The bilingual cost in speech production

Studies of phonological and articulatory processes

Jasmin Sadat Schaffai

TESI DOCTORAL UPF / 2012
DOCTORAT DE L'UNIVERSITE D'AIX-MARSEILLE

DIRECTORS DE LA TESI:

Dr. Albert Costa

(Departament de Tecnologies de

la informació i les Comunicacions,

Universitat Pompeu Fabra, Spain)

and

Dr. F.-Xavier Alario

(Laboratoire de Psychologie Cognitive,

Aix-Marseille Université & CNRS, France)

Department of Experimental and Health Sciences

Ecole Doctorale 356: Cognition, Langage, Education

UNIVERSITAT POMPEU FABRA

Aix-Marseille université

To Nurollah and Theresa Sadat Schaffai
who speak with my favorite foreign accents

Acknowledgements

The present dissertation results from a long and adventurous journey on speech production and bilingualism. It all started while I was living in Sofia where Elena and Armina introduced me to the exciting field of psycholinguistics. I would have never started this journey without your knowledge and encouragement - thank you! It was during this time when I discovered the work of Albert and Xavier whose research has continuously inspired me since then. Therefore, my foremost thanks go to each of them: I am grateful for their inspiration and insights throughout my studies, and for the opportunity to learn from them. Special thanks to Albert for introducing me to the world of research, for spreading great ideas, for teaching me the importance of being concise and to wise up. Enormous thanks to Xavier! I am very grateful to him for always focusing on the good things, for always keeping a global view, for believing in me, for so many explanations, especially in times when I got lost in statistical analyses and random forests, and for all these hypothetical games like “ok, let’s imagine, I am a word and I have many neighbors…”. Xavier is a very bright person and an amazing mentor! His availability and support over the years have been crucial to the realization of this thesis!

I am grateful to Nuria Sebastián Gallés, Jonathan Grainger and Johannes Ziegler for welcoming me into their labs and for making me feel part of a large research team. Many thanks to Xavi, Sylvia, and Loïc for their help on technical issues, and to Natalia, Neus, Santa, Cristina, Florencia, Jaume, Sonia Testa, and Frederic for their help on so many bureaucratic issues and pages of formalities necessary to realize this international joint project. I would also like to thank Luis Fuentes and his research group for receiving me in their lab in Murcia and for helping me to find monolingual participants.

I would like to acknowledge Sharon Peperkamp, Irene Castellón, Niels Janssen, Christophe Pallier, and Ramón Ferrer i Cancho for agreeing to serve on my committee and for providing me with helpful comments.

Over the years, many people have provided company along the journey of my dissertation. Special thanks to the people I first met in Mundet and Sant Joan: Iva, Mireia, Kristof, Manu, Jordi, Judith, Sara, Cris Baus, Elena, Fatima, Barbara, Bego, Elin, Phil, Volker, Carol, and Alexis. It is always a huge pleasure to hang out with you guys! Thanks to Melania for taking such good care of me when I had just arrived in this new environment. I am indebted to Clara for being there all along my dissertation, for very helpful advice and guidance, and for so much encouragement. Thanks also to all my new lab mates at the Pompeu, especially Paula, Xavi, and Francesca. I would like to thank Bahare for her admirable company through the urban life and for Iranian bonding. Huge thanks to Abi for being such a pure person, and for sharing worlds and chill times.

My journey continued in the remarkable city of Marseille where I was lucky to be surrounded by some very ingenious spirits.

Special thanks to Pauline, Stephanie, and Mauricio for spending some really great times inside and outside of the lab. Thanks to Rita for coming to the lab when everybody else had left and for many enjoyable moments. I would like thank Svetlana and Dash for their delicious brownies, for insightful views on life, and for making me feel optimistic about the next "generation" of researchers working with Xavier! Thanks to Anaïs and Carlos for good vibes and for making me discover the neuroscience of language production with their intra-cranial recordings. I am deeply grateful to Laure for her positive multilingual support, her vivid spirit, her delicious cooking, and the precious moments shared in my little garden and outside the lab. Heartfelt thanks to Delphine for her incredible energy and optimism, for constantly reassuring me and organizing my time schedule, and for making the work in our loft a truly enjoyable experience. I thank Sylvain for his kindness and for always understanding my worries about analysis problems. Many thanks to Isa for Columbian cheerfulness and compassion. I would also like to acknowledge all my other labmates who made me enjoy the daily life in the lab: Maria, Catarina, Aurélie, Leni, Michael, Seb, Pierre-Yves, Anahid, Romain, and many more. Thanks to Nico for his incredible faith to leave me the keys to his apartment in Marseille, although we just had met for 30 minutes. I am grateful to Olivier for fun times and always inspiring hope, even on a Saturday night at the PMU in Simiane. Many thanks to Elena for suffering with me through the nightmares of CheckVocal, and for her support when work seemed endless. Thanks to Stan for showing me how to efficiently write a thesis and to enjoy your surroundings. I am deeply grateful to Gabriel and Selva, my Marseillan godparents, for always welcoming me in their place, and for sharing so many beautiful moments. Immense thanks to Brandon for his delightful company, for providing grounding, for his support and care, during good and bad times, and for making me happy!

Many people from far away have also contributed in their ways to this dissertation. In particular, I would like to acknowledge Isa for being there each time after my Murcia stays, who cared for and spoiled me, and for being such an energetic and faithful companion! Big thanks to Marie for sharing crucial moments, for always inspiring my sense of exploration, and for seriously making me consider writing my thesis in the hinterlands of the Central African Republic. Last but not least, I would like to thank my incredible family - for their curiosity, for their craziness, for constantly reminding me of the most important things in life, for respecting and supporting my choices over all these years, for their positive energy, for teaching me the importance of friendships and of being there for each other, and for making impossible things possible. You are the best - Thank you so much!

The work of the present dissertation has been supported by several grants from the German, Catalan, French, and Spanish governments, for which I am very grateful.

Abstract

The main objective of this dissertation is to examine the consequences of bilingualism on speech production. Previous research has shown that bilingual speakers experience a cost compared to monolinguals in a variety of linguistic experiments. We investigated the origins of the bilingual cost by exploring influences of particular variables such as phonological similarity. Moreover, we investigated the scope of the bilingual cost by assessing speech performance, focusing on articulatory durations and noun-phrase production. We provide evidence that increased phonological similarity among words within one language slows speech, whereas increased phonological similarity across translations helps bilinguals to overcome the bilingual cost. In addition, our results show that the bilingual cost generalizes to articulatory durations and noun-phrase production. The current dissertation provides a more specific understanding of speech processing at phonological and articulatory stages in mono- and bilinguals, and extends our knowledge on the bilingual cost in speech production.

Resumen

El objetivo principal de esta tesis es examinar las consecuencias del bilingüismo en la producción del habla. Estudios anteriores han demostrado que los hablantes bilingües presentan una mayor dificultad que los monolingües durante la realización de diferentes experimentos lingüísticos. Investigamos los orígenes del coste del bilingüismo, explorando los efectos de variables como la similitud fonológica. Además, investigamos el alcance del coste, evaluando las duraciones articulatorias durante el habla y la producción de sintagmas nominales. Mostramos que una mayor similitud fonológica entre palabras de una lengua ralentiza el habla, mientras que una mayor similitud entre traducciones ayudó a los bilingües a superar el coste. Finalmente, demostramos que el coste bilingüe se extiende a las duraciones articulatorias y a la producción de sintagmas nominales. Esta tesis aporta nuevas evidencias acerca del efecto que la similitud fonológica tiene sobre la producción del habla y proporciona un conocimiento más específico sobre cómo el bilingüismo influye durante las últimas fases del procesamiento del habla.

Preface

Bilingualism has become an increasingly common phenomenon in our globalized world. Nowadays, more than 60 percent of the world's population is growing up bilingual and there are more speakers using English as a second than as a first language (Sampat, 2001). Accordingly, research on the consequences of bilingualism has experienced an immense growth during the last two decades. Although the first studies on bilinguals, dating back to the first half of the last century, almost uniformly reported negative consequences on cognitive abilities (e.g., Pintner, 1932; Saer, 1923; Smith, 1939), research throughout the last decades has provided a more precise and differentiated picture of the implications of bilingualism. Many recent studies have reported beneficial effects, either in non-linguistic or meta-linguistic domains, of mastering two languages. However, a few recent studies also found negative consequences in the linguistic performance of bilinguals, referring to this phenomenon as the bilingual cost or bilingual disadvantage. In this dissertation, we will refer to the bilingual cost when linguistic performance differs between mono- and bilinguals. However, it is important to emphasize that this cost refers to any different linguistic behavior between these two speaker groups (mono- and bilinguals), rather than a true drawback in communication. Despite any linguistic disadvantages associated with bilingualism, it is evident that knowing more than one language represents a true advantage in the ability to interact with more people and to immerse in different cultures. Nevertheless, these performance differences require a detailed empirical investigation to provide an appropriate description of speech production in bilinguals and to advance our knowledge on this omnipresent phenomenon.

In the current literature, bilinguals are defined as individuals who use two languages in their everyday life either actively (comprehending and producing) or passively (only comprehending, e.g., Grosjean, 1982). This definition relaxes the popular notion of a bilingual who performs almost native-like in both languages. In the current dissertation, however, we will make use of the latter definition of bilinguals given the language community under study. In Catalonia, Catalan and Spanish are both official languages and education in the two languages is offered to children beginning at the kindergarten level. The current education system requires that at the end of primary school, children are able to read, write, speak, and understand both Catalan and Spanish. Furthermore, in many families both languages are spoken. Thus, growing up in Catalonia usually means that the two languages are present in everyday life from very early on, and that bilinguals have high levels of proficiency in both languages. Most studies have focused on bilingualism’s effects on a later-learned second language. However, how bilingualism impacts speech production when the two languages are acquired nearly simultaneously and at an early age has received little attention up to now. The current dissertation focuses precisely on speech performance in such early and highly-proficient bilinguals.

A particular characteristic of Spanish-Catalan bilinguals is that they speak two Romance languages that share many important properties. Approximately 70% of the vocabulary can be considered to have similar sounding translations. Additionally, grammar and syntax relate closely across the two languages. Still, significant differences are found between Catalan and Spanish in their phonological repertoires. Spanish has fewer vowels than Catalan and some consonants exist only in Spanish, while others are specific to Catalan. Therefore, these two languages provide an ideal model for examining the role of phonological processes in language production of bilinguals. One central aspect of the current dissertation is the impact that the interacting phonological systems within a bilingual may have on the ability to master two languages.

The present research aims to investigate two important facets of the bilingual cost: First, we will explore the origins of the established bilingual cost in single word production and whether they may relate to particular variables, such as phonological similarity. Second, we will investigate the scope of the bilingual cost by assessing performance in articulatory durations and in speech contexts beyond single word production. The goal of the present dissertation is to provide a deeper understanding of the phonological and articulatory processes involved in the speech of mono- and bilinguals.

Before presenting the experimental section of this dissertation, the most important aspects of language production for mono- and bilinguals are introduced in Chapter 1. This chapter forms the theoretical background for the dissertation and describes the outstanding questions of speech production in mono- and bilinguals.

Table of contents

List of figures

List of tables

1. GENERAL INTRODUCTION: Current description of the bilingual cost in speech production

Research on speech production in bilinguals builds on a large body of knowledge acquired from monolingual speech production. At the same time, speech production in bilinguals is deemed special, because it needs to accommodate for the fact that the same speaker can successfully handle two different language systems for communication. This raises additional questions for speech processing relative to monolinguals. Some of the main issues in speech production in bilinguals are whether the non-spoken language is co-activated, and if so, to what extend does this co-activation impact on the spoken language?

Many studies have provided strong evidence that the two languages of a bilingual interact during speech production. In addition, several studies have reported a bilingual cost in linguistic experiments when compared to monolingual speakers. These findings imply that the fact of having two interacting language systems may have direct consequences on the linguistic performance of bilinguals. The present dissertation aims at investigating the origin and the scope of such a bilingual cost by exploring phonological and articulatory processes. Our main focus will be on the interplay between the two bilingual language systems at these late stages of speech processing, and the consequences they may entail for performance.

In what follows, we will present an overview of the core ideas from monolingual language production that can be transferred to bilinguals. Then, we will turn to the particularities of the bilingual language production system and describe the different levels at which inter-language influences may occur. We will put a special focus on phonological and articulatory processes in the speech of bilinguals. Moreover, we will review the previous evidence on the bilingual cost and present the current explanations regarding its origin. Finally, we will point out some limitations about our knowledge on the bilingual cost beyond single word production.

1.1 Speech production in monolinguals

Language production is an essential part of our daily interactions. The apparently simple act of speaking contrasts with the complex and unperceived cognitive processes that need to be coordinated before articulation can take place. The main theories established in the literature distinguish between three levels of mental representations involved in speech production (Caramazza, 1997; Dell, 1986; Levelt, 1989). The first step to produce a single word involves the retrieval of a concept that matches the intended idea we want to communicate (e.g., the idea of a feline pet). Subsequently, the word corresponding to this idea has to be retrieved from the mental lexicon, along with its grammatical properties (e.g., “cat”, noun, neuter). Once the word has been accessed, the phonological information belonging to it has to be retrieved (e.g., /k/, /ae/, /t/). This process will allow sending the respective information to the articulatory organ and to elicit the speech signal that corresponds to the idea to express. These are the three major steps that have to be undergone to produce a single word.

Articulation of several ideas and words in sentences entails even more processes than the previously described three stages. For example, several words are combined according to specific rules of the spoken language, such as number and gender agreements between the individual words which have to be computed on the fly (see section 1.5 for speech production beyond single words). Despite the complexity of the language production system, it is surprising how efficient healthy individuals are in producing speech and how rarely they commit errors. Although the proportion of speech errors in overall speech is not very high, they are highly informative regarding activation flow within the language production system. For example, slips-of-the-tongue in which two words with highly related meanings are substituted alludes to the dynamics that underlay word selection (e.g., the intention to say "fork" results in the production of "spoon"). The occurrence of such type of errors hints to parallel activation of the word intended for communication and additional words that are closely related to it.

The idea of co-activated related items around the intended concept represents one of the core principles for the dynamics in the speech production system: spreading activation (Collins & Loftus, 1975). As mentioned above, when a speaker wants to produce a word like “cat”, the semantic information related to the intended concept is activated, but in addition semantically related information is also activated (e.g., concept of related domestic animals like dog, the action to meow, etc.). Importantly, spreading activation also applies to the activation flow between the successive representational levels of speech production. The main theories agree on the assumption that the activated semantic information (i.e., target and related concepts) spreads proportional activation to the corresponding lexical nodes (e.g., the concept of cat activates the lexical representation of cat, but a related representation like dog is also activated to a smaller extend). However, speech production models differ with respect to spreading activation flow after the lexical level. Discrete two-stage processing models (Levelt et al., 1991; Schriefers, Meyer, & Levelt, 1990) assume that only the selected lexical target node sends activation to its phonological representation, whereas unidirectional cascaded-processing models (Humphreys, Riddoch, & Quinlan, 1988; Peterson & Savoy, 1998) and interactive activation models (Dell, 1986) agree that the activated semantic related information spreads proportional activation up to the corresponding phonological representations. The latter two models then differ in their assumption of possible feedback connections: Purely cascaded models do not allow phonological information to sent activation back to the lexical level, whereas interactive activation models allow such interactions through bidirectional links (see Figure 1 for a schematic architecture of an interactive spreading activation model of speech production).

