(02), 225–251. https://doi.org/10.1017/S0272263108080327

Brown, A., & Gullberg, M. (2011). Bidirectional cross-linguistic influence in event conceptualization? Expressions of Path among Japanese learners of English. *Bilingualism: Language and Cognition*, *14*(01), 79–94. https://doi.org/10.1017/S1366728910000064

Brown, A., & Gullberg, M. (2013). L1–L2 convergence in clausal packaging in Japanese and English. *Bilingualism: Language and Cognition*, *16*(03), 477–494. https://doi.org/10.1017/S1366728912000491

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, *109*(2), 204–223. https://doi.org/10.1037/0033-2909.109.2.204

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition*, *13*(03), 359–371. https://doi.org/10.1017/S1366728909990344

Calvet, L.-J. (1991). L'argot comme variation diastratique, diatopique et diachronique (autour de Pierre Guiraud). *Langue française*, *90*(1), 40–52. https://doi.org/10.3406/lfr.1991.6194

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Caramazza, A., & Brones, I. (1980). Semantic classification by bilinguals. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *34*(1), 77–81. https://doi.org/10.1037/h0081016

Cenoz, J., Hufeisen, B., & Jessner, U. (2007). Why investigate the multilingual lexicon? In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *The multilingual lexicon* (pp. 1–9). Kluwer Academic Publishers. http://accesbib.uqam.ca/cgi-bin/bduqam/transit.pl?&noMan=25127878

Chafe, W. L., & Nichols, J. (1986). *Evidentiality: The linguistic coding of epistemology*. Ablex.

Chang, C. B. (2019). Phonetic Drift. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 191–203).

Christensen, R. H. B. (2019). *ordinal: Regression Models for Ordinal Data* (Version 2019-12-10) [Computer software]. https://CRAN.R-project.org/package=ordinal

CNRT. (2020, May 29). *Centre National des Ressources Textuelles et lexicales, version 2.1., [version 23.1 en ligne]. Https://www.cnrtl.fr/definition/*.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance* (pp. 535–555). Lawrence Erlbaum, Mahwah, NJ, USA.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Cook, V. (2003). *Effects of the second language on the first*. Multilingual Matters.

Cook, V. (2013). *Second language learning and language teaching* (4. ed). Routledge.

Cook, V., & Bassetti, B. (2011). *Language and bilingual cognition*. Psychology Press. http://site.ebrary.com/id/10447776

Corominas, J., & Pascual, J. A. (2008). *Breve diccionario etimológico de la lengua castellana*. Editorial Gredos.

Costa, A., Heij, W. L., & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, *9*(2), 137–151. https://doi.org/10.1017/S1366728906002495

Davis, C. J., & Taft, M. (2005). More words in the neighborhood: Interference in lexical decision due to deletion neighbors. *Psychonomic Bulletin & Review*, *12*(5), 904–910. https://doi.org/10.3758/BF03196784

De Bot, K. (2001). Language use as an interface between sociolinguistic and psycholinguistic processes in language attrition and language shift. In Klatter-Folmer & P. van Avermaet (Eds.), *Theories on maintenance and loss of minority languages: Towards a more integrated explanatory framework* ((eds), pp. 65–81). Waxmann.

De Bot, K. (2004). The Multilingual Lexicon: Modelling Selection and Control. *International Journal of Multilingualism*, *1*(1), 17–32. https://doi.org/10.1080/14790710408668176

De Groot, A.M.B. (1992). Bilingual lexical representation: A closer look at conceptual representations. In *Orthography, phonology, morphology, and meaning* (pp. 389–412). North-Holland. https://doi.org/10.1016/S0166-4115(08)62805-8

De Groot, A.M.B., & Hoeks, J. C. J. (1995). The Development of Bilingual Memory: Evidence from Word Translation by Trilinguals. *Language Learning*, *45*(4), 683–724. https://doi.org/10.1111/j.1467-1770.1995.tb00458.x

De Groot, A.M.B., & Poot, R. (1997). Word Translation at Three Levels of Proficiency in a Second Language: The Ubiquitous Involvement of Conceptual Memory. *Language Learning*, *47*(2), 215–264. https://doi.org/10.1111/0023-8333.71997007

De Groot, A.M.B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, *30*(1), 90–123. https://doi.org/10.1016/0749-596X(91)90012-9

De Jong, N. H., Schreuder, R., & Harald Baayen, R. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes*, *15*(4–5), 329–365. https://doi.org/10.1080/01690960050119625

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dewaele, J.-M. (2018). *Glimpses of semantic restructuring of English emotion-laden words of American English L1 users residing outside the USA*. Lab.15046.Dew. Retrieved July 13,

2018, from https://benjamins.com/catalog/lab.15046.dew

Dewaele, J.-M. (2018). Glimpses of semantic restructuring of English emotion-laden words of American English L1 users residing outside the USA. *Linguistic Approaches to Bilingualism*, *8*(3), 320–342. https://doi.org/10.1075/lab.15046.dew

Dewaele, J.-M., & Pavlenko, A. (2003). Productivity and lexical diversity in native and non-native speech: A study of cross-cultural effects. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 120–141). Multilingual Matters.