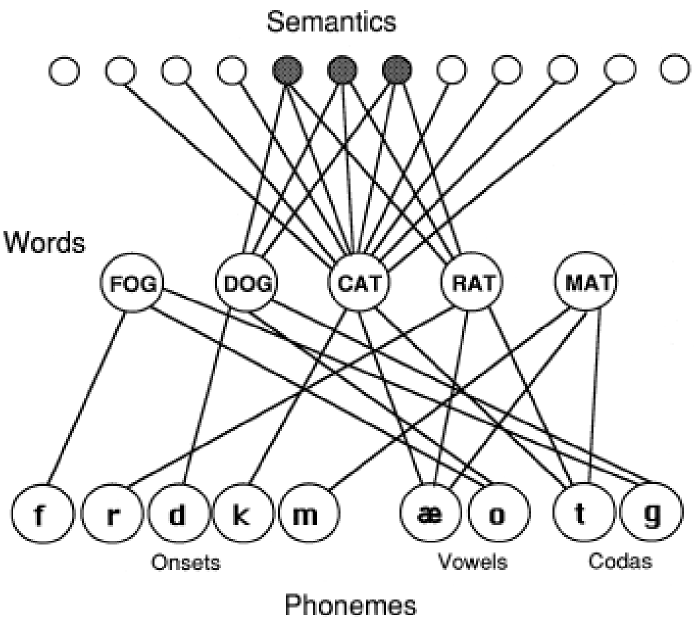

Figure 1: Illustration of an interactive spreading activation model of speech production with three layers of representations: semantic features, lexical words, and phonemes. Connections between the layers are bidirectional and excitatory (the aphasia model taken from Dell, Chang, & Griffin, 1999).

A second core principle in speech production is the assumption that word selection depends on the activation level of lexical representations and is accomplished by competition between co-activated words. That is, the time required to select the intended word is a function of the difference in activation levels between the intended and the related lexical representations. As long as several lexical nodes are highly activated, they all compete for selection, and retrieval of the target word is slowed down. On the contrary, if not many related words are co-activated, selection of the intended word can proceed unrestrained (e.g., La Heij, 1988; Levelt, Roelofs, & Meyer, 1999; but see Finkbeiner & Caramazza, 2006; Janssen, Schirm, Mahon, & Caramazza, 2008; Mahon, Costa, Peterson, Vargas, & Caramazza, 2007, for an alternative explanation).

In most studies of language production, the idea of word selection among co-activated words is discussed with respect to selection among semantically similar words. The current dissertation, however, puts a special focus on phonological processing, and thus on how words that are phonologically related to each other may impact on word selection. Although speech production proceeds from meaning to articulation, there are studies indicating that form properties nevertheless influence word selection. One such piece of evidence for such influences is the mixed error effect (e.g., Dell & Reich, 1981; Harley, 1984; Martin, Gagnon, Schwartz, Dell, & Saffran, 1996). Mixed errors share both semantic and phonological similarity with the intended word (e.g., "rat" for "cat"). Overall, they occur more often than predicted based on rates of purely semantic (e.g. "cat" for "dog") or phonological errors (e.g., "cab" for "cat"). In this context, the interactive activation model by Dell (1986) is of high importance (see Figure 1). This type of model allows interactive feedback between the phonological and lexical level and thus provides a straightforward explanation to phonological influences on word retrieval. We will revisit this issue later when discussing the effects of phonological similarity on speech production in mono- and bilinguals.

A widely used laboratory task to study language production processes is the picture naming task. This task consists of presenting an object picture to the participants and to ask them to name the depicted object as fast and as accurate as possible. Picture naming tasks are mostly used to elicit to single word production. Although this is an oversimplification of everyday language use, it nevertheless engages the basic processes involved in speech production. To verbalize a response to a depicted object, the participants have to recognize the picture and access the concept to name. As described earlier, following concept selection, lexical, grammatical and phonological information has to be retrieved before the response can be articulated.

One of the aims of previous research on language production has been to relate various characteristics of both the picture and its word label with the particular stages involved in word production. These studies allowed researchers to determine the most important predictors of speech production and to clarify the extent to which the different levels of processing are involved in speech production (see Figure 2 for a schematic overview).

Over the years, assessment methods for the determinants of naming performance have been continuously improving, yet the localization of the effects of each of the variables in speech processing is still an ongoing field of research. Since many years, a

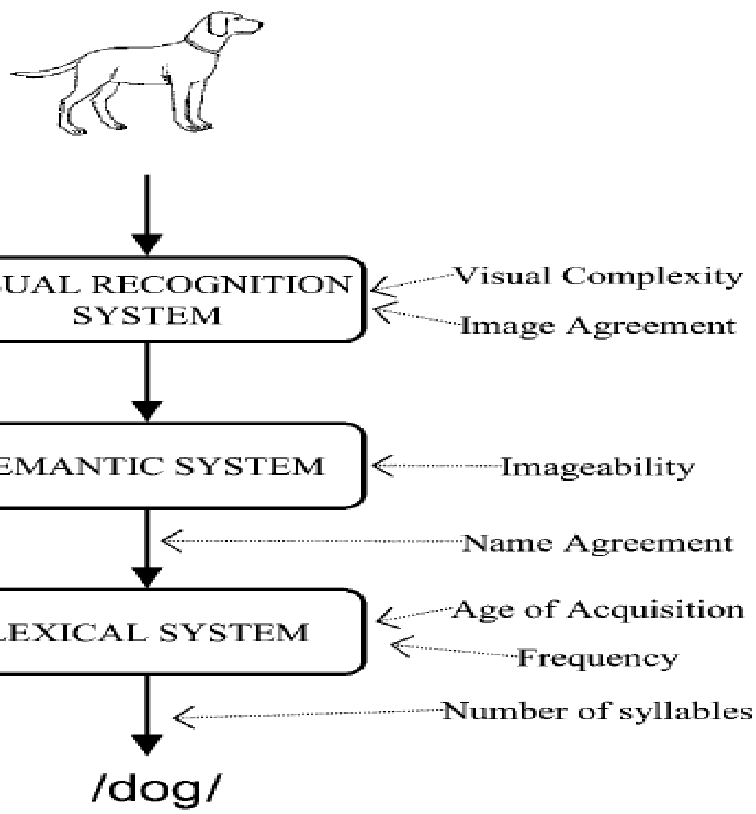

Figure 2: Relationship between different variables influencing word production and their suggested loci in a model of picture naming (taken from Alario, Ferrand, Laganaro, New, Frauenfelder, & Segui, 2004).

large bulk of psycholinguistic research aimed at revealing the most important variables influencing picture naming speed (e.g. Barry, Morrison, & Ellis, 1997; Cycowicz, Friedman, Rothstein, & Snodgrass, 1997; Ellis & Morrison, 1998; Vitkovitch & Tyrrell, 1995). Among this research, some of the most important findings are for example the seminal study by Oldfield & Wingfield (1965) that reported a negative linear relation between the time to name a picture and the logarithm of word frequency. Carroll & White (1973) were the first to show that the estimated age of acquisition of words was an important predictor of picture naming speed in adults. The studies of Lachman and colleagues explored the importance of uncertainty measures (i.e., name agreement and number of alternative names) on naming performance (Lachman, 1973; Lachman, Shaffer, & Hennrikus, 1974), and the work of Snodgrass and colleagues (Snodgrass & Vanderwart, 1980; Snodgrass & Yuditsky, 1996) prepared further promising ground for investigations of the cognitive processes that take place during picture naming.

One main focus of the present dissertation are the phonological processes during speech production. Thus, special attention will be paid to the effects of variables indexing processes that occur at this late stage of word retrieval. Previous studies have shown that phonological properties of the word to utter affected lexical retrieval both in mono- and bilinguals (e.g., Costa, Caramazza, & Sebastián-Gallés, 2000; Colomé, 2000; Colomé & Miozzo, 2010; Ferreira & Griffin, 2003; Harley, 1984; Martin et al., 1996; Rapp & Samuel, 2002; Vitevitch, 2002). However, note that the definition of this variable differs according to the speaker group of interest. For monolinguals, phonological similarity usually captures the amount of phonological similarity of a given word relative to other words within one language (i.e., phonological neighborhood density), whereas for bilinguals, it usually refers to phonological similarity across translations (i.e., cognate status). The findings of influences from phonological properties on word retrieval support interactive models of speech production (Dell, 1986). These kind of models can readily account for such effects through their feedback connections from the phonological to the lexical level (see Figure 1).

In particular, the effect of phonological neighborhood density has received a lot of attention in the monolingual literature. A word that sounds highly similar to other words in that language is said to come from a dense neighborhood, whereas words that have only few similar sounding words are said to come from sparse neighborhoods. In an influential study by Vitevitch (2002), the author showed that words from dense neighborhoods were named faster than words from sparse neighborhoods. Vitevitch attributed this finding to the increased phonological overlap of the words from dense neighborhoods that facilitated lexical retrieval compared to words from sparse neighborhoods. However, this evidence and its explanation have not gone unchallenged. In the experimental section, we will further introduce the role of phonological neighborhood density for monolingual speech production. In what follows, we will turn to speech production in bilinguals and point out how phonological similarity across translations has been shown to be influential.

1.2 Speech production in bilinguals

In the case of speech production in bilinguals, current models assume that the two languages share a common semantic representation, while two different lexical representations (one for each language) are associated to it (De Bot, 1992; Finkbeiner, Nicol, Greth, & Nakamura, 2002; Green, 1986; Green, 1998; Kroll & Stewart, 1994; Poulisse & Bongaerts, 1994; but see Paivio & Desrochers, 1980; Van Hell & De Groot, 1998, for language-dependent semantic representations). The question that arises is how do lexical access and selection operate when two lexical representations are associated to the concept a bilingual speaker wants to express? A large amount of research on bilingualism addressed the question of potential inter-language interplay at the lexical (i.e., selecting the right word in the intended language) and at the phonological level (i.e., selecting the right sounds of the intended language). In addition, recent research has reported costs in the speech of bilinguals compared to monolinguals. That is bilingual speakers have been shown to perform worse in several linguistic laboratory tasks than monolinguals. Thus, it appears evident that the interactions between the two languages of a bilingual may be at the source of the linguistic costs that is associated with bilinguals. However, it remains unclear whether the bilingual cost originates from inter-language influences at the lexical, or phonological level, or both. Below, we will briefly introduce the most popular accounts to explain how bilinguals select the right word in the right language, with a special emphasis on the Inhibitor Control Model (Green, 1998).

Given that there are two fully specified language systems in a bilingual, the easiest solution for successful communication in one language would be to turn off the unintended language – as if a bilingual is two monolinguals in one person. However, that this is not the case has been shown by several studies reporting influences from the unintended language on lexical and phonological representations of the intended language (e.g., Colomé, 2001; Costa et al., 2000; Costa, Miozzo, & Caramazza, 1999; De Bot, 1992; Green, 1986; Kroll & Stewart, 1994; for a review see Costa, 2005). Despite these inter-language interactions, the number of intrusions from the unintended language is very scarce (Poulisse & Bongaerts, 1994), and bilinguals effectively manage communication in each of their languages.

Two different accounts can explain how bilingual speakers select the right word in the intended language without language intrusions. One of the most prevalent models in bilingual research is the Inhibitory Control Model (Green, 1998). It is based on the idea of interference between the two activated languages and states that lexical selection is language non-specific. This means that in bilinguals the activated representations of both the intended and unintended language would compete with each other. To achieve successful language production, an overall and active suppression of the unintended language representations would be required. Thus, lexical representations in the unintended language would not enter into competition, and lexical selection would be restricted only to words from the intended language. Importantly, this account claims that the process of lexical selection is fundamentally different between mono- and bilingual speakers. For the purpose of the current research, it is important to note that the Inhibitory Control Model has also been used to explain the bilingual cost in speech production. We will come back to this explanation when discussing the possible origin behind this cost (see section 1.4).

According to an alternative proposal by Costa and colleagues, lexical selection is language-specific (Costa & Caramazza, 1999; Costa et al., 1999). This means that during lexical selection only the activated lexical representations of the intended language are considered. Lexical representations of the unintended language would be ignored, and thus would not enter into competition. Following this assumption, lexical selection would be accomplished directly within the intended language, and so word selection processes between mono- and bilinguals would not be different.

Despite many attempts to understand how bilinguals overcome intrusions from the unintended language, none of the two proposals can fully account for the wide range of empirical findings in bilingual speech production (for a recent discussion see Runnqvist, Strijkers, Alario, & Costa, 2012). Thus, explanations relying on inhibitory control as the origin for a bilingual cost remain disputed.

The focus of present dissertation is on bilingualism effects on phonological and articulatory processes and how they relate to the bilingual cost. In what follows, we will describe inter-language influences at these late stages of language processing, discussing cognate effects, foreign accent, and bilingual articulatory durations. In particular, we will outlay how inter-language interactions at the phonological and articulatory level may affect the presence of a performance cost for bilinguals.

1.3 Phonological and articulatory processes in the speech production of bilinguals

More obvious evidence on inter-language interference in bilinguals comes from processes anchored at the phonological and articulatory level. As mentioned above, a large body of evidence indicates that during speech production in bilinguals the lexical and phonological representations of the two languages interact. Thus, inter-language influences can emerge during phonological processing in bilinguals. For example, in a study by Colomé (2001), bilinguals were presented with object pictures and were asked to decide whether a specific phoneme was the initial phoneme of the picture's name or not. The results showed that the participants' responses were faster for object names that began with the same phonemes in both languages (e.g., "gat" and "gato" ["cat" in Catalan and Spanish]) than those who had differing onsets (e.g., "gos" and "perro" ["dog" in Catalan and Spanish]). Importantly, this effect was absent in a monolingual control group. Colomé's results indicated that the activation of the overlapping phonemes across the two languages converged, thus facilitating decisions for situations in which onsets overlapped as compared to different onsets across languages (see also Colomé & Miozzo, 2010).

Relatedly, Costa et al. (2000) investigated phonological overlap across translations. In their study, bilingual participants had to name pictures that were manipulated for phonological overlap across translations. Bilinguals were faster in naming the picture names that had high phonological overlap across translations (i.e., cognates, e.g., tomato – tomate [in Spanish]) as compared to translations that did not share many sounds across the two languages (i.e., non-cognates, e.g., apple – manzana [in Spanish]). Again, the monolingual control group did not show any difference in naming times for these two types of pictures (see also Hoshino & Kroll, 2008; Ivanova & Costa, 2008) .

Similar performance differences between cognate and non-cognate words were also encountered in studies of bilingual language comprehension (e.g., Dijkstra, Grainger, & Van Heuven, 1999). In addition, cognate words seem to be easier to learn, recall and are more resistant to retrieval failures than non-cognates in patients and healthy bilinguals (e.g., Costa et al., 2012; Gollan & Acenas, 2004; Kohnert, 2004; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009; Van Hell & De Groot, 1998). Moreover, cognate effects have been shown to add up across several non-target languages as in the case of triple cognates (i.e., words sharing sounds across the three languages of a trilingual being processed faster than bilingual cognates, Lemhöfer, Dijkstra, & Michel, 2004).

The findings on cognate effects in word production of bilinguals suggest that the lexical items of the two languages spread activation to their common phonological representations. The most prominent explanation regarding the origin of the cognate effect relies on the interactive nature during lexical access (e.g., Costa et al., 2000; Costa, Santesteban, & Caño, 2005; but see also Sánchez-Casas & García-Albea, 2005, or Van Hell & De Groot, 1998, for alternative explanations at the morphological and conceptual level respectively). Costa and colleagues postulate representational overlap at the phonological level for cognates compared to non-cognates, and interactivity between the phonological and lexical level. Accordingly, the convergence of activation of the two languages at the phonological level will speed up lexical retrieval for cognates when compared to non-cognates (see Figure 3).

In addition, several studies showed that cognate effects are larger for bilinguals speaking in their second language (L2) than their first and dominant one (L1; Costa et al., 2000; Ivanova & Costa, 2008). The explanation for this difference refers to the supposedly stronger links between semantic and lexical nodes in the L1 compared to the L2 of a bilingual (Kroll and Stewart, 1994). When a bilingual produces a cognate in L2, the strongly activated lexical representation in L1 spreads activation to its phonological segments, thus facilitating the retrieval of the overlapping phonological segments. On the contrary, when bilinguals produce a cognate in L1, the activation sent by the lexical representation in L2 is weaker than in the previous situation. Thus, retrieval of a cognate word in L2 benefits more from the additional activation received from a strong L1 than vice versa (Costa et al., 2000).

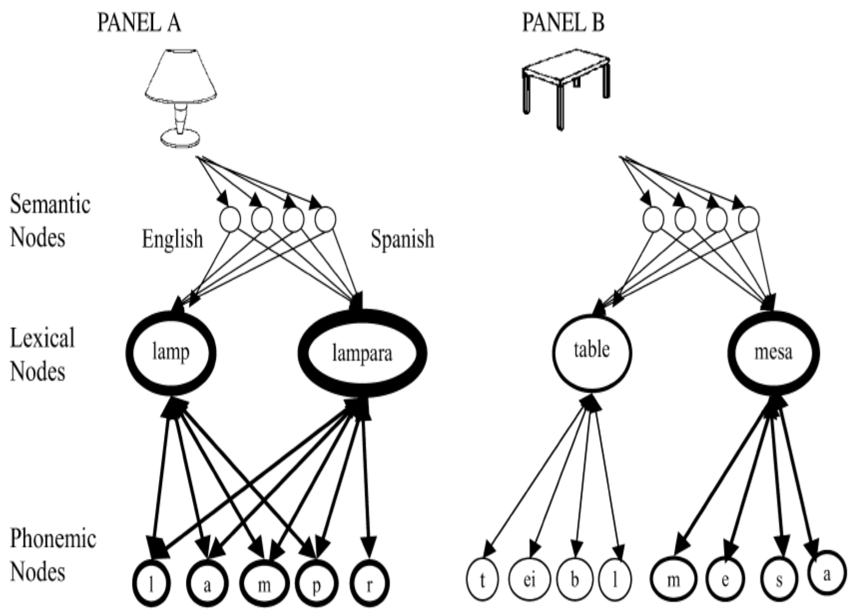

Figure 3: Illustration of an interactive spreading activation model for a Spanish-English bilingual in the case of cognate (Panel A) and non-cognate word production (Panel B; taken from Costa et al., 2005).

Recently an alternative proposal has been provided to explain the origin of cognate effects in the speech of bilinguals. In a study by Strijkers, Costa, & Thierry (2010) using event-related-potentials, they found divergent amplitudes at a similar point for the effects of lexical frequency and translation overlap. Their results indicated that these two effects may stem from relatively similar and early processes during lexical access. The authors suggested that cognate effects could be understood as a frequency effect in disguise (see also Sandoval, Gollan, Ferreira, & Salmon, 2010). Given that the shared phonological segments of cognates feedback activation to the lexical level, cognates are supposed to develop higher lexical activation levels over time as compared to non-cognates. That is, each time a cognate word is used, irrespective of the language of response, it would augment the frequency of use of this word. As a result, the cognate effect is suggested to emerge already at the lexical level, because activation levels of the representations between cognate and non-cognates differ.

One assumption common to the two explanations of cognate effects is that similar phonological segments between the two languages of a bilingual are shared. Yet, little is known about how bilinguals process language at the phonological level. Evidence suggesting that the phonological and phonetic representations are shared and that the two languages influence each other mainly comes from studies investigating speech accents.

When interacting with bilinguals, it is often times striking to hear their foreign accent or to see how much effort they need to put into pronunciation processes. There is ample evidence that L2 speakers show influences from their native language system on the newly learned one (e.g., Flege & Eefting, 1987; Flege, Schirru, & MacKay, 2003). Clearly, one important factor underlying such interactions is the age at which a second language is learnt. The younger a speaker is when learning a second language, the less accented is his or her speech (e.g., Flege, Munro, & MacKay, 1995a; Flege, Munro, & MacKay, 1995b; Munro, Flege, & MacKay, 1996). Therefore, a great deal of the discussions about foreign accents focused on the idea of a critical period for language acquisition. According to this view, the ability to speak a foreign language without a foreign accent seems to be constrained by a loss of plasticity arising from neural maturation (e.g., Lenneberg, 1967).

However, several studies have documented the existence of late L2 speakers whose speech performance was judged as being similar to natives (e.g., Abu-Rabia & Kehat, 2004; Bongaerts, 1999). These results indicate that an early start of foreign language learning is not mandatory to obtain native-like proficiency. Moreover, it has been shown that age is not the only determinant of speech accents, and that there are additional factors relevant to it. For example, in a study by Flege, Frieda, and Nozawa (1997), the authors showed that the perception of an L2 speech accent depended on the amount of L1 usage of the bilinguals (see also Guion, Flege, & Loftin, 2000; MacKay, Flege, Piske, & Schirru, 2001; Piske, MacKay, & Flege, 2001; Piske, Flege, MacKay, & Meador, 2002; Yeni-Komshian, Flege, & Liu, 2000). Thus, more recent proposals regarding the deviant pronunciations of bilinguals capitalize rather on the interactions between the phonetic representations of a bilingual than on a critical period. Specifically, current explanations focus on how prior experience of perceptual processing influences the representations of later acquired speech sounds across the life span.

From a developmental perspective, Kuhl (2000) suggested that perceptual representations of phonemes are stored in the infant's memory which in turn will guide subsequent native speech production. Kuhl and colleagues claim that initial coding of native-language "maps" will interfere with the learning of later learnt "maps" of a foreign language. In a study by Iverson et al. (2003), they tested how early language experience interfered with the later acquisition of phonemes of a foreign language (e.g., the perception of the English /r/-/l/ contrast for native Japanese speakers that do not distinguish between these two phonetic categories). Their results showed that early language-specific perceptual processing altered the salience of phonetic categories of a foreign language. Japanese native speakers differed in sensitivity at low-level perceptual dimensions from English natives, and thus were insensitive to distinguish the English /r/-/l/ boundary. This means that early perceptual experience influenced the perception of non-native sounds. More importantly, these early acquired perceptual categories are supposed to shape the way how sounds are produced later on (see also Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004). However, for bilingual infants who learn two languages at the same time, two different "mappings" are acquired simultaneously and thus interference will be minimal (Kuhl, 2000).

In the same vein, Flege proposed speech learning model that focuses notably on the acquisition of a second language (Flege, 1995). One main assumption of this model is that the bilingual phonetic systems are shared between the two languages and necessarily influence one another in both directions. Importantly, this view of a common phonological space for the two languages of a bilingual predicts that even L1 speech differs from monolinguals, because of the mere presence of an additional language system. Accordingly, phonetic categories of L1 and L2 change (dissimilate and assimilate) depended on language dominance and usage during speech production in bilinguals. Flege and colleagues claim that the degree of similarity between L1 and L2 phonetic categories determines whether merged or new categories are established for the sounds of the two languages. If the two sounds are very similar, they will be assimilated into a single already established category (see Flege, 1987; MacKay et al., 2001). This will occur relatively more often in the case of late L2 speakers and thus explains the frequently observed foreign accents in L2. On the other hand, if the two sounds are not close enough to be accurately merged, they will dissimilate and create phonetic categories with slightly exaggerated characteristics to distinguish them (see Flege, Schirru, & MacKay, 2003). The dissimilation process is thought to occur because bilinguals strive to maintain phonetic contrasts among all elements in their shared phonological space. Overall, these two processes at the phonetic level provide a straightforward explanation of accented speech, apparent in both L1 and L2 speech.

Several studies have addressed these phonetic inter-language influences by measuring acoustic properties. In a study on voice-onset-times (VOT; i.e., the time between air release and the moment at which the vocal cord starts vibrating in consonant production), Flege (1987) observed that the realizations of /t/ of English-French bilinguals in both L1 and L2 differed from monolinguals for each of these languages. That is, the bilingual VOTs in English were shorter than those from English monolinguals, and the bilingual VOTs in French were longer than those of French monolinguals. Similar results of changed VOTs for bilinguals' L1 and L2 when compared to monolinguals have been reported in several studies, ranging from early to late bilinguals, even after brief immersion periods (e.g., Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Flege & Eefting, 1987; Fowler, Sramko, Ostry, Rowland, & Hallé, 2008; Sancier & Fowler, 1997; but see also Mack, 1989). Additionally, Amengual (2012) recently reported VOT differences in the case of cognate production of bilinguals compared to monolinguals. He assessed the production of /t/ in several groups of Spanish-English bilinguals and one group of Spanish monolinguals. Bilinguals speaking in Spanish produced /t/ in cognates with longer VOTs (more English-like) when compared to non-cognates. His results suggest that the phonetic representations of the two languages of a bilingual influence one another, and importantly that these phonetic alterations are enhanced by the cognate status of the word to produce. This finding implies that articulation is sensitive to factors like conceptual overlap in addition to the mere inter-language interactions in phonetic categories. This issue will be further discussed when revisiting the origin of the bilingual cost in articulation. Altogether, studies on VOTs indicate inter-language influences and changes in the phonetic representations of a bilingual due to the experience with an additional language system.

A few studies also addressed the nature of phonological representations in bilingual production by means of experimental manipulations. Roelofs (2003) showed that bilingual naming latencies decreased for the repeated use of phonemes within as well as across languages. This result suggests that for the late bilinguals tested in his study shared phonemic representations were repeatedly used to produce words in the two languages. Reladetly, Alario, Goslin, Michel, & Laganaro (2010) manipulated syllabic frequency in early and late bilinguals' speech production. Their results showed that while both speaker groups were affected by the syllable frequency of the spoken language, only the speech of late bilinguals was also affected by the syllabic frequency of the unspoken language. The authors suggested that syllable representations in early bilinguals are separate, whereas late bilinguals use the same representation for speech in the two languages. This finding provides further evidence of why early bilinguals are better in approaching monolingual pronunciations when compared to late bilinguals. It is also in line with previous explanations of how early perceptual experience in bilinguals may shape native phonetic representations and influence later language production.

The last dimension of speech that can be investigated for bilingualism effects is the level of speech-motor control and articulation. Overall, studies on articulatory durations concordantly reported slower speech rates in late L2 speakers when compared to native speakers or early bilinguals (e.g., Chakraborty, Goffman, & Smith, 2008; Guion, Flege, Liu, & Yeni-Komshian, 2000; Mackay & Flege, 2004; Munro & Derwing, 1995; Nissen, Dromey, & Wheeler, 2007; see also Simmonds, Wise, Dhanjal, & Leech, 2011, for a recent review on motor-sensory aspects of bilingual compared to monolingual articulations; and Jones et al., 2012, for evidence of increased articulation demands in bilinguals as compared to monolinguals). Two explanations for the bilingual cost during articulatory processes are plausible: First, as described above, articulatory costs could be attributed to the phonetic inter-language influences during speaking (e.g., Guion et al., 2000). On the other hand, several studies have shown that the ease with which word retrieval is achieved may have an impact not only on the speed with which utterances are initiated (onset latencies) but also on the corresponding articulatory durations (Balota, Boland, & Shields, 1989; Kawamoto, Kello, Jones, & Bame, 1998; Kawamoto, Kello, Higareda, & Vu, 1999; Kello, Plaut, & MacWhinney, 2000; Kello, 2004, but see Damian, 2003). Hence, if bilinguals experience a cost already at earlier stages of lexical retrieval, one can assume that it could percolate to articulatory durations. Whether articulatory durations are influenced by phonetic interactions or carry-over effects from lexical access remains unclear. Nevertheless, it motivates an investigation of articulatory processes in speech production of bilinguals.

From the above review, it is clear that the two language of a bilingual interact, and that these influences are especially apparent at late stages of processing (i.e., phonetic, phonological, and articulatory levels). In other words, experience with an additional language impacts on the way speech sounds are stored in memory and produced overtly. Thus, there is good reason to assume that these interactions will contribute to a performance difference between mono- and bilinguals (see also Indefrey, 2006; Hanulová, Davidson, & Indefrey, 2011). Although most of the literature focused on such performance differences in late L2 speakers, it is important to establish whether these interactions are evident also in the case of L1 speakers. The current dissertation addresses this gap by investigating the contributions of phonological and articulatory processes to the bilingual cost by always including a group of bilinguals using their first-learnt and dominant language. This will provide a more detailed picture of the bilingual cost and will establish whether such costs are the result of having two interacting language systems, or a more direct consequence of a lack of experience (which is typically the case in L2 speaking). In what follows, we will introduce previous evidence of the bilingual cost in speech production together with the three currently proposed accounts explaining its origin.

1.4 Current explanations of the bilingual cost

Studies on speech production in bilinguals show ample evidence for inter-language influences at different levels of speech processing. These interactions are supposed to provoke performance differences between the speech production of mono- and bilinguals, which is characterized by poorer linguistic performance of the bilinguals (i.e. bilingual cost). As mentioned in the previous section, the bilingual cost is most obvious when performance of L2 speakers is compared to L1 bilinguals or monolinguals. Studies over a wide range of linguistic laboratory tasks (e.g., picture naming, sentence repetition, tongue twister elicitation) have shown that speaking in a non-native language is overall slower, less accurate, and takes longer than in a native language (e.g., Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007; Gollan & Goldrick, 2012; Guion et al., 2000; Ivanova & Costa, 2008; Kohnert, Hernandez, & Bates, 1998; Mackay & Flege, 2004; Mägiste, 1979; Roberts, Garcia, Desrochers, & Hernandez, 2002; for an overview see Hanulová et al., 2011). In addition, further studies have reported that L2 bilinguals suffered more word retrieval failures than L1 bilinguals or monolinguals in tip-of-the-tongue elicitation tasks (a tip-of-the-tongue state is the feeling of knowing an infrequent object's name, but being unable to retrieve it immediately; Gollan & Acenas, 2004; Gollan & Brown, 2006; Gollan, Montoya, & Bonanni, 2005). Additionally, in verbal fluency tasks in which participants are asked to generate as many exemplars of a given semantic category (e.g., animals), non-native speakers have been shown to have slower first response times and retrieve less category members than natives (e.g., Gollan, Montoya, & Werner, 2002; Portocarrero, Burright, & Donovick, 2007; Rosselli et al., 2000; Rosselli & Ardila, 2002; Sandoval, Gollan, Ferreira, & Salmon, 2010; but see Bialystok, Craik, & Luk, 2008a; Luo, Luk, & Bialystok, 2010). The naming cost associated with L2 speech has been observed in many different languages and across different age groups (e.g., Chen, Cheung, & Lau, 1997; Gollan et al., 2008).