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, *41*(4), 496–518. https://doi.org/10.1006/jmla.1999.2654

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, *62*(3), 284–301. https://doi.org/10.1016/j.jml.2009.12.003

Dong, Y., Gui, S., & Macwhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition*, *8*(3), 221–238. https://doi.org/10.1017/S1366728905002270

Dubois, J., Mitterand, H., & Dauzat, A. (Eds.). (2011). *Dictionnaire étymologique & historique du français: Une précieuse histoire des mots pour comprendre la langue d'aujourd'hui à travers celle d'hier*. Larousse.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, *45*(4), 1246–1258. https://doi.org/10.3758/s13428-013-0326-1

Durgunoğlu, A. Y., & Roediger, H. L. (1987). Test differences in accessing bilingual memory. *Journal of Memory and Language*, *26*(4), 377–391. https://doi.org/10.1016/0749-596X(87)90097-0

Ecke, P. (2004). Language attrition and theories of forgetting: A cross-disciplinary review. *International Journal of Bilingualism*, *8*(3), 321–354. https://doi.org/10.1177/13670069040080030901

Ecke, P. (2009). The Tip-of-the-Tongue Phenomenon as a Window on (Bilingual) Lexical Retrieval. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 185–208). Multilingual Matters.

Ecke, P., & Hall, C. J. (2013). Tracking tip-of-the-tongue states in a multilingual speaker: Evidence of attrition or instability in lexical systems? *International Journal of Bilingualism*, *17*(6), 734–751. https://doi.org/10.1177/1367006912454623

Ellis, N., C. (2005). Introduction to Part 1. In Judith F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 3–8). Oxford University Press.

Ervins, S. M., & Osgood, C. E. (1954). Second language learning and bilingualism. *Journal of Abnormal and Social Psychology*, 139–145.

Escudero, P., & Boersma, P. (2004). Bridging the gap between l2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*(04). https://doi.org/10.1017/S0272263104040021

Fabbro, F. (1999). *The neurolinguistics of bilingualism: An introduction*. Psychology Press.

Fishman, J. A. (2012). Language Maintenance, Language Shift, and Reversing Language Shift. In *The Handbook of Bilingualism and Multilingualism* (pp. 466–494). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118332382.ch19

Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of english. *Speech Communication*, *6*(3), 185–202. https://doi.org/10.1016/0167-6393(87)90025-2

Forster, K. I., & Azuma, T. (2000). Masked priming for prefixed words with bound stems: Does submit prime permit? *Language and Cognitive Processes*, *15*(4–5), 539–561. https://doi.org/10.1080/01690960050119698

Forster, K. I., & Taft, M. (1994). Bodies, antibodies, and neighborhood-density effects in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 844–863. https://doi.org/10.1037/0278-7393.20.4.844

Frost, R., & Grainger, J. (2000). Cross-linguistic perspectives on morphological processing: An introduction. *Language and Cognitive Processes*, *15*(4–5), 321–328. https://doi.org/10.1080/01690960050119616

García-Albea, J. E., Sánchez-Casas, R., & Valero, T. (1996). *Form and meaning contribution to word recognition in Catalan-Spanish bilinguals. Conference of European Society for Cognitive Psychology.*

Gass, S. M., & Selinker, L. (1992). *Language Transfer in Language Learning*. John Benjamins Publishing.

Gathercole, V. C. M., & Moawad, R. A. (2010). Semantic interaction in early and late bilinguals: All words are not created equally*. *Bilingualism: Language and Cognition*, *13*(4), 385–408. https://doi.org/10.1017/S1366728909990460

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(2), 305–315. https://doi.org/10.1037/0278-7393.15.2.305

Glanzer, M., & Duarte, A. (1971). Repetition between and within languages in free recall. *Journal of Verbal Learning & Verbal Behavior*, *10*(6), 625–630. https://doi.org/10.1016/S0022-5371(71)80069-5

Goggin, J., & Wickens, D. D. (1971). Proactive interference and language change in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *10*(4), 453–458. https://doi.org/10.1016/S0022-5371(71)80046-4

Gollan, Tamar H., & Acenas, L.-A. R. (2004). What Is a TOT? Cognate and

Translation Effects on Tip-of-the-Tongue States in Spanish-English and Tagalog-English Bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 246–269. https://doi.org/10.1037/0278-7393.30.1.246