Given that the amount of inter-language influences in bilinguals varies depending on factors like proficiency, age of acquisition, and the specific languages, one could assume the bilingual cost to vary accordingly. In particular, one could argue that the early and highly proficient L1 bilinguals tested in the current research should have supposedly adapted their language system from early on, and thus they should show very small or no effects in terms of a cost when compared to monolinguals. However, there is previous evidence that has shown that speech production is affected even in the case of such early highly proficient bilinguals. In a picture naming study by Ivanova & Costa (2008), they tested participants from the same bilingual community as the present one. The authors reported that onset latencies in single word production were slower for bilinguals using their L1 than in a monolingual control group. Their result imply that mastering two languages leads to a cost in word production, even in a first learnt and dominant language.

Altogether, the above studies document a bilingual cost over a variety of linguistic tasks. Moreover, due to the absence of any theoretical reasons to distinguish between the bilingual cost reported for L2 and L1 speech, one may suggests a common origin of this phenomenon. On this view, any performance differences between mono– and bilinguals in linguistic tasks should be due to the fact that there are inter-language interactions in the bilingual system that are absent in monolinguals. The questions that immediately arise at this point are: what mechanism is responsible for such costs, and at which level during speech processing do(es) the cost(s) occur?

Three major accounts have been put forward to explain linguistic performance differences between mono- and bilinguals. As introduced above, according to the executive control account, the bilingual cost is the consequence of applying language control mechanisms during speech production (Green, 1998). Since lexical representations of the two languages are co-activated during lexical access, bilinguals are exposed to potential competition between translations. To avoid interference between languages, bilinguals have to apply language control mechanisms. This additional processing would slow down lexical access and selection in bilinguals compared to monolinguals. Thus, this account relies on the fact that bilinguals have to constantly control two languages, and explains linguistic costs in terms of additional executive control processes when compared to monolinguals.

An alternative but not mutually exclusive explanation relates linguistic costs to general language usage and to the fact that bilinguals produce speech in each of their languages overall less often than monolinguals who always speak only one language. Basically, it assumes that the bilingual cost is a frequency effect in disguise (Gollan et al., 2008). That is, frequency-of-use of the bilinguals' lexical representations should be lower than those of monolinguals', and therefore weaker links would be created between the semantic and phonological representations in bilinguals. This account makes several clear predictions regarding lexical frequency effects. Given that word frequency negatively correlates with the speed of lexical retrieval (Oldfield & Wingfield, 1965), bilinguals would show a cost in lexical access relative to monolinguals. Due to the logarithmic shape of the effect of lexical frequency on naming times, decreasing language usage would lead to increasing effects of lexical frequency. Thus, effects of lexical frequency should be larger in L2 than in L1 bilinguals, which in turn should be larger than in monolinguals. Additionally, reduced language use would mainly affect words of lower frequency. In this conception, the bilingual cost in speech production would emerge from the same mechanism that accounts for lexical frequency effects in monolinguals. This is, it takes longer to produce those words that over time have been practiced less, affecting rarely used words more strongly than frequently used ones.

Finally, a recent proposal to explain the bilingual cost in language production is based mainly on empirical evidence on inter-language interactions in bilinguals. This account derives from a review of the literature on non-native and native language processing and relates performance differences to post-lexical processing such as phonological encoding, syllabification, and articulation (Hanulová et al., 2011; Indefrey, 2006). Accordingly, Indefrey and collaborators claim that phonological encoding in non-native speech may be more effortful if for example the phonotactic constraints on syllable structure of the native language are applied to the non-native language, resulting in costs during non-native speech. Alternatively, this account may also incorporate each of the two previously described explanations, but with an exclusive restriction to post-lexical processes. As a result, the mechanisms responsible for the more effortful non-native speech may be explained in terms of frequency (e.g., syllable frequency, motor-program frequency etc.) and/or the need to apply language control to avoid interference from the non-target phonemes.

Up to now, it is still unclear what mechanisms are responsible for the bilingual cost in language processing, and at what stage they emerge. Evidence that stages following concept selection are responsible comes from Gollan et al. (2005). In their study, mono- and bilingual speakers were asked to classify picture names into categories. Their results showed that both speaker groups performed equally well on this task, implying that the bilingual cost does not originate at the semantic level where concepts, at least in the case of concrete objects, are shared.

Given the available evidence on the bilingual cost and the various explanations to it, it is still a debated issue which of the stages after concept selection would be affected and how. Some authors proposed that the bilingual cost mainly originates during lexical access (e.g., Gollan et al., 2008; Green, 1998; Strijkers et al., 2010), while others suggest that they stem exclusively from late post-lexical levels during speech production (Hanulová et al., 2011; Indefrey, 2006). One major difficulty in disentangling between effects from different processing stages relates to the cascading nature of the language production system. Costs observed at late levels of processing could be a mere consequence of processing difficulties occurring at earlier levels, a point that was made earlier for articulatory durations. In this vein, costs at the lexical level might be carried over into late stages such as phonological encoding and articulation. Moreover, these costs could also be reinforced by additional processing demands during late stages, indicating an emergence at both levels. Alternatively, in keeping with an interactive model of speech production, effects at late stages may also impact on word selection and so influence word retrieval. In sum, there is no clear agreement about the origin and mechanism behind the bilingual cost. In a recent review by Runnqvist, Strijkers, Sadat, & Costa (2011), the authors examined behavioral and neuroscientific evidence on the bilingual cost and concluded that the cost most probably originates during lexical access, but affects subsequent stages as well. Further investigations will have to clarify the independent contributions from each of the processing stages involved in the bilingual cost. A first step in this direction is provided by the present dissertation that elaborates on the relationship between the bilingual cost and stages of phonological and articulatory processing. It is important to note that the approach of the current dissertation is very similar to the ideas proposed in the post-lexical account. Given evident inter-language influences at late stages of language processing in bilinguals, the current dissertation focuses on the phonological and articulatory processes to explain the bilingual cost. However, unlike the post-lexical account, the present dissertation does not specify that the bilingual cost emerges at these late levels. Rather, we will explore these late processes in bilinguals to provide further insights about inter-language influences in the speech of bilinguals and describe how they relate to performance differences relative to monolinguals.

Before turning to the experimental section of this dissertation, we will introduce further open questions about the scope of the bilingual cost in multiword production.

1.5 The bilingual cost beyond single word production

As in the monolingual literature, most studies on speech production in bilinguals have assessed linguistic performance during single word production. Consequently, most evidence on the bilingual cost stems from tasks requiring single word production (e.g., picture naming and verbal fluency tasks). It is certain that single word production entails many core processes of language production (see section 1.1). However, it has little bearing on everyday speech. For example, in order to produce a simple noun phrase in English (i.e., "the red car"), the speaker not only needs to retrieve the two lexical items corresponding to the object (car) and the property (red), but also the grammatical and syntactic rules of the language in use (correct determiner, order of the words etc.). Importantly, at present we still do not know whether and how these additional processes (e.g., grammatical and syntactic encoding) may modulate the bilingual cost in speech production.

As reviewed in the above section, previous studies observed poorer performance for bilinguals than monolinguals in single word production. It is of high importance to address the scope of such a cost and to relate it to contexts of more natural speech production.

One possibility is that the influence of the non-spoken language could be reduced when more than one word has to be selected. This idea is inspired by studies showing that context effects in language comprehension can reduce inter-language influences (e.g., Elston-Güttler, Paulmann, & Kotz, 2005; Schwartz & Kroll, 2006). Thus the bilingual cost could be reduced when a multiword context restricts production to the spoken language. If the bilingual cost is minimized in more complex utterances, the consequences of this cost for the regular use of language might be negligible.

However, given previous results in the monolingual literature where single word production has been contrasted with noun phrase production, one could tentatively predict that the bilingual cost will be larger in the latter context. For example, a study by Alario, Costa, & Caramazza (2002a) showed that noun and adjective frequency effects are both present in English noun phrase productions like “the blue kite”. According to the authors, the retrieval of all the elements of a noun phrase is already performed before the utterance is initiated (see Alario, Costa, & Caramazza, 2002b, and Levelt, 2002, for a discussion about the scope of speech planning). Thus, it can be suggested that onset latencies in noun phrase production would depend on how fast both the adjective and the noun are retrieved (see also Ayora & Alario, 2009; Costa & Caramazza, 2002; Janssen & Caramazza, 2011). In such a scenario, one would expect the bilingual cost observed in single word retrieval to be larger in noun phrase production, as bilinguals might be slower in retrieving the different words that compose the utterance.

Additional evidence that the bilingual cost might be increased in multiword production comes from a study on articulatory durations. Flege & Hojen (2004) compared the articulatory durations of speech in late L2 speakers to those of monolinguals. Their results showed that L2 bilinguals took more time to articulate an utterance than monolinguals. Importantly, they reported that with increasing complexity of the utterance to produce (from noun phrase to simple and complex sentences), the differences in articulatory durations between the two speaker groups increased as well. Thus, if the bilingual cost in language production is maximized in multiword utterances, then consequences of bilingualism for speech performance could be even more important than previously thought. Note that previous studies on articulatory durations focused on how using a late acquired L2 influences articulation compared to monolinguals or early bilinguals. Up to now, we lack knowledge on whether the bilingual cost is also present in early highly proficient bilingual speakers articulating speech in their L1, and if so, whether this cost may be generalized to multiword productions.

To conclude, the above review of the literature revealed that there are important inter-language influences at late stages of speech processing in bilinguals, but it is unclear how they relate to the bilingual cost. Thus, the major goal of this dissertation is to study phonological and articulatory processes underlying speech production in mono- and bilinguals. Given the influential role of phonological properties on speech production in mono- and bilinguals (i.e., effects of phonological neighborhoods and cognates), it needs to be established up to what extent and how these properties may affect the bilingual cost. Finally, it is unclear how the bilingual cost extends to articulatory processes of early highly proficient bilinguals using their L1 and how such a cost could be explained. In addition, information is missing on how the bilingual cost may generalize beyond single word production. These issues are important to obtain more detailed information on the origin and on the scope of the bilingual cost. Specifically, the studies presented in the experimental section will approach this goal by addressing the following two main questions:

1) How does phonological similarity affect speech production in mono- and bilinguals?

2) Do articulatory processes contribute to the bilingual cost in speech production?

2. EXPERIMENTAL SECTION: Studies of phonological and articulatory processes

As already mentioned in the General Introduction, the phonological properties of a word have been shown to impact on the word retrieval process in both mono- and bilinguals. Thus, one important aspect in explaining difficulties during speech performance is to consider the extent to which a given word bears more or less phonological similarity with other words within one language (for monolinguals) and across languages (for bilinguals). Hence variables like phonological neighborhood density and cognate status become central. In previous studies, increasing similarity for both of these variables has been shown to facilitate word retrieval. Strikingly, explanations for the facilitatory effect of cognates in bilinguals have been closely related to the facilitatory effect observed for phonological neighborhood density (e.g., Costa et al., 2005; Runnqvist, FitzPatrick, Strijkers, & Costa, 2012). Both facilitatory effects have been explained by increased overlap at the phonological level that feeds back activation to the lexical representation of the word. Given that the explanations for the effects of cognates and phonological neighborhood density both rely on the principle of interactive activation in speech production, it appears likely that a common dynamic underlies them.

However, recent studies on the effect of phonological neighborhood density in speech production have provided conflicting results: some studies reported that naming responses were delayed for words from dense neighborhoods (e.g., Arnold, Conture, & Ohde, 2005; Taler, Aaron, Steinmetz, & Pisoni, 2010; Vitevitch & Stamer, 2006, 2009), whereas others reported faster naming responses (e.g., Baus, Costa, & Carreiras, 2008; Perez, 2007). Specifically, studies reported contradictory results in the case of Spanish speech production (Baus et al., 2008; Perez, 2007; Vitevitch & Stamer, 2006, 2009), the language of testing in the current research. Before turning to the effects of phonological properties in the speech of bilinguals, we first explored how phonological similarity among words within one language impacts on speech production. We did so by assessing the performance of monolinguals in a large-scale picture naming study in Spanish. In this study, we attempted to reconcile the highly conflicting literature on the effect of phonological neighborhood density in speech production. We concluded that there are task-dependent differences in the way phonological properties affect word retrieval in monolinguals.

The second study of this dissertation concerns the influence of phonological similarity across translations (i.e., cognate status) on speech production in bilinguals. Studies with bilinguals have consistently shown better performance for phonologically similar translations (i.e., cognates) as compared to non-similar sounding translations (i.e., non-cognates; see section 1.3 of General Introduction). Thus, we investigated the influence of phonological properties on the bilingual cost, in order to understand how these two bilingual phenomena may relate to each other. We attempted to clarify the rather neglected role of phonological processing on the bilingual cost, and explored the origin of phonological similarity effects in bilingual speech.

The third study of this dissertation addressed the scope of the bilingual cost and aimed to provide new evidence in the domains of articulatory duration and noun phrase production. As mentioned in the General Introduction, several studies have already shown that articulatory durations of late L2 bilinguals were slower than those of native speakers. Here we focused on whether the bilingual cost would also be apparent in early highly proficient bilinguals using their L1. In addition, we also investigated how the bilingual cost may generalize from single word to noun phrase production.

2.2 Reconciling phonological neighborhood effects in speech production through single trial analysis

Sadat, J., Martin, C.D., Costa, A., & Alario, F.-X. (submitted). Reconciling phonological neighborhood effects in speech production through single trial analysis. *Journal of Memory and Language.*

Reconciling phonological neighborhood effects in speech production through single trial analysis

Jasmin Sadat [a, b], Clara D. Martin [a, c, d], Albert Costa [a, e], F.-Xavier Alario [b]

[a] Departament de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra, Barcelona, Spain

[b] Laboratoire de Psychologie Cognitive, CNRS and Université Aix-Marseille, France

[c] Basque Center on Cognition, Brain and Language, Donostia-San Sebastian, Spain

[d] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

[e] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Running head: Phonological neighborhood effects in speech production

Address correspondence to:

Jasmin Sadat

Departament de Tecnologies de la Informació i les Comunicacions

Universitat Pompeu Fabra

C/ Roc Boronat 138

08018 Barcelona, Spain

Phone: +34 - 93 542 13 81

Email: jasmin.sadat@upf.edu

Abstract

Phonological neighborhood density (PhND) refers to the number of words that differ from a given word by a single phoneme. Previous studies on the effect of PhND in speech production have reported conflicting results, notably within and across languages. The aim of the present study was to clarify the role of PhND in a large-scale picture naming experiment. Our results showed that increasing PhND has a detrimental effect on naming latencies. Furthermore, several re-analyses of independent data sets provided evidence for an inhibitory effect of dense phonological neighborhood on naming latencies. In addition, we highlight that the effect of PhND differs according to the task and performance measure at hand. We argue that when naming speed is tested, latencies are influenced by competitive processes involving similar sounding words, while in accuracy tasks phonological similarity plays a facilitative and protective role. We conclude that the lexical network underlying speech production should not be described solely on the basis of static representational properties such as phonological similarity. The dynamics of the retrieval process have an influential impact on how these properties surface in speech performance.

Keywords: Speech production; Phonological similarity; Neighborhood density; Lexical access; Mental lexicon

Acknowledgements

This research was supported by a grant from the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 Grant agreement n◦ 263575), two grants from the Spanish government (PSI2008-01191, Consolider Ingenio 2010 CSD2007-00012), and the Catalan government (Consolidado SGR 2009-1521). Jasmin Sadat was supported by a pre-doctoral fellowship from the Spanish Government (FPU-2008). Clara Martin was supported by the Spanish government (grant Juan de la Cierva). We thank Laure Minier and Elena Guerra for checking a portion of the vocal responses, and Luis Fuentes for providing access to the laboratory. We are very grateful to Matt Goldrick, Gary Dell, and Harald Baayen for helpful comments on earlier versions of this work.

Reconciling phonological neighborhood effects in speech production through single trial analysis

Native speakers of a language know a myriad of different words. This so-called mental lexicon is often described as an interconnected network in which representation distance may depend on meaning or form similarities among the words. One crucial step for understanding this network is to describe which kind of similarity influences language processing and how it modulates performance. The goal of the present research is to advance our understanding of how form similarity affects word retrieval during speech production.

Compared to the large amount of studies concerning form similarity in visual (e.g., Andrews, 1989, 1992; Carreiras, Perea, & Grainger, 1997; Grainger, O'Regan, Jacobs, & Seguí, 1989, 1992; Forster & Shen, 1996; Sears, Hino, & Lupker, 1995; Grainger, 1990; Grainger & Seguí, 1990; Perea & Pollatsek, 1998; Pollatsek, Perea, & Binder, 1999) and auditory word recognition (e.g., Dufour & Frauenfelder, 2010; Goldinger, Luce, & Pisoni, 1989; Landauer & Streeter, 1973; Luce, 1986; Luce & Pisoni, 1998; Vitevitch & Rodríguez, 2005; Ziegler & Muneaux, 2007; Ziegler, Muneaux, & Grainger, 2003; Vitevitch & Luce, 1998, 1999), research in the field of speech production is scarce and conflicting. Our goal here is to establish how form similarity influences access to the lexical network during speech production. Phonological neighborhood density (PhND) provides an approximation of how similar or interconnected a word is within the lexical network. It refers to the number of words that can be formed from a given word by substituting, adding or deleting one phoneme (Luce, 1986). For example, the word “bat” sounds similar to many other words (e.g., “cat”, “fat”, “rat”, “mat”, “bad”, “sat”, etc.; i.e., dense neighborhood), whereas the word “elk” shares its phonological form with only a few words (“ilk”; “elm”, “elf”, “else”; i.e., sparse neighborhood).

In what follows, we will first present previous research on the effect of phonological similarity in speech production. We will then describe how speech production models account for the effect of phonological similarity. Finally, we will illustrate the inconsistent pattern of results reported for PhND.

Given that speech production proceeds from meaning to articulation, it is not immediately evident that words related in sound should affect word retrieval. Yet there are studies indicating that form properties nevertheless influence word selection. One such piece of evidence is the mixed error effect (e.g., Dell & Reich, 1981; Harley, 1984; Martin, Gagnon, Schwartz, Dell, & Saffran, 1996). Mixed errors share both semantic and phonological similarity with the intended word (e.g., “rat” for “cat”). They occur more often than predicted based on rates of purely semantic (e.g. “cat” for “dog”) or phonological errors (e.g., “cab” for “cat”). Relatedly, Miceli, Capasso and Caramazza (1999) showed that semantic errors of aphasic patients were influenced by recently produced words in the same (oral) or different (written) modality.

This finding indicates that lexical selection may be constrained by previously activated form information (phonologic or orthographic; see also Alario, Schiller, Domoto-Reilly, & Caramazza, 2003). Further evidence for the influence of phonological properties on word selection comes from priming studies (Ferreira & Griffin, 2003; see also Rapp & Samuel, 2002, and Jaeger, Furth, & Hilliard, 2012). Ferreira and Griffin asked participants to name pictures following sentences that primed a semantic or homophone competitor (e.g., "nun" or "none" respectively for the picture of a "priest"). As expected, pictures were more often misnamed after presentation of a semantic priming sentence (i.e., saying "nun" instead of "priest") than after an unrelated sentence. Interestingly, the homophone priming sentences had the same effect as the semantic ones. Thus for the word "none" to be a successful intruder its phonological similarity had to combine with the semantics of its homophone "nun". Therefore the authors suggested that phonological similarity can promote errors at the lexical level.

To account for these observations of form influences on lexical selection, it is generally assumed that there are feedback links between the phonological and lexical representational levels in speech production models (i.e., interactivity; Dell, 1986; Harley, 1993; Rapp & Goldrick, 2000). According to interactive language production models, words with many neighbors receive more activation than words with few neighbors due to increased feedback from their shared phonological segments (Dell, 1986; Dell & Gordon, 2003). This means that there is bidirectional excitatory activation spread (forward and backward) between the phonological and lexical level. For example when the word “bat” has to be produced, its phonological representation co-activates phonologically similar lexical items (see examples above). These co-activated lexical items send back some activation to their phonological representations, thereby reinforcing the activation of the shared phonemes with the target word. Note however that such phonological influences in speech can also be explained by an alternative framework. In terms of strictly feed-forward models of lexical access, such effects stem from monitoring processes (Levelt, 1989; Levelt, Roelofs, & Meyer, 1999). Proponents of such an account assume a monitor that covertly filters out some utterances more often than others. In general, it is assumed that the monitor employs the comprehension system to compare a to-be produced word with its intended form. For example, in the case of mixed errors, both a target and a co-activated intruder are erroneously selected, resulting in a blend response. Given the additional form similarity of mixed errors, they will be overall more likely to slip by the monitor than a purely semantic error. Altogether, the fact that phonological activation influences word retrieval, whether it is through interactivity or monitoring mechanisms, motivates a detailed investigation of the role of PhND in word production.

A seminal study conducted by Vitevitch (2002) reported an investigation of the role of PhND in speech production. His results showed that naming latencies were faster for English picture names from dense neighborhoods than for those from sparse ones. However, since Vitevitch (2002), several studies have reported conflicting results in English, in which no effect of PhND was observed (e.g., Bernstein Ratner, Newman, & Strekas, 2009; Newman & Bernstein Ratner, 2007; Vitevitch, Armbruster, & Chu, 2004), or most remarkably, inhibitory effects were obtained on production measures like naming speed or accuracy (e.g., Arnold, Conture, & Ohde, 2005; Munson, Swenson, & Manthei, 2005; Newman & German, 2002, 2005; Taler, Aaron, Steinmetz, & Pisoni, 2010). Thus at present the effect of PhND on the most studied language (English) is still unclear.

Unfortunately, evidence from other languages is also controversial. Vitevitch and Stamer (2006, 2009) observed an inhibitory effect of PhND in Spanish picture naming: the denser the neighborhood, the slower the picture naming latencies. These contrasting cross-linguistic effects were originally attributed by Vitevitch and Stamer to differences between the richness of the morphological system of Spanish and that of English (see Ziegler & Perry, 1998, for a similar argument in word recognition). However, a more recent study by Baus, Costa, and Carreiras (2008) challenged the observations of Vitevitch and Stamer. Baus et al. reported a facilitative effect of PhND in Spanish (i.e. faster naming for words from dense than sparse neighborhoods). The authors concluded that the PhND effect is facilitative in Spanish as previously observed by Vitevitch (2002) in English, and they argued for a language-independent interactive spreading activation model of speech production. A similar facilitatory effect of PhND in Spanish has been reported by Pérez (2007), albeit the focus of his study was on frequency effects (see also Rodriguez-Gonzalez, 2012, for an apparent facilitatory effect in certain cases of verb production). Lastly, in a study conducted with Dutch native speakers, Bien, Baayen, and Levelt (2005) reported a non-linear effect of PhND, whereby naming responses were fastest for very sparse and very dense PhND of the initial constituents of compound words (although see Tabak, Schreuder, & Baayen, 2010, for no PhND effect at all in Dutch). In summary, the available evidence does not provide a consistent pattern for the effect of PhND on speech production performance across various languages. Furthermore, the explanations given to the discrepancies between studies (e.g., cross-linguistic properties, inherent property of interactive systems, etc.) do not seem to provide a satisfactory account to the inconsistent pattern of results that has been observed within and across languages.

The Present Study

In the present article, our goal is to clarify the effect of PhND in language production by reporting new data, re-analyzing data from previously reported studies and establishing a detailed summary of the available evidence. This should allow an evaluation of the theoretical accounts put forward for the effect of this variable.

Before doing so, an important caveat should be made explicit. PhND is an intrinsically between-item variable. Examining its effect requires that the performance across words is compared, and thus that the possible contribution from other variables that may affect performance is partialled out or controlled for. This is only possible if many of these variables (e.g., word length, syllable frequency, etc.) as well as PhND estimates themselves are available for the language under study. Computing reliable estimates for such variables is an endeavor of itself, and indeed large and broad psycholinguistic databases are only available for a handful of languages. This together with the availability of naming data are the main factors setting the scope of the current proposal.

Below we report a large-scale experiment on picture naming in Spanish in which we estimated the effect of PhND on naming latencies to test whether co-activation of similar sounding words facilitates or hinders retrieval of the desired word. We did so by measuring the naming latencies of word production in a standard picture naming paradigm. In addition, we will report re-analyses of several published data sets on Spanish, French and Dutch.

Our current study incorporates several improvements compared to previous studies on PhND in speech production. First, previous research has been limited to relatively small sets of items in factorial designs with PhND as a between-item manipulation (i.e., comparing performance for words from dense versus sparse neighborhoods; although see Baus et al., 2008, for a cross-linguistic control within "items"). The data set we collected and analyzed is an order of magnitude larger than that of the studies conducted previously (31,980 trials here vs. 1,482 trials in Baus et al., or 1,152 trials in Vitevitch & Stamer, 2006). The importance of this shift in the amount of evidence can hardly be overestimated. Second, the data was analyzed with regression models performed at the single trial level, providing a highly fine-grained approach in which the properties of each individual word and participant are considered explicitly. Third, we included several additional variables in our analyses to control for possible confounds for the effect of PhND. The additional variables were selected according to two groups of previous research. First, previous studies on speech production identified some of the most important predictors of naming performance (e.g., Alario, Ferrand, Laganaro, New, Frauenfelder, & Seguí, 2004; Barry, Morrison, & Ellis, 1997; Cycowicz, Friedman, Rothstein, & Snodgrass, 1997; Ellis & Morrison, 1998; Lachman, Shaffer, & Hennrikus, 1974; Severens, Van Lommel, Ratinckx, & Hartsuiker, 2005; Snodgrass & Yuditsky, 1996; Vitkovitch & Tyrrell, 1995). Based on these studies, we included several influential predictors such as name agreement or age of acquisition to investigate whether PhND would impact speech performance beyond the effects of these variables. Second, other variables such as neighborhood frequency (i.e., the average lexical frequency of all neighbors), onset density (i.e., number of neighbors that share the onset with the given word), syllable frequency or word length have been proposed as potential confounds for PhND (see Storkel, 2004; Vitevitch & Sommers, 2003; Vitevitch et al., 2004). For example, shorter words tend to have more neighbors than longer words, and words consisting of high frequency syllables tend to have more phonological neighbors than words consisting of low frequency syllables. Given the intermingled relationships between these confounding variables, the present study considered them all as potential regression predictors to ensure a reliable and independent effect of PhND. Finally, we took especial care to address issues of collinearity between the large numbers of variables included, for example by comparing results of alternative methods.

Method

Participants

Thirty native Spanish monolinguals participated in the experiment. They were all students from the University of Murcia in Spain. They grew up in Spanish speaking families and used only Spanish for daily communication. All participants had normal or corrected-to-normal vision, and received a monetary reward (20 Euros) for participating in the experiment.

Stimuli

533 black-and-white line drawings of common objects were selected from various picture databases (Bonin, Peereman, Malardier, Méot, & Chalard, 2003; Pérez & Navalón, 2003; Snodgrass & Vanderwart, 1980; Székely et al., 2004). Spanish picture names met the following criteria: (a) they consisted of a single word; (b) they were present in the Spanish database BuscaPalabras (Davis & Perea, 2005); (c) they had no other meanings with higher frequency usage to be confused with (e.g. “tienda”, meaning “tent” or “shop” in English, or “sobre”, meaning

"envelope" or "on"/"over"/"about"); and (d) they had relative high name agreement in Spanish (based on previously established name agreement values if available in Cuetos, Ellis, & Alvarez, 1999, and through offline pretests with four Spanish monolinguals from among the university staff). Pictures had black outlines and white surfaces and were presented 300 pixels wide x 300 pixels high on a white rectangle with a monitor resolution of 800 x 600 pixels.

For the multiple regression analysis, the following item-related predictors were collected from the Spanish lexical database BuscaPalabras (Davis & Perea, 2005) if not specified otherwise:

- Phonological neighborhood density (PhND; range [0, 37] number of neighbor words, *M* = 5, *SD* = 7), as defined in the second paragraph of the Introduction.

- Neighborhood frequency (range [0, 3.03] log occurrences per million, *M* = 0.70, *SD* = 0.69). It refers to the average lexical frequency of all neighbors. A logarithmic transformation was applied to avoid the undue influence of extreme values in the regression.

- Onset density (range [0, 35] number of onset neighbor words, *M* = 4, *SD* = 5). It refers to the number of neighbors sharing the first phoneme.

- Word length measured in phonemes (range [2, 11] number of phonemes, *M* = 6, *SD* = 2) and syllables (range [1, 5] number of syllables, *M* = 3, *SD* = 1).

- First syllable frequency (range [0.08, 4.61] log occurrences per million, *M* = 3.33, *SD* = 0.91). Values were taken from Alario, Goslin, Michel, and Laganaro (2010). A logarithmic transformation was applied to avoid the undue influence of extreme values in the regression.

- Word form frequency with values for written (range [0.07, 2.80] log occurrences per million, *M* = 0.98, *SD* = 0.54) and spoken frequency (range [0.00, 3.12] log occurrences per million, *M* = 1.02, *SD* = 0.60). A logarithmic transformation was applied to avoid the undue influence of extreme values in the regression.

- Objective visual complexity (range [4,048, 48,874] bytes, *M* = 14715, *SD* = 7557). Values were calculated based on the compressed JPEG file size (Székely & Bates, 2000).

- Name agreement (range [10, 100] percentage, *M* = 85, *SD* = 18), number of alternative names (range [0, 8] number of alternative words, *M* = 1, *SD* = 1), and H statistics (range [0, 2.6] unit of response agreement, *M* = 0.4, *SD* = 0.5; Snodgrass & Vanderwart, 1980). Values were established on the basis of the analysis of the data we collected (see below for response scoring).

- Subjective estimates of age of acquisition from adult ratings (range [2.3, 10.3] years; *M* = 4.6, *SD* = 1.4). Questionnaires consisted of four randomized lists of all picture names together with a 1 to 7 point rating scale (Barbón Gutiérrez &

Cuetos Vega, 2006), where 1 corresponded to an age of acquisition before 2 years old and 7 to an age of acquisition after 12 years old. They were filled in by 50 graduate Spanish monolinguals from different universities.

Design and Procedure

Six experimental lists were created with each of the 533 pictures being presented twice. The order of presentation of the pictures was pseudo-randomized with the following restrictions: (a) the 1066 pictures appeared in two successive sets of 533, with each picture presented only once per set; (b) picture names in two successive trials were neither semantically nor phonologically related. Participants were randomly assigned to one of the six lists.

Participants were tested in a sound-proof room. Stimulus presentation and the software voice-key were controlled via DMDX (Forster & Forster, 2003). The sensitivity of the voice-key was adjusted for each participant. Each trial started with a fixation cross displayed at the centre of the computer screen for 500 ms. After a 300 ms blank screen, the picture of the object to name was displayed. The picture remained on the screen until either the voice key detected the response or a 2500 ms deadline was reached without any overt response detected. The next trial began 700 ms after the recording period finished.