Gollan, Tamar H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Gollan, Tanar H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*(7), 1220–1234. https://doi.org/10.3758/BF03193224

Goral, M., Campanelli, L., & Spiro, A. (2015). Language dominance and inhibition abilities in bilingual older adults*. *Bilingualism: Language and Cognition*, *18*(1), 79–89. https://doi.org/10.1017/S1366728913000126

Goral, M., Libben, G., Obler, L. K., Jarema, G., & Ohayon, K. (2008). Lexical attrition in younger and older bilingual adults. *Clinical Linguistics & Phonetics*, *22*(7), 509–522. https://doi.org/10.1080/02699200801912237

Graham, C. R., & Belnap, R. K. (1986). The Acquisition of Lexical Boundaries in English by Native Speakers of Spanish. *IRAL - International Review of Applied Linguistics in Language Teaching*, *24*(1–4). https://doi.org/10.1515/iral.1986.24.1-4.275

Grainger, J., & Dijkstra, T. (1992). On the Representation and Use of Language Information in Bilinguals. In *Advances in Psychology* (Vol. 83, pp. 207–220). Elsevier. https://doi.org/10.1016/S0166-4115(08)61496-X

Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Perception & Psychophysics*, *47*(2), 191–198. https://doi.org/10.3758/BF03205983

Graves, S., Piepho, H.-P., & Selzer, L. (2019). *multcompView: Visualizations of Paired Comparisons* (version 0.1-8) [R package]. https://CRAN.R-project.org/package=multcompView

Green, D. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, *27*(2), 210–223. https://doi.org/10.1016/0093-934X(86)90016-7

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

Groot, A. M. B. de. (2011). Comprehension Processes: Word Recognition and Sentence Processing. In *Language and cognition in bilinguals and multilinguals: An introduction* (pp. 155–219). Psychology Press. http://www.123library.org/book_details/?id=74712

Groot, Annette M. B. de. (2013). Bilingual Memory. In F. Grosjean & P. Li, *The*

psycholinguistics of bilingualism (pp. 171–191). Wiley-Blackwell.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3–15. https://doi.org/10.1016/0093-934X(89)90048-5

Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.

Grosjean, F. (2018). Multilinguisme: Diversité des approches. *Revue francaise de linguistique appliquee*, *Vol. XXIII*(2), 7–14.

Grosjean, F., Li, P., & Bialystok, E. (2013). *The psycholinguistics of bilingualism*. Wiley-Blackwell/John Wiley & Sons.

Grosjean, F., & Py, B. (1991). La restructuration d'une première langue: L'intégration de variantes de contact dans la compétence de migrants bilingues. *La Linguistique*, *27*(2), 35–60.

Gumperz, J.J. and Hernández-Chávez, E. (1970). *Cognitives aspects of bilingual communication. In Hernández-Chávez et al. (Eds), El lenguaje de los Chicanos.* (Arlington Center for Apllied linguistics).

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory* (pp. xix, 335). Wiley.

Heij, W. L., Hooglander, A., Kerling, R., & van der Velden, E. (1996). Nonverbal Context Effects in Forward and Backward Word Translation: Evidence for Concept Mediation. *Journal of Memory and Language*, *35*(5), 648–665. https://doi.org/10.1006/jmla.1996.0034

Heredia, R. R. (2008). Mental Models of Bilingual Memory. In J. Altarriba & R. R. Heredia (Eds.), *An introduction to bilingualism: Principles and practies* (pp. 39–67). Lawrence Erlbaum Associates.

Heredia, R. R., & Brown, J. M. (2006). Bilingualism: Memory, Cognition, and Emotion. In T. K. Bhatia (Ed.), *The handbook of bilingualism* (pp. 225–249). Blackwell Pub.

Heredia, R. R., & Brown, J. M. (2012, October 3). *Bilingual Memory*. https://doi.org/10.1002/9781118332382.ch11

Hervé, M. (2020). *RVAideMemoire: Testing and Plotting Procedures for Biostatistics* (version 0.9-78) [R package]. https://CRAN.R-project.org/package=RVAideMemoire

Hohenstein, J., Eisenberg, A., & Naigles, L. (2006). Is he floating across or crossing afloat? Cross-influence of L1 and L2 in Spanish–English bilingual adults. *Bilingualism: Language and Cognition*, *9*(3), 249–261. https://doi.org/10.1017/S1366728906002616

Hohenstein, J. M., Naigles, L. R., & Eisenberg, A. R. (2004). Keeping Verb Acquisition in Motion: A Comparison of English and Spanish. In *Weaving a lexicon* (pp. 569–602). MIT Press.

Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Language Learning & Language Teaching* (Vol. 32, pp. 1–20). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.32.01hou

Hulsen, M. E. H. (2000). *Language loss and language processing: Three generations of Dutch migrants in New Zealand*. Hulsen.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected Readings* (pp. 269–293). Penguin Books.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277–288. https://doi.org/10.1016/j.actpsy.2007.06.003

Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*, *19*(1), 57–84. https://doi.org/10.1191/0265532202lt220oa

Jarvis, S. (2003). Probing the Effects of the L2 on the L1: A Case Study. In V. Cook (Ed.), *Effects of the second language on the first*. Multilingual Matters.

Jarvis, S. (2019). Lexical attrition. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 241–250).

Jarvis, S., & Pavlenko, A. (2010). *Crosslinguistic influence in language and cognition* (paperback ed). Routledge.

Jiang, N. (2002). Form–meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, *24*(4), 617–637. https://doi.org/10.1017/S0272263102004047

Jiang, N. (2004). Semantic Transfer and Its Implications for Vocabulary Teaching in a Second Language. *The Modern Language Journal*, *88*(3), 416–432. https://doi.org/10.1111/j.0026-7902.2004.00238.x

Joseph, Briatte, F., & Chl. (2019). *larmarange/analyse-R: Version du 6 mai 2019* (Version 2019-05-06) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.2669067

Kartushina, N., Frauenfelder, U. H., & Golestani, N. (2016). How and When Does the Second Language Influence the Production of Native Speech Sounds: A Literature Review: L2 Influences on L1: A Literature Review. *Language Learning*, *66*(S2), 155–186. https://doi.org/10.1111/lang.12187

Kasper, G., & Kellerman, E. (Eds.). (1999). *Communication strategies: Psycholinguistic and sociolinguistic perspectives* (2. impr). Longman.

Keatley, C. W., Spinks, J. A., & De Gelder, B. (1994). Asymmetrical cross-language priming effects. *Memory & Cognition*, *22*(1), 70–84. https://doi.org/10.3758/BF03202763

Kellerman, E. (1995). Crosslinguistic Influence: Transfer to Nowhere? *Annual Review of Applied Linguistics*, *15*, 125–150. https://doi.org/10.1017/S0267190500002658

Klatter-Folmer, & Avermaet, P. van (Eds.). (2001). *Theories on maintenance and loss of minority languages: Towards a more integrated explanatory framework* ((eds)). Waxmann.

Kolers, P. A., & Gonzalez, E. (1980). Memory for words, synonyms, and translations. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(1), 53–65. https://doi.org/10.1037/0278-7393.6.1.53

Köpke, B. (2002). Activation Thresholds and non-pathological first language attrition. In F. Fabbro (Ed.), *Advances in the neurolinguistics of bilingualism: Essays in honor of Michel Paradis* (pp. 119–142). Forum.

Köpke, B. (2009). *Approches neuropsycholinguistiques de la gestion des langues chez le sujet plurilingue*. Toulouse : UTM.

Köpke, B. (2020). Language, Interaction, and Acquisition Brain plasticity in the interaction between languages: What does language attrition tell us? *[Unpublished Manuscript]. Université Toulouse Jean Jaures.*

Köpke, B., & Genevska-Hanke, D. (2018). First Language Attrition and Dominance: Same Same or Different? *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01963

Köpke, B., & Keijzer, M. (2019). Introduction to psycholinguistic and neurolinguistic approaches to language attrition. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 63–72).

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language*, *33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kroll, Judith F., & Groot, A. M. B. de. (1997). Lexical and Conceptual Memory in the Bilingual: Mapping Form to Meaning in Two Languages. In Annette M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 169–199). Lawrence Erlbaum Associates Publishers.

Kroll, Judith F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, *18*(2), 137–171. https://doi.org/10.1191/0267658302sr201oa

Kroll, Judith F, & Tokowicz, N. (2005). Models of Bilingual Representation and Processing: Looking Back and to the Future. In Judith F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 531–553). Oxford University Press.

Kroll, Judith F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and*

Cognition, *13*(3), 373–381. https://doi.org/10.1017/S136672891000009X

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B., & Pødenphant Jensen, S. (2019). *lmerTest: Tests in Linear Mixed Effects Models* (version 3.1-1) [Computer software]. https://CRAN.R-project.org/package=lmerTest

La Heij, W. (2005). Selection Processes in Monolingual and Bilingual Lexical Access. In Judith F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 289–307). Oxford University Press.

Labov, W. (2006). *The social stratification of English in New York City* (2nd ed). Cambridge University Press.

Lalor, E., & Kirsner, K. (2001). The representation of "false cognates" in the bilingual lexicon. *Psychonomic Bulletin & Review*, *8*(3), 552–559. https://doi.org/10.3758/BF03196191

Lang, B., & Davidson, L. (2017). Effects of Exposure and Vowel Space Distribution on Phonetic Drift: Evidence from American English Learners of French. *Language and Speech*, *62*(1), 30–60. https://doi.org/10.1177/0023830917737111

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. https://doi.org/10.18637/jss.v069.i01

Levelt, W. J. M. (1993). *Speaking: From Intention to Articulation*. MIT Press.