The experiment consisted of a short training followed by two sessions that were separated by a large break. First, participants were asked to name eight practice black-and-white pictures similar to the materials used in the experiment. They were instructed to name the pictures as fast and as accurately as possible using single nouns. After that, in the first session, they had to name the first set of 533 object pictures divided into eight short blocks. The responses were monitored by the experimenter and if participants gave another name for the picture than the intended one, they were corrected by the experimenter at the end of the first session. In the second session, the same 533 pictures were presented in the same way as in the first session, but in a different order. Participants' responses were automatically recorded by the computer as digitized sound files, and errors were noted online by the experimenter. Each session lasted about 45 minutes. In total, the experiment lasted about two hours.

Response Scoring

All 31,980 vocal responses and onset markers (533 pictures x 2 presentations x 30 participants) were visually checked offline with the software Check-Vocal (Protopapas, 2007) and corrected if necessary. Responses other than the intended one were classified as errors and excluded from naming latency analyses. To establish name agreement values from the present responses and to properly characterize the naming behavior of the participants, responses were coded into seven categories:

1. The produced name was the target name.
2. The response was a morphological variant of the target name, defined as a variation that shares the word root without changing the word's core meaning. Examples are clippings (e.g., "tele" for "television"- television) or plural/singular alternations (e.g., "ojos" for "ojo"- eyes / eye).
3. The response was a synonym for the target name, not sharing its word root (e.g., "frigorifico" for "nevera" - fridge).
4. The response included hyponyms (e.g. "fruta"-fruit for "naranja"- orange), semantic coordinates that share the same class but do not have the target word's core meaning (e.g., "cebolla"- onion for "ajo"- garlic), part-whole relations at the visual-semantic level (e.g., "mano"- hand for "brazo"- arm), and visual errors (e.g., "cacahuete"- peanut for "patata"- potato).
5. The response was the target name, but included a phonological error.
6. The response was the target name, but started with a hesitation (e.g., "aeh") followed by participant's auto-correction.
7. The response was omitted or was a non-word.

Data analyses

Due to the large number of variables known to affect naming performance and their potentially high level of collinearity, two statistical procedures were applied to select the most important variables and to reduce collinearity. We first ran a random forest (e.g., Breiman, 2001) using the package party in R and the function cforest (Hothorn, Buehlman, Dudoit, Molinaro, & Van Der Laan, 2006; Strobl, Boulesteix, Kneib, Augustin, Zeileis, 2008; Strobl, Boulesteix, Zeileis, & Hothorn, 2007). A random forest is a collection of classification trees providing a single measure of importance for each predictor. The following variables were included in the random forest analysis: name agreement, number of alternative names, H index, age of acquisition, concreteness, familiarity, written and spoken frequency, imageability, visual complexity, number of phonemes and syllables, PhND, neighborhood frequency, onset density, initial syllable frequency, and naming latencies. Secondly, correlations among variables were assessed through a hierarchical clustering analysis (using the Hmisc package in R and the varclus function; Harrell & with contributions from many other users, 2010). When several variables appeared to be highly correlated given the cluster analysis, we selected the one variable in the cluster having the higher measure of variable importance in the random forest analysis to represent the cluster.

Finally, before entering the selected predictors in the linear mixed-effects models, we systematically tested for correlations between them. Predictors that were correlated above 0.25 were orthogonalized by running a linear model in which one variable was used to predict the other variable. The residuals of these linear models were entered as fixed effects in the linear mixed-effects model. Since our focus of analyses was to explore the influence of PhND, we tried to remove from PhND any source of variance that was shared with other correlated variables. Therefore we run a regression model in which PhND was predicted by phoneme length and neighborhood frequency and the residuals of this model were used as PhND variable (i.e., residualized PhND). Furthermore, we used the residuals of a model in which phoneme length was predicted by neighborhood frequency (i.e., residualized phoneme length) and the residuals of a model in which onset density was predicted by PhND (i.e., residualized onset density; see Discussion section for performing the reverse regression). Given the subjective nature of age-of-acquisition ratings, lexical frequency and name agreement were regressed out of age of acquisition (i.e., residualized age of acquisition). This way, variables could be included in the model without introducing unreasonable amounts of collinearity.[1] We report variance inflation factors (VIF) of each of the predictors in the models to ensure reliability of the analyses (using the function vif in the package HH in R, Heiberger, 2009).

---

[1] As alternative prevention against collinearity, we also conducted an analysis in which the predictors were normalized (i.e., centered and scaled), a procedure which reduces collinearity. This analysis yielded the same results as those obtained from non-normalized orthogonalized predictors. Since the interpretation of the results is more difficult with normalized predictors, these results are not reported in the main text (see Appendix A for details).

VIF indicates how much the variance of an estimated coefficient is increased due to collinearity in the regression model.[2]

The previously described statistical procedures lead us to select eleven variables to be included in the linear mixed-effects models. The regression models always included fixed effects for residualized phoneme length, neighborhood frequency, residualized onset density and first syllable frequency, in order to control for possible confounds with PhND. Predictors motivated by previous naming studies on monolingual naming performance were also entered as fixed effects (name agreement, residualized age of acquisition, written lexical frequency, visual complexity). Variables accounting for general fatigue effects in the naming latencies were included by coding for trial order presentation (from 1 to 533) and session (first or second). All these predictors were entered in the model before PhND. This procedure ensured that any effect of the theoretically central predictor of PhND was significant over and above the variation explained by other secondary predictors. Several models were fitted and compared progressively by means of a log likelihood test to identify the optimal linear mixed-effects model.

One important aspect of the applied analyses is that naming latencies and accuracy rates were analyzed by mixed regression models at the single trial level, and not on averages (Baayen, Davidson, & Bates, 2008). In addition to fixed predictors considered in simple linear regressions, linear mixed-effects models

[2] VIF values around 2 are considered as problematic for a regression coefficient.

account for random variation induced by specific words or speakers. We introduced by-participant random intercepts, by-participant random slopes for all significant fixed effects estimated within participants (except for the control variable trial order), by-item random intercepts, and finally by-item random slopes for session.[3] All statistical analyses were run with the statistical software R (R Development Core Team, 2011) and linear mixed-effects models were computed with the package lme4 in R (Bates, Maechler, & Bolker, 2011). P-values were validated by Markov chain Monte Carlo simulations using the function pvals.fnc in the package “languageR” (Baayen, 2008) whenever possible (i.e., in the case of models without random slopes). The Box-Cox test (using the function boxcox in the package MASS in R, Venables & Ripley, 2002) indicated that the reciprocal transformation of the latencies was the most appropriate transformation for the data to reduce skewness and approximate a normal distribution. We used -1000/RT as an order (and to some extent magnitude) preserving transformation to facilitate the interpretation of our results. We followed Baayen's (2008) procedure of model criticism in which trials whose standardized residual value is above 2.5 are removed and the model is recomputed. Only such recomputed models are reported.

[3] A log-likelihood test indicated significant improvement by including such maximal random effect structure over the simple random intercept effect model. Trial order could not be included as a random effect, because the model did not converge. Although the contribution of some of the fixed effects (e.g., session) was reduced by including random slopes, it never resulted in the disappearance of a main effect.

Results

Naming latencies

After removing non-target responses (i.e., only category 1 responses were used) and outliers, 27 581 responses remained for analyses. The average naming latency was 910 ms (*SD* = 117). Results showed that naming latencies increased with the phonological density of a word's neighborhood (i.e., inhibitory PhND effect). Figure 4 presents an estimate of the variation of this effect across individuals. The effect of neighborhood frequency was marginally significant: words with more frequent neighbors tended to be named slower than words with less frequent neighbors.

As expected, name agreement, age of acquisition, and lexical frequency contributed significantly towards predicting naming speed: earlier learned words were named faster than later learned ones, words with high percentages of name agreement were named faster than words with lower percentages, and high frequent words were named faster than low frequent ones. The effects of trial order and session were significant, showing that responses to pictures became slower with increasing trial order and that responses in the second session were faster than in the first. The effects of visual complexity, phoneme length, first syllable frequency, and onset density were not significant. Predictors not showing significant effects were removed from the model. Statistical values for the final linear mixed-effect model together with the variance inflation factors of the predictors are reported in Table 1.

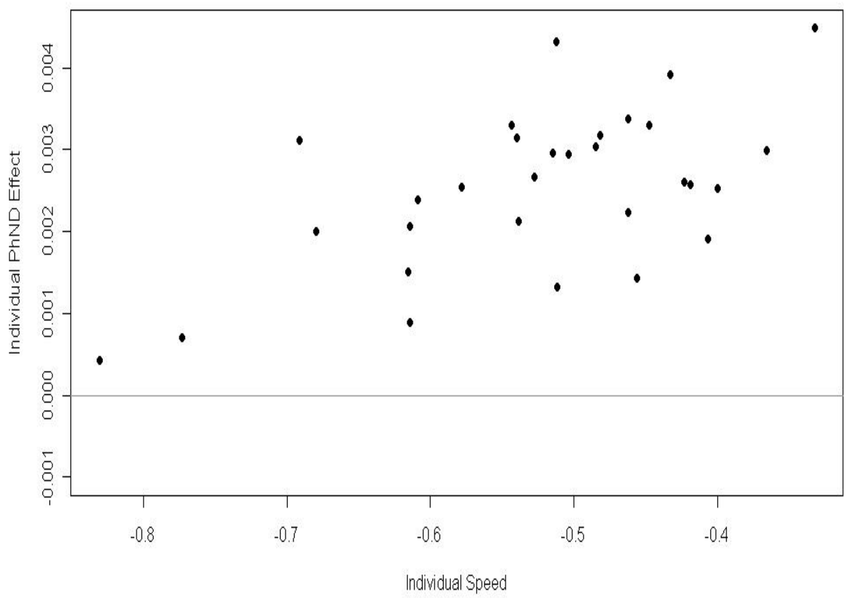

Figure 4: Relationship between magnitude of phonological neighborhood density (PhND) effect and overall speed estimated at the individual level. Each point represents a participant. Larger values in the x-axis indicate slower responding participants. Larger values in the y-axis indicate positive coefficients for the PhND effect (i.e. stronger inhibition). Note that all participants showed an inhibitory effect.

Table 1: Beta coefficients, effect magnitude in ms, standard errors (SE), t- and p-values together with the variance inflation factor (VIF) for each of the predictors in the final model.

Predictor	Effect magnitude	raw β	SE β	t-value	p-value	VIF
Intercept		-5.26x10-1	$3.35x10^{-2}$	-15.70	< 0.001	
Session	40.93	-1.26x10-1	$9.38x10^{-3}$	-13.40	< 0.001	1.0
Trial order	0.07	$9.20x10^{-5}$	$9.14x10^{-6}$	10.06	< 0.001	1.0
Name agreement	9.54	$-6.97x10^{-3}$	$3.54x10^{-4}$	-19.68	< 0.001	1.1
Age of acquisition	19.93	$2.74x10^{-2}$	$5.26x10^{-3}$	5.21	< 0.001	1.1
Lexical Frequency	29.74	$-4.26x10^{-2}$	$1.21x10^{-2}$	-3.51	0.001	1.1
Neighborhood Frequency	10.37	$1.41x10^{-2}$	$8.12x10^{-3}$	1.73	0.045	1.1
PhND	1.88	$2.53x10^{-3}$	$1.01x10^{-3}$	2.50	0.008	1.0

Note: PhND = phonological neighborhood density; Effect magnitude shows the increase in ms per scaling unit for each of the predictors.

Since ratings for concreteness, familiarity and imageability were available only for a subset of the 533 words (concreteness and familiarity: 385 words; imageability: 377 words), we conducted an additional analysis on this restricted data set (20 492 data points). The same random effect structure as in the previous model was included. When entering familiarity, imageability, and concreteness into the model, along with the other predictors, the effect of PhND remained significant ($\beta = 2.34x10^{-3}$, SE = $1.11x10^{-3}$, t(20091) = 2.10, p = 0.021; VIF = 1.0). The effect of concreteness was marginally significant (β = -1.68×10^{-2}, SE = 1.03×10^{-2}, t(20091) = -1.63, p = 0.056; VIF=1.3). The effects of familiarity and imageability were not significant.

Analysis of accuracy rates

There were 12% of non-target responses overall (3,934 out of 31,980). Most of them (5%) were of category 4 (i.e., semantic variations around the target, see Methods). Categories 5, 6, and 7 (responses containing speech errors; 1,681 trials) were contrasted from error-free responses of category 1 (28,046 trials) to accurately predict the probability of an error-free response. The generalized maximal linear mixed-effects model, computed over 28,930 trials, estimated how the predictors modulate the odds of error.[4] These were smaller for words that have higher name agreement (β = -0.13, SE = 0.01, z[28908] = -20.52, p = < 0.001; VIF = 1.2), higher lexical frequency (β = -1.4, SE = 0.36, z[28908] = -4.27, p = < 0.001; VIF = 1.2), and are learned earlier (β = 0.52, SE = 0.10, z[28908] = 5.06, p = < 0.001; VIF = 1.1). The effect of session was significant (β = -2.43, SE = 0.11, z[28908] = -21.81, p = < 0.001; VIF = 1.0). There was no significant effect of word length, first syllable frequency, neighborhood density, onset density, or PhND (β = 0.03, SE = 0.03, z[28908] = 1.06, p = 0.290; VIF = 1.1).

[4] The control variable trial order was not included because the model would not converge.

Discussion

The results of this large scale experiment show that increasing PhND has a detrimental effect on naming latencies, i.e. an inhibitory effect. This is the case even when many other important variables affecting the speed of speech production are taken into account. Incidentally, these findings confirm previous studies on predictors of naming performance, showing that name agreement, age of acquisition, and lexical frequency were the three most influential variables for predicting naming behavior (e.g. Alario et al., 2004; Barry et al., 1997; Bonin, Chalard, Meot, & Fayol, 2002; Ellis & Morrison, 1998; Gilhooly & Gilhooly, 1979; Kremin, Hamerel, Dordain, De Wilde, & Perrier, 2000; Lachman, 1973; Lachman et al., 1974; Pind & Tryggvadottir, 2002; Snodgrass & Yuditsky, 1996; Severens et al, 2005; Vitkovitch & Tyrrell, 1995).

Our main finding is in line with Vitevitch & Stamer (2006, 2009), and supports the view of increased competition during the production of dense PhND words in Spanish. Our results are at odds with those of Baus et al. (2008) and Pérez (2007) who reported a facilitative effect of PhND in Spanish naming performance - the same language of response we used in our experiment. Given the discrepancy of the present results with these two findings, we conducted several re-analyses of our data, the data of Baus et al. and that of Pérez. Moreover, we conducted analyses of the effect of PhND on three datasets from previously published papers: a naming study in Spanish (Sadat, Martin, Alario, & Costa, 2012), a large-scale naming study in French (Alario et al., 2004), and an additional large-scale naming study in Dutch (Severens et al., 2005).

Further Evidence from Chronometric Studies

Our set of 533 object pictures comprised most of the pictures used in the studies by Baus et al. (2008). Thus, we selected from our dataset the trials that corresponded to those pictures and applied (1) the same analyses as in the original study (Student t-test). We also applied the analyses used in the current study (2) to this subset of trials from our dataset and (3) to the original data from Baus et al. The Box-Cox test indicated a logarithmic transformation of the data of Baus et al. as the most appropriate to approximate a normal distribution. In the following re-analyses, linear mixed model analyses were conducted with PhND and all relevant control variables (word length, first syllable frequency, name agreement, age of acquisition, etc.) and in case of multicollinearity the same residualization procedures were used as in the main analyses.[5]

Four stimuli of the set used by Baus et al. (2008) were not part of our stimuli (two of each of their PhND categories). Applying the same statistical analysis (Student t-test) used by Baus et al. to

[5] Note that random by-participant and by-item slopes were introduced only when log-likelihood tests indicated significant improvement in the model's fit. Including these individual slopes never resulted in the disappearance of a main effect.

our data led to the same pattern of results that they reported: words coming from dense neighborhoods were named faster than words coming from sparse neighborhoods ($t_1[29] = 2.8$, $p < 0.01$; magnitude of the effect in our subset: 32 ms and in Baus et al.: 33 ms). This means that for an almost identical set of stimuli as used in Baus et al. (and even with more observations: 1,920 trials in our subset versus 1,238 trials in Baus et al.), our data also showed a facilitative effect of PhND. However, when using linear mixed model analysis the previously observed facilitative effect of PhND disappeared in the original data of Baus et al. ($\beta = 1.41 \times 10^{-3}$, SE = 2.25×10^{-3}, $t(1199) = 0.63$, $p = 0.27$; VIF = 1.0) and showed a trend towards inhibition in our selected subset ($\beta = -6.61 \times 10^{-3}$, SE = 4.11×10^{-3}, $t(1881) = 1.61$, $p = 0.08$; VIF = 1.1). Note that in the re-analyzes of both sets, the effect of word length was significant, indicating increasing naming latencies with increasing word length. Thus, the disappearance of the facilitatory effect could be due to a better fit of individual variations and residual noise often present in such relatively small datasets as well as a better control of confounding variables by including them in the model explicitly.[6] On the basis of further examples (see below), we will argue that the effect of PhND is often very small and difficult to detect. More

[6] A reviewer suggested that the disappearance of the facilitatory effect might be due to the transformation of the data (i.e., normalizing the data distribution in regression but not Student t-tests). We conducted the same analysis as in Baus et al. (2008) on the log-transformed data and still obtained a significant facilitative effect. This means that the applied data transformation is not responsible for the disappearance of the effect when analyzing the data of Baus et al. by means of linear mixed models.

powerful designs with a large number of trials and more fine-grained analyses methods provide better tools to reveal such subtle effects. For now we note that the facilitatory effect reported by Baus et al. did not survive a fine-grained analysis.

Additional previous evidence for a facilitative effect of PhND in Spanish naming was reported by Pérez (2007). Although the focus of his study was on frequency effects, the author also reported a facilitative effect of PhND in Spanish speech production. We selected the subset of our data (N_{items}=146) that overlaps with the stimuli by Pérez (N_{items}= 161). Applying the same statistical analysis (linear regression) as used by Pérez to our data did not yield the same results as he reported: there was no effect of PhND on naming latencies (β = 1.5, SE = 1.2, t[133] = 1.2, p = 0.23; VIF = 1.1). This means that for a very similar set of stimuli and same analyses as used in Pérez (and even with more observations: 8,092 trials in our subset versus 5,937 trials in Pérez), our data did not replicate the facilitatory effect of PhND reported by Pérez. However, when introducing the predictor variables used by Perez together with control variables into a maximal linear mixed model, the re-analyses of our subset showed a marginally significant inhibitory PhND effect ($\beta = 2.70\text{x}10^{-3}$, SE $=1.5\text{x}10^{-3}$, t(7742) = 1.77, p = 0.04; VIF = 1.0; number of observations after outlier removal: 7,766 trials). In short, our subset of data similar to Perez' stimuli showed an inhibitory effect when analyzed with fine-grained measures on more observations.

To gain further insights, we considered a third set of Spanish picture naming data that had not been tested previously for an effect of PhND (Sadat et al., 2012). In addition to the above mentioned control predictors, we entered the variables from the original study into the maximal linear mixed effect model.[7] We used -1000/RT for transforming the latencies as indicated by The Box-Cox test. The results on 21,264 trials showed an inhibitory effect of PhND on naming latencies (effect magnitude: 2.1 ms per unit [range: 0-26]; $\beta = 6.01 \times 10^{-3}$, SE = 2.71×10^{-3}, $t(21252) = 2.22$, $p = 0.02$; VIF = 1.0).

Additionally, we explored speech production data from another Romance language besides Spanish using a large set of French naming data (Alario et al., 2004). We introduced the variables of the original study together with first syllable frequency (taken from Alario et al., 2010) and PhND (from the French database LEXIQUE; New, Pallier, Ferrand, & Matos, 2001) into a maximal linear mixed model. The Box-Cox test indicated a logarithmic transformation of the data of Alario et al. as most appropriate. Since word length and PhND were highly correlated in this dataset ($r = -0.78$), we chose to include the original PhND variable and to enter word length as residualized fixed effect. The model showed a significant inhibitory effect of PhND on naming latencies of 16,258 trials (effect magnitude: 1.2 ms per unit [range: 0 – 62]; $\beta = 1.41 \times 10^{-3}$, SE = 6.12×10^{-4}, $t(16248) = 2.30$, $p = 0.01$; VIF = 1.2). However, when the opposite residualization was

[7] We used the data from two speaker groups (mono- and bilingual) because there was no significant interaction of group with PhND.

performed, there was no effect of PhND (β = 7.74x10^{-4}, SE = 7.77x10^{-4}, t(16248) = 1.00, p = 0.25; VIF = 1.0), whereas the effect of word length was significant, indicating *decreasing* naming latencies with *increasing* word length (as reported in the original study). This means that the effect of PhND in this dataset can be explained by the variance common to both word length and PhND. However, the direction of the observed length effect is opposite to what would be expected under theoretical hypothesis that predict an effect of this variable (i.e., word length was facilitatory, which corresponds to an inhibitory PhND effect), we are inclined to attribute the effect to PhND.

Finally, we explored the effect of PhND in a non-Romance language, using a large-scale naming dataset in Dutch (Severens et al., 2005). We introduced the variables of the original study together with PhND and summed neighborhood frequency (from the Dutch database CELEX; Baayen, Piepenbrock, & Rijn, 1993) into a maximal linear mixed model. We used -1000/RT for transforming the latencies as indicated by The Box-Cox test. The results on 14,277 trials showed no effect of PhND at all on naming latencies (β =5.87x10^{-4}, SE=1.1x10^{-3}, t(14247) = 0.55, p = 0.29; VIF = 1.0). As suggested by the study by Bien, Baayen, and Levelt (2005) in Dutch, we also explored the data for non-linear effects of PhND. The results indicated a non-significant linear component and a marginally significant quadratic term with a positive coefficient (i.e., an U-shaped PhND effect; β =2.00x10^{-4}, SE=1.3x10^{-3}, t(14524) = 1.56, p = 0.07; VIF = 1.2).[8] This result hints to an inhibitory component of the PhND effect that is present even in a non-Romance language like Dutch.

In summary, we presented evidence from five studies on the influence of PhND on naming latencies. In the case of Baus et al. (2008), we replicated their findings with their analyses method, but showed that with fine-grained analyses, the previously reported facilitatory effect disappeared. Regarding the data of Pérez (2007), we could not replicate the facilitatory effect of PhND, but observed an inhibitory effect with more detailed analyses. Likewise, when exploring the data of Sadat et al. (2012) for PhND we observed an inhibitory effect. The reanalysis of naming data in French (Alario et al., 2004) provides additional evidence, if somewhat more ambiguous, to support the claim of an inhibitory effect of PhND on speech production. The results from Dutch did not show any clear effect, but indicated a trend towards inhibition related to a non-linear component of the PhND effect. Overall, our results together with the re-analyses reported above support the view proposed by Vitevitch & Stamer (2006, 2009) that PhND shows an inhibitory effect in morphologically rich languages like Spanish and French. The evidence previously reported in favor of a facilitative effect in Spanish was shown to be inconclusive.

[8] Note that none of the PhND effects reported in the previous analyses showed significant non-linear effects.

General Discussion

In the present study we used the recently introduced analysis at the single trial level (Baayen et al., 2008) to assess the impact of PhND on word production performance. The result of a picture naming paradigm with a very large set of materials revealed an inhibitory effect of PhND on naming latencies. Five previous studies were reanalyzed using these statistical methods, which are arguably more fine-grained than those used in the original reports. The new analyses challenged previous conflicting findings from two studies and provided new results from three published datasets. These results confirmed the presence of an inhibitory effect. Thus, we conclude that the effect of PhND on word production speed is inhibitory in the languages considered.

Are there Cross-linguistic Differences in the Direction of the PhND Effect?

The above conclusion is in line with that of other studies assessing the influence of PhND on naming latencies of Spanish word production (Vitevitch & Stamer, 2006, 2009). In English, however, Vitevitch (2002) showed a *facilitatory* effect of PhND. In addition, following that seminal article, several studies in the speech production literature have since then reported an inhibitory effect in this language (see Table 2 for an overview of studies of PhND on naming latencies). For example, when assessing children's speech production latencies in English, naming was slower for words from dense phonological neighborhoods compared to sparse ones (e.g., Arnold et al., 2005; Munson et al., 2005). The authors appealed to developmental change considerations to explain the discrepancy between their inhibitory effect in children and the facilitatory result reported in adults by Vitevitch. The idea was that the impact of PhND changes with vocabulary growth during development (but see Coady & Aslin, 2003, for re-analyses showing that children's vocabulary contains even more similar sounding words than previous studies indicated). Further studies reported an inhibitory effect of PhND in English speech production latencies (Luce & Pisoni, 1998; Taler et al., 2010; Vitevitch & Luce, 1998). In these studies, participants were asked to repeat as fast as possible a word or sentence presented acoustically (i.e., shadowing task). The shadowing task involves both speech perception and speech production, and hence may reflect contributions of either process (Bates & Liu, 1996). For example, shadowing experiments are used to investigate the relationship between input and output phonological representations (e.g., Mitterer & Ernestus, 2008; Taler et al., 2010). Therefore, attributing the results from shadowing studies to perception versus production processes should be done cautiously.

To our knowledge, there is no study besides Vitevitch (2002) that has reported a facilitative effect of PhND on English naming latencies. Further studies in this language would be useful to test the robustness of these findings. In view of the relatively small data set sizes of previous studies, we highlight that large-scale studies will certainly aid in clarifying such effects. We speculate, on

Table 2: Overview of speech production studies of the effects of phonological neighborhood density on naming latencies

Study	Manipulation	Task	Population	Nr. of Trials	Language	Direction of Effect	Remarks
Vitevitch (2002) - Exp. 5	Factorial	Button-press followed by naming	young adults	1,200	English	Facilitation	
Vitevitch (2002) - Exp. 4	Factorial	Picture naming	young adults	1,200	English	Facilitation	
Vitevitch (2002) - Exp. 3	Factorial	Picture naming	young adults	1,632	English	Facilitation	
Baus et al. (2008) - Exp. 2	Factorial	Picture naming	young adults	1,482	Spanish	Facilitation	but see Reanalysis in current study
Perez (2007)	Continous	Picture naming	young adults	6,440	Spanish	Facilitation	but see Reanalysis in current study
Tabak et al. (2010)	Continous	Picture naming	young adults	2,890	Dutch	ns	
Reanalysis of Severens et al. (2005)	Continous	Picture naming	young adults	16,575	Dutch	ns	
Jescheniak & Levelt (1994)	Continous	Picture naming	young adults	?	Dutch	ns	
Bernstein Ratner et al. (2009)	Factorial	Picture naming	children (control group)	660	English	ns	
Newman & Bernstein Ratner (2007)	Factorial	Picture naming	adults (control group)	1,100	English	ns	
Vitevitch et al. (2004) - Exp. 3	Factorial	Picture naming	young adults	1,100	English	ns	
Munson et al. (2005)	Factorial	Word repetition	younger children	512	English	ns	
Marian et al. (2008) - Exp. 1	Factorial	Picture naming	young adult bilinguals in first and dominant language	855	German	ns	
Vitevitch & Stamer (2009)	Factorial	Picture naming	young adults	22,400	Spanish	ns	Reanalysis of Bates et al. (2003); trend for inhibition
Frank et al. (2007)	Factorial	Picture naming	young adults	3,840	Artificial	Inhibition	
Arnold et al. (2005)	Factorial	Picture naming	children (control group)	180	English	Inhibition	
Gordon (2011)	Continous	Picture naming	younger and older adults	14,600	English	Inhibition	ns for youngest adults
Munson et al. (2005)	Factorial	Word repetition	older children	480	English	Inhibition	
Vitevitch & Luce (1998)	Factorial	Shadowing	young adults	4,500	English	Inhibition	
Taler et al. (2010)	Factorial	Sentence repetition	younger and older adults	4,440	English	Inhibition	
Luce & Pisoni (1998) - Exp. 3	Factorial	Shadowing	young adults	7,200	English	Inhibition	
Reanalysis of Alario et al. (2004)	Continous	Picture naming	young adults	20,720	French	Inhibition	
Vitevitch & Stamer (2006)	Factorial	Picture naming	young adults	1,152	Spanish	Inhibition	
Reanalysis of Sadat et al. (2011)	Continous	Picture naming	young adults	11,200	Spanish	Inhibition	
Current Study	Continous	Picture naming	young adults	31,980	Spanish	Inhibition	

Note: Exp.= experiment; ns = not significant; nr = not reported the basis of admittedly scarce evidence, that in an English large-scale naming study the effect of PhND should turn out to be inhibitory. In fact, an inhibitory effect of PhND has very recently been observed in older adults naming pictures in English (Gordon, 2011).

The above cross-linguistic discussion notwithstanding, PhND *per se* may not be detailed enough to characterize the phenomena under scrutiny. Vitevitch and Stamer (2006) suggested that the difference between language production in Spanish and in English lies in the location within the word where neighbors can be formed in either language (e.g., frequently the rhyme in English, but the onset in Spanish: rhyme vs. onset density). A study by Vitevitch et al. (2004) showed that when PhND is controlled, English words with a high number of onset neighbors are produced more slowly than words with few onset neighbors. This suggests that phonological neighborhood effects in English may have several facets: overall neighborhood has a facilitatory contribution (Vitevitch, 2002) whereas onset neighborhood has an inhibitory one (Vitevitch et al., 2004). It is possible that the inhibitory effect of PhND in Spanish can be described more appropriately as an initial overlap effect, given the highly inflectional nature of Spanish.

In fact, our analyses do not reveal an independent effect of onset density and PhND in the datasets we report. Note that the relative contributions of PhND and onset density are difficult to tease apart, because the latter count is a subset of the former, and the two are highly correlated in Spanish and in our materials (r =

0.97; 82 % of all neighbors were onset neighbors). In all the models we reported, onset density was residualized against PhND. This means that the variance common to both variables was attributed to PhND. We also ran a series of models in which the residualization was performed the opposite way (i.e., PhND residualized on onset density). In these models, residualized onset density was significant, whereas PhND was no longer significant.[9] This indicates that the reported *inhibitory* effect can be attributed to the variance common to PhND and onset density. The results of the random forest analysis performed with unresidualized variables showed that onset density was ranked as being of slightly higher importance towards predicting naming latencies than PhND. To properly address the influence of onset density, one possibility would be to assess this measure relative to an entire language corpus (i.e., onset cohort as used in spoken word recognition; Marslen-Wilson & Welsh, 1978), and not only relative to the set of phonological neighbors differing by one phoneme. Following Vitevitch and Stamer's (2006) theoretical proposal, establishing the relative and concurrent contributions of such fine-grained positional effects to the PhND effect seems to require further empirical scrutiny. One recent study addressing this issue is Bien, Levelt, & Baayen (2011). Their results on naming latencies showed independent effects for onset cohort and rhyme neighbors, these being facilitatory and inhibitory respectively. However, note that the paradigm they used (i.e., a position-response association task) differs largely to standard

[9] In particular, PhND was not facilitatory in this analysis, as could have been suggested by the dual-faceted phenomenon present in English (see above).

picture naming as it involves recently learned arbitrary stimulus-response associations and a large number of item repetitions.