Levshina, N. (2015). *How to do Linguistics with R*. John Benjamins Publishing Company. https://benjamins.com/catalog/z.195

Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting Your Native Language: The Role of Retrieval-Induced Forgetting During Second-Language Acquisition. *Psychological Science*, *18*(1), 29–34. https://doi.org/10.1111/j.1467-9280.2007.01844.x

Lickley, R. J. (2015). Fluency and Disfluency. In M. A. Redford (Ed.), *The Handbook of Speech Production* (pp. 445–474). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118584156.ch20

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing Access to the Native Language While Immersed in a Second Language: Evidence for the Role of Inhibition in Second-Language Learning. *Psychological Science*, *20*(12), 1507–1515. https://doi.org/10.1111/j.1467-9280.2009.02480.x

Lindqvist, C. (2012). Advanced learners' word choices in French L3. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Studies in Bilingualism* (Vol. 46, pp. 255–280). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.46.15lin

López, M., & Young, R. K. (1974). The linguistic interdependence of bilinguals. *Journal of Experimental Psychology*, *102*(6), 981–983. https://doi.org/10.1037/h0036544

Lynch, A. (2017). Bilingualism and Second Language Acquisition. In N. Van

Deusen-Scholl & S. May (Eds.), *Second and Foreign Language Education* (pp. 43–55). Springer International Publishing. https://doi.org/10.1007/978-3-319-02246-8_5

Major, R. C. (1992). Losing English as a First Language. *The Modern Language Journal*, *76*(2), 190–208. JSTOR. https://doi.org/10.2307/329772

Malt, B. C., Jobe, R. L., Li, P., Pavlenko, A., & Ameel, E. (2016). What constrains simultaneous mastery of first and second language word use? *International Journal of Bilingualism*, *20*(6), 684–699. https://doi.org/10.1177/1367006915583565

Malt, B. C., Li, P., Pavlenko, A., Zhu, H., & Ameel, E. (2015). Bidirectional lexical interaction in late immersed Mandarin-English bilinguals. *Journal of Memory and Language*, *82*, 86–104. https://doi.org/10.1016/j.jml.2015.03.001

Malt, B. C., & Sloman, S. A. (2003). Linguistic diversity and object naming by non-native speakers of English. *Bilingualism: Language and Cognition*, *6*(1), 47–67. https://doi.org/10.1017/S1366728903001020

Malt, B. C., Sloman, S. A., & Gennari, S. P. (2003). Universality and language specificity in object naming. *Journal of Memory and Language*, *49*(1), 20–42. https://doi.org/10.1016/S0749-596X(03)00021-4

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLOS ONE*, *7*(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Mehotcheva, T., & Köpke, B. (2019). Introduction to L2 attrition. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 331–348).

Meriläinen, L. (2006). *Lexical transfer errors in the written English of Finnish Upper Secondary School students.* Unpublished licentiate thesis, University of Joensuu.

Meuter, R. (2009). Neurolinguitic Contributions to Undertanding the Bilingual Mental Lexicon. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 1–25). Multilingual Matters.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227–234. https://doi.org/10.1037/h0031564

Mohammed, E. A., Naugler, C., & Far, B. H. (2015). Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics. In *Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology* (pp. 577–602). Elsevier. https://doi.org/10.1016/B978-0-12-802508-6.00032-6

Montrul, S. (2011). Multiple interfaces and incomplete acquisition. *Lingua*, *121*(4), 591–604. https://doi.org/10.1016/j.lingua.2010.05.006

Montrul, S., & Polinsky, M. (2019). Introduction to heritage language development. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 419–433).

Mulder, K., Dijkstra, T., Schreuder, R., & Baayen, H. R. (2014). Effects of primary and secondary morphological family size in monolingual and bilingual word processing. *Journal of Memory and Language*, *72*, 59–84.

Mulder, K., Schreuder, R., & Dijkstra, T. (2013). Morphological Family Size effects in L1 and L2 processing: An electrophysiological study. *Language and Cognitive Processes*, *28*(7), 1004–1035. https://doi.org/10.1080/01690965.2012.733013

Müller, O., Duñabeitia, J. A., & Carreiras, M. (2010). Orthographic and associative neighborhood density effects: What is shared, what is different? *Psychophysiology*, *47*(3), 455–466. https://doi.org/10.1111/j.1469-8986.2009.00960.x

Murphy, G. L. (2004). *The big book of concepts* (1. MIT Press paperback ed). MIT Press.

Myers-Scotton, C., & Jake, J. L. (2005). Matching lemmas in bilingual language competence and production model: Evidence from intrasentential code-switching. In L. Wei (Ed.), *The bilingualism reader* (Reprinted, pp. 281–320). Routledge.

Nespoulous, J.-L. (2004). Linguistique, pathologie du langage et cognition. In C. Fuchs (Ed.), *La linguistique cognitive* (pp. 171–194). Éditions de la Maison des sciences de l'homme. https://doi.org/10.4000/books.editionsmsh.7055

Nocaudie, O. (2016). *Imitation et contrôle prosodique dans l'entraînement à la remédiation phonétique: Évaluation, mesure et applications pour l'enseignant en langue étrangère* [These de doctorat, Toulouse 2]. http://www.theses.fr/2016TOU20123

Ortega, L. (2013). *Understanding second language acquisition*. Routledge, Taylor & Francis Group.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *45*(3), 255–287. https://doi.org/10.1037/h0084295

Paivio, A., Clark, J. M., & Lambert, W. E. (1988). Bilingual dual-coding theory and semantic repetition effects on recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 163–172. https://doi.org/10.1037/0278-7393.14.1.163

Paivio, A., & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *34*(4), 388–399. https://doi.org/10.1037/h0081101

Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of" interference" in bilingual speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics*.

Paradis, M. (1997). The cognitive neuropsychology of bilingualism. In A. M. B. De Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 331–354). Lawrence Erlbaum Associates Publishers. https://psycnet.apa.org/record/1997-08536-012

Paradis, M. (2004). *A Neurolinguistic Theory of Bilingualism* (Vol. 18). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.18

Paradis, M. (2007). L1 attrition features predicted by a neurolinguistic theory of bilingualism. In B. Köpke, M. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language attrition: Theoretical perspectives*. Benjamins; 121-134.

Pavlenko, A. (2000). L2 Influence on L1 in Late Bilingualism. *Issues in Applied Linguistics*, *11*(2). https://escholarship.org/uc/item/7gs944m5

Pavlenko, A. (2002). Bilingualism and emotions. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication*, *21*(1), 45–78. https://doi.org/10.1515/mult.2002.004

Pavlenko, A. (2003). Eyewitness memory in late bilinguals: Evidence for discursive relativity. *International Journal of Bilingualism*, *7*(3), 257–281. https://doi.org/10.1177/13670069030070030301

Pavlenko, A. (Ed.). (2009). *The bilingual mental lexicon: Interdisciplinary approaches*. Multilingual Matters.

Pavlenko, A. (2014). *The bilingual mind: And what it tells us about language and thought*. Cambridge Univ. Press.

Pavlenko, A., & Jarvis, S. (2002). Bidirectional Transfer. *Applied Linguistics*, *23*(2), 190–214. https://doi.org/10.1093/applin/23.2.190

Pavlenko, A., & Malt, B. C. (2011). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian–English bilinguals. *Bilingualism: Language and Cognition*, *14*(01), 19–45. https://doi.org/10.1017/S136672891000026X

Phillips, N. A., Klein, D., Mercier, J., & de Boysson, C. (2006). ERP measures of auditory word repetition and translation priming in bilinguals. *Brain Research*, *1125*(1), 116–131. https://doi.org/10.1016/j.brainres.2006.10.002

Pistono, A. (2017). *Exploration du discours dans le vieillissement typique et la maladie d'Alzheimer: Liens avec les modifications neurocognitives sous-jacentes* [Phd, Université de Toulouse, Université Toulouse III - Paul Sabatier]. http://thesesups.ups-tlse.fr/3678/

Poplack, S. (2005). Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. In L. Wei (Ed.), *The bilingualism reader* (Reprinted, pp. 221–256). Routledge.

Potter, M. C., So, K.-F., Eckardt, B. V., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, *23*(1), 23–38. https://doi.org/10.1016/S0022-5371(84)90489-4

Poulisse, N. (2000). Slips of the tongue in first and second language production. *Studia Linguistica*, *54*(2), 136–149.

Psychology Software Tools, Inc. (E-Prime 2.0). (2016). [Computer software]. Retrieved

from https://www.pstnet.com.

Quilis, A., Adell, M., Arranz, C., Cabezas, M., Fernández, J. A., García, B., Gatón, S., Gil Craviotto, F., Gutiérrez, P., Herreras, J. C., Navarro, J., Planelles, P., & Plaza, I. (1982). *Interferencias lingüísticas en el habla de los niños españoles emigrantes en Francia*. Ministerio de Educación y Ciencia.

Quilis, Antonio. (1981). *Fonética acústica de la lengua española*. Gredos.

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

RAE. (2020, May 29). *Diccionario de la lengua española, 23.ª ed., [versión 23.3 en línea]. Https://dle.rae.es*.

Riehl, C. M. (2019). Language contact and language attrition. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 314–328).

Ringbom, H. (1987). *The role of the first language in foreign language learning*. Multilingual Matters Ltd.

Ringbom, H. (2001). Lexical Transfer in L3 Production. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psychological perspectives* (pp. 59–68). Multilingual Matters.

Ringbom, H. (2007). *Cross-linguistic similarity in foreign language learning*. Multilingual Matters.

Roelofs, A. (2003). Shared phonological encoding processes and representations of languages in bilingual speakers. *Language and Cognitive Processes*, *18*(2), 175–204. https://doi.org/10.1080/01690960143000515

Romaine, S. (1989). *Bilingualism*. B. Blackwell.

Salcedo, C. S. (2010). The phonological system of spanish. *Revista de Lingüística y Lenguas Aplicadas*, *5*(1). https://doi.org/10.4995/rlyla.2010.769

Sánchez-Casas, R., & García-Albea, J. E. (2005). The Representation of Cognate and Noncognate Words in Bilingual Memory: Can Cognate Status Be Characterized as a Special Kind of Morphological Relation? In Judith F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 226–250). Oxford University Press.

Sánchez-Casas, R., Suarez-Buratti, B., & Igoa, J. M. (1992). *Are Bilingual Lexical Representations Interconnected?* [Paper presented at]. Fifth conference of the European Society for Cognitive Psychology, Paris., Paris, France.

Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behavior*, *23*(1), 84–99. https://doi.org/10.1016/S0022-5371(84)90519-X

Schepens, J. J., Slik, F. van der, & Hout, R. van. (2016). L1 and L2 Distance Effects in Learning L3 Dutch. *Language Learning*, *66*(1), 224–256.

https://doi.org/10.1111/lang.12150

Schmid, M. S. (2011). *Language attrition*. Cambridge Univ. Press.

Schmid, M. S., & Jarvis, S. (2014a). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, *17*(4), 729–748. https://doi.org/10.1017/S1366728913000771

Schmid, M. S., & Jarvis, S. (2014b). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition*, *17*(4), 729–748. https://doi.org/10.1017/S1366728913000771

Schmid, M. S., & Köpke, B. (2007). Bilingualism and attrition. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Studies in Bilingualism* (Vol. 33, pp. 1–7). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.33.02sch

Schmid, M. S., & Köpke, B. (2009). L1 Attrition and the Mental Lexicon. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 209–238). Multilingual Matters.

Schmid, M. S., & Köpke, B. (2017). The relevance of first language attrition to theories of bilingual development. *Linguistic Approaches to Bilingualism*, *7*(6), 637–667. https://doi.org/10.1075/lab.17058.sch

Schreuder, R., & Baayen, R. H. (1997). How Complex Simplex Words Can Be. *Journal of Memory and Language*, *37*(1), 118–139.

Schwanenflugel, P. J., Akin, C., & Luh, W.-M. (1992). Context availability and the recall of abstract and concrete words. *Memory & Cognition*, *20*(1), 96–104. https://doi.org/10.3758/BF03208259

Seliger, H. W. (1991). Language attrition, reduced redundancy, and creativity. In H. W. Seliger & R. M. Vago (Eds.), *First Language Attrition* (1st ed., pp. 227–240). Cambridge University Press. https://doi.org/10.1017/CBO9780511620720.015

Sharwood Smith, M. (1983). Cross-linguistic Aspects of Second Language Acquisition. *Applied Linguistics*, *4*(3), 192–199. https://doi.org/10.1093/applin/4.3.192

Sharwood Smith, M. (2019). Language attrition as a special case of processing change: A wider cognitive perspective. In M. S. Schmid & B. Köpke (Eds.), *The Oxford handbook of language attrition* (pp. 77–87).

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer Between Picture Naming and Translation: A Test of Asymmetries in Bilingual Memory. *Psychological Science*, *6*(1), 45–49. https://doi.org/10.1111/j.1467-9280.1995.tb00303.x

Simon, J. R., & Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus-response correspondence and age. *Ergonomics*, *6*(1), 99–105. https://doi.org/10.1080/00140136308930679

Singleton, D. (1999). *Exploring the Second Language Mental Lexicon* (1st ed.).

Cambridge University Press. https://doi.org/10.1017/CBO9781139524636

Slobin, D. I. (1993). Adult language acquisition: A view from child language study. In C. Perdue (Ed.), *Adult language acquisition: Cross-linguistic perspectives* (pp. 238–252). Cambridge University Press.

Steinkrauss, R., & Schmid, M. S. (2017). Entrenchment and language attrition. In H.-J. Schmid (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge.* (pp. 367–383). American Psychological Association. https://doi.org/10.1037/15969-017

Szumilas, M. (2010). Explaining Odds Ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *19*(3), 227–229.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530–12535. https://doi.org/10.1073/pnas.0609927104

Thierry, Guillaume, & Sanoudaki, E. (2012). Activation syntaxique non-sélective à la langue chez le bilingue précoce. *Revue française de linguistique appliquée*, *XVII*(2), 33. https://doi.org/10.3917/rfla.172.0033

Treffers-Daller, J., & Silva-Corvalán, C. (2016). *Language Dominance in Bilinguals: Issues of Measurement and Operationalization*. Cambridge University Press.

Vaissière, J. (2010). *Le français, langue à frontières par excellence*. ⟨halshs-00456280⟩

Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*(3), 193–211. https://doi.org/10.1017/S1366728998000352

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

Van Heuven, W. J. B., Dijkstra, T., Grainger, J., & Schriefers, H. (2001). Shared neighborhood effects in masked orthographic priming. *Psychonomic Bulletin & Review*, *8*(1), 96–101. https://doi.org/10.3758/BF03196144

Van Heuven, W. J. B., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language Conflict in the Bilingual Brain. *Cerebral Cortex*, *18*(11), 2706–2716. https://doi.org/10.1093/cercor/bhn030

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequencyinevent-related brainpotentials. *Memory & Cognition*, *18*(4), 380–393. https://doi.org/10.3758/BF03197127

Voga-Redlinger, M. (2005). Amorçage masqué cognate inter-alphabet. Rôle des facteurs morphologique et phonologique. *Revue d'intelligence Artificielle*, *19*(1–2), 407–426. https://doi.org/10.3166/ria.19.407-426

Wang, Y., & Shaw, P. (2008). *Transfer and universality: Collocation use in advanced Chinese and Swedish learner English*. *32*, 32.

Weinreich, U. (1979). *Languages in contact: Findings and problems* (9. print). Mouton.

Weinreich, U., & Martinet, A. (1979). *Languages in contact: Findings and problems* (9. print). Mouton.

Winter, B. (2019). *Statistics for linguists: An introduction using* R. Routledge.

Woutersen, M., Cox, A., Weltens, B., & De Bot, K. (1994). Lexical aspects of standard dialect bilingualism. *Applied Psycholinguistics*, *15*(4), 447–473. https://doi.org/10.1017/S0142716400006871

Abstract

L'influence translinguistique de la L2 sur la L1 : le cas des extensions sémantiques chez des bilingues tardifs espagnol-français en situation d'immersion

L'influence de la L1 sur la L2 est un phénomène largement étudié dans le domaine du bilinguisme, contrairement à l'influence de la L2 sur la L1. Nous proposons l'étude des extensions sémantiques (SE) comme phénomène illustrant ce dernier cas chez les bilingues tardifs espagnols-français. Nous comparons un groupe de bilingues tardifs en immersion en France (durée moyenne d'immersion = 4,6 ans) avec un groupe de monolingues colombiens (naïfs en L2 français). Ils ont accompli une série de tâches expérimentales visant à étudier les SE 1) en production, via une tâche de complétion de phrases, 2) en reconnaissance de mots, avec une tâche de décision lexicale et 3) en compréhension, avec une tâche de jugement d'acceptabilité. Nous faisons l'hypothèse que, contrairement à ce qui est préconisé par la littérature, les SE seraient le résultat d'un transfert au niveau du lexème (concernant l'information phonologique et orthographique des mots) et non exclusivement le résultat d'un transfert au niveau du lemme (au niveau des liens sémantiques et syntaxiques entre deux ou plusieurs lemmes). Nous postulons donc que la compétition formelle entre les langues joue également un rôle au moment du transfert lexical. Notre matériel linguistique oppose deux facteurs psycholinguistiques : la densité du voisinage et la taille de la famille morphologique des mots en L1 et en L2. Ces variables sont présentées dans différentes conditions expérimentales en opposant des associations au niveau de la forme et des liens sémantiques des mots. Les résultats en production orale montrent quelques SE et d'autres transferts lemmatiques et lexémiques. Les résultats en décision lexicale valident notre hypothèse selon laquelle les bilingues seraient sensibles aux stimuli de la L2. Un effet facilitateur apparaît ici pour les stimuli à voisinage dense en L2, révélant ainsi un transfert au niveau du lexème. Ces résultats suggèrent une coactivation des langues chez les bilingues en reconnaissance de mots et en production. En revanche, en compréhension, les bilingues adoptent un mode monolingue dans leurs évaluations des SE. Nous concluons que les SE permettent d'illustrer l'influence de la L2 sur la L1 chez le bilingue tardif espagnol-français.

Mots clés: accès lexical, transfert lexémique, CLI, bilingues tardif, voisinage orthographique