Another property that could modulate the effect of PhND is semantic and morphological similarity. Due to the inflectional properties of Spanish, semantically similar words often bear high phonological resemblance with differences in word affixes being smaller (e.g., “mejor” [better] -“mejorar” [improve]; see Arbesman, Strogatz, & Vitevitch, 2010, for a comparison of Spanish and English morphological networks in relation to PhND). Phonologically similar words are thus likely to receive additional activation through semantic similarity. Depending on further assumptions, this may increase competition at stages of lexical selection. Note that this explanation would also allow purely cascaded models of speech production (i.e. without feedback from phonology to the lexical level) to explain an inhibitory effect of PhND on speech production. Thus, one general caveat regarding an independent effect of phonological similarity is to consider the morphological structure of the words. To better characterize how the PhND effect relates to semantic similarity, future studies will need to take into account measures like morphological family size (again see Bien, et al., 2011, for some steps in this direction).

Based on the present results and the review of previous studies suggesting an inhibitory effect of PhND on onset latencies in naming tasks, one may wonder why the common assumption in language production research is that increasing PhND speeds up word production.[10] When reviewing the literature on PhND in speech production, it appears that many previous studies assessed accuracy measures like word finding rates or speech errors, not naming latencies. Several studies have reported that words with dense neighborhoods tend to be less amenable to tip-of-the-tongue (ToT) states than words with sparse neighborhoods (e.g., Brennen, Baguley, Bright, & Bruce, 1990; Brown & McNeill, 1966; Burke, MacKay, Worthley, & Wade, 1991; Dell & Gordon, 2003; Harley & Bown, 1998; James & Burke, 2000; Vitevitch & Sommers, 2003). Along the same lines, Meyer and Bock showed that phonologically related cues helped resolution of ToT states as compared to unrelated ones. In turn, semantically related cues did not show any superior effect on resolution efficiency of a ToT. Original studies investigating properties of ToT states have shown that lexical access is disrupted mainly at the stage of phonological retrieval (Badecker, Miozzo, & Zanuttini, 1995). This means that when representations have been successfully selected during previous stages (in some models, the lemma level), phonological cues will jolt the missing word’s segments for production.

[10] See for example the assumption made by Chen and Mirman (2012) with respect to the modelisation of PhND effects in speech production.

Activation from/to phonological similar words will reverberate within the phonological level and converge onto the shared phonemes of the missing target word, thereby facilitating its access. Interactive activation theories propose that TOT states occur when the connections to a word's phonological segments are weakened, resulting in insufficient activation to enable word retrieval. In this sense, words from dense neighborhoods will be easier to access during ToT states, because they are more interconnected at the phonological level and thus compensate better for deficient information at that level. In short, although words from dense neighborhoods are retrieved more slowly because of increased competition, in ToT states their phonology is accessed more easily because of the previous retrieval of the target's lexical representation.

In addition, several studies have reported that speech errors tend to be less likely for words from dense than sparse neighborhoods. This finding holds for healthy participants or individuals with speech production deficits in spontaneous speech or error-elicitation tasks (e.g., Best, 1995; German & Newman, 2004; Goldrick, Folk, & Rapp, 2010; Gordon, 2002; Kittredge, Dell, Verkuilen, & Schwartz, 2008; Middleton & Schwartz, 2010; Mirman, Kittredge, & Dell, 2010; Stemberger, 2004; Vitevitch, 2002; see Vitevitch, 1997, for an interaction between lexical frequency and PhND in malapropisms). The contrast between speech error and response time data is discussed more in detail in the next section.

Altogether, the literature shows that PhND effects in ToT and error elicitation tasks are facilitatory and not inhibitory, as we found for chronometric data in the present study. In other words, speed and accuracy of lexical access are affected differently by phonological similarity. Therefore, one main criterion to distinguish the different effects of PhND relates to a combination of strategies (simple naming vs. active searching) and measures (chronometric vs. accuracy data) required in the speech production task at hand. Given the evidence presented in the current studies, we conclude that the direction of the PhND effect orderly depends on the language production task you are engaged in (see Andrews, 1997, for a similar claim in visual word recognition).

Theoretical Implications of the Inhibitory Latency Effect of PhND

As already mentioned in the Introduction, PhND effects are often explained by interactive models of speech production (Dell, 1986). Such effects are assumed to result from feedback between shared phonological representations and the intended lexical representation, and are therefore predicted to be facilitative. This means that for a word with many phonological neighbors, many shared phonological representations send feedback that contributes to activating the word to be produced. However, while the phonological segments of the desired word send some activation to all lexical representations with similar phonology, in some cases these co-activated items can also trigger lexical competition. This requires close attention to the dynamics and the parameters of the model. Dell and Gordon (2003) addressed this issue with computational models simulating lexical retrieval. They showed that their model could explain facilitatory effects in accuracy rates. However, they also showed that with certain model parameters (e.g., in a weight-lesioned model), the activation of phonological neighbors could hamper lexical selection because these neighbors reached an activation level sufficiently close to the activation level of the target word, hence yielding an inhibitory PhND effect on accuracy.

These parameter settings may presumably also model the inhibitory effect in naming latencies. Given the high accuracy rates of our healthy young participants, however, we are reluctant to claim that lexical competition would arise only in the case of lesioned lexical connections. Instead, we suggest that the inhibitory effect of PhND observed in the present study reflects a general competition mechanism that is at work during lexical access (among semantically and phonologically similar words), and that surfaces in onset latencies. Although this mechanism could be reinforced by communication context (e.g. Jaeger et al., 2012) and developmental changes (e.g., Gordon, 2011), it may well constitute an underlying principle of the core process of lexical selection in non-impaired individuals.

Alternatively, our findings also fit with Dell's (1988) connectionist model of phonological encoding, where competition occurs at the phonological level. This account assumes that there is an initial parallel activation of an intended word's phonemes followed by a sequential left-to-right selection process. Initial parallel activation is increased by words sharing common phonemes. The sequential "left-to-right" manner of selection will however trigger more competition for words sharing onsets than endings. Such predictions will thus differ according to the inflectional properties of a given language (see the discussion above on the neighborhood position effects and the role of morphology). This view is in line with previous research showing that there is more interference when producing word pairs that overlap at the beginning (e.g., "storage" - "story") than words that rhyme (e.g., "glory"- "story"; O'Seaghdha & Marin, 2000; see also Sevald & Dell, 1994; Sullivan & Riffel, 1999; Wheeldon, 2003; Yaniv, Meyer, Gordon, Huff, & Sevald, 1990).

A recent simulation study by Chen & Mirman (2012) showed how the same lexical connectionist framework could account for PhND effects to be inhibitory in some cases and facilitatory in others. In their simulations, the direction of the effect depended on the strength of the co-activated neighbors: an inhibitory effect was observed when they were strongly activated (in their proposal, during word recognition) and a facilitative effect was observed when they were weakly activated (in their proposal, during word production). Regarding the high overlap of semantic and phonological similarity in Spanish (as discussed earlier), it would be reasonable to assume that phonological neighbors become activated strongly enough to exert an inhibitory effect on lexical selection in word production. In this context, it is also important to highlight the marginally significant effect of neighborhood frequency that we observed independently of the PhND effect. In the current study, words with activated neighbors of high mean frequency were named more slowly than words with neighbors of low mean frequency. This finding indicates that the neighbors of the to-be produced word behave as strong lexical competitors when their activation level is increased, as presumably indexed by high mean frequency values.

In contrast to the theoretical frameworks above, it is not obvious how to account for the independent influence of PhND on word production in terms of a feed-forward model of speech production with a monitoring system (Levelt, 1989; Levelt et al., 1999). One could hypothesize that monitor processes are sensitive to neighborhood structure. Words from dense neighborhoods are overall more word-like within the lexical network, and thus they might pass the monitor more quickly compared to words from sparse neighborhoods. However, this prediction is at odds with the present results that showed slower naming times for words from dense than sparse phonological neighborhoods.

Alternatively, a monitoring account of neighborhood effects could build on the general hypothesis that monitoring processes rely on the speech comprehension system[11]. In studies of speech comprehension, PhND manipulations have shown cross-linguistic differences, with facilitatory effects in auditory word recognition in Spanish (Vitevitch & Rodriguez, 2005) and inhibitory effects in

[11] We thank a reviewer for bringing up this point.

English (e.g., Vitevitch & Luce, 1998). Thus, such an account could explain the inhibitory effect of PhND in English speech production (if it were to be confirmed, see above), but could not explain the inhibitory effect in Spanish that is reported in the current article.

Finally, feed-forward architectures with a monitoring system cannot easily account for empirical evidence from studies on speech errors (i.e. a facilitatory effect). Nooteboom (2005) extended the monitoring account by postulating that the monitor carries out comparisons based on phonological similarity between the error and its intended form. Thus, the more an error sounds like its intended word, the more likely it is to be mistaken and to slip through the monitor (see also Slevc & Ferreira, 2006). When committing an error on a dense neighborhood word, one is more likely to produce another real word, since words from dense neighborhoods are overall more similar to other words in the lexicon. Accordingly, errors stemming from dense neighborhood words would be less likely to be filtered out by the monitoring system. This prediction contrasts with the published results showing that participants commit fewer errors on words from dense neighborhoods. Altogether, we conclude that predictions derived from the monitoring account about PhND, whether in standard naming or error-elicitation tasks, do not fit the empirical evidence. Further explicit assumptions would be needed to explain how the monitoring system could explain the effect of PhND on word production.

Relationship to Other Effects in Word Production

There are two additional findings that have to be discussed in the light of an inhibitory PhND effect. Both findings can be described as facilitatory effects driven by phonological similarity, and therefore they are in apparent contrast with the inhibitory effect observed in the present study. First, bilingual speakers name cognate words (translations having a high phonological overlap like "tomate" [Spanish] -"tomato" [English]) faster than non-cognates (e.g., "mesa" [Spanish] – "table" [English]; Costa, Caramazza, & Sebastián-Gallés, 2000). This means that the high phonological overlap across translations facilitates naming performance. The present findings could only be considered consistent with the facilitative cognate effect if one assumes that the lexical representations of a bilingual's two languages do not engage in cross-linguistic lexical competition (see Runnqvist, Strijkers, Alario, & Costa, 2012, for extensive discussion).

The second observation refers to a phonological facilitation effect reported in studies using the picture-picture paradigm (e.g., Morsella & Miozzo, 2002). In this task participants are asked to name a picture while ignoring a superimposed distractor picture whose name is phonologically-related to the target. Results showed that naming latencies were shorter for pictures that bear a phonological relationship with the distractor picture than for pictures that have no relation. This finding indicates that speakers access the phonological information of unattended stimuli and that this overlap in phonology facilitates the production of the intended word. One caveat here is that the phonological relationship between the tested words is not as precisely defined as in standard picture naming studies. Results of these picture-picture studies would be more comparable to the present results if the picture words would be either phonological neighbors (i.e., differing by only one phoneme) or controlled for positional overlap (i.e., only differing at the onsets or endings). Moreover, although participants were selectively attending to the target picture, the representation of the distractor was nevertheless processed at early perceptual stages (Bles & Jansma, 2008). One may need to differentiate between paradigms in which phonologically similar words are externally pre-activated or primed, from paradigms in which phonological activation spreads naturally within the network (i.e., without explicit distractors or primes as in the present study). This discrepancy points out again the importance of distinguishing task properties of the experimental paradigms at hand.

Conclusion

In the present study, we observed that word production is slower for words with many similar sounding neighbors. We suggest that this detrimental effect of phonological similarity is due to competitive processes during lexical selection. In addition, we highlight that the effect of PhND varies depending on task and performance measure. Dense PhND plays a facilitative and protective role with respect to lexical search and error rates, whereas it slows naming. This provides a view of the mental lexicon as a dynamic entity whose properties influence performance depending on the way it is accessed.

2.3 Breaking down the bilingual cost in speech production

Sadat, J., Martin, C.D., Magnuson, J., Alario, F.-X., & Costa, A. (in preparation). Breaking down the bilingual cost in speech production. *Manuscript at Aix-Marseille University and Pompeu Fabra University.*

Breaking down the bilingual cost in speech production

Jasmin Sadat [a, b], Clara D. Martin [a, c, d], James S. Magnuson [e],
F.-Xavier Alario [b], & Albert Costa [a, f]

[a] Departament de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra, Barcelona, Spain

[b] Laboratoire de Psychologie Cognitive, CNRS and Université Aix-Marseille, France

[c] Basque Center on Cognition, Brain and Language, Donostia-San Sebastian, Spain

[d] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

[e] Department of Psychology, University of Connecticut, Storrs, USA

[f] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Running head: BILINGUAL COST IN SPEECH PRODUCTION

Address correspondence to:

Jasmin Sadat

Departament de Tecnologies de la Informació i les Comunicacions

Universitat Pompeu Fabra

C/ Roc Boronat 138

08018 Barcelona, Spain

Phone: +34 - 93 542 13 81

Email: jasmin.sadat@upf.edu

Abstract

Bilinguals have been shown to perform worse than monolinguals in several linguistic tasks. The current study investigated this bilingual cost in a large-scale picture naming study in Spanish. We compared naming performance of Spanish-Catalan bilinguals speaking in their dominant and non-dominant language to Spanish monolinguals. In particular, we focused on the contributions of lexical frequency and phonological overlap across translations to explain the bilingual cost. Naming latencies were analyzed by means of linear mixed-model models accounting for individual effects at the participant and item level. Our results showed that the most important predictor determining the bilingual cost in speech production was phonological overlap across translations. The bilingual cost disappeared when naming phonological highly overlapping translations. In turn, increasing lexical frequency was shown to decrease naming latencies in mono– and bilinguals, but seemed to be less influential for explaining the bilingual cost in speech production. Implications of our results for the effects of phonological overlap across translations within the bilingual language production system are discussed.

Acknowledgements

This research was supported by a grant from the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 Grant agreement n◦ 263575), two grants from the Spanish government (PSI2008-01191, Consolider Ingenio 2010 CSD2007-00012), and the Catalan government (Consolidado SGR 2009-1521). Jasmin Sadat was supported by a pre-doctoral fellowship from the Spanish Government (FPU-2008). Clara Martin was supported by. We thank Laure Minier and Elena Guerra for checking a portion of the vocal responses, and Paula Marne for help with the Spanish vocabulary ratings.

Breaking down the bilingual cost in speech production

Being able to communicate in two languages is an obvious asset, but being a bilingual also has some negative consequences on linguistic performance. Recent research on bilingual language production has documented several costs in bilinguals compared to monolingual speakers. Bilinguals show slower naming latencies, take longer to articulate, make more errors, and experience more tip-of-the-tongue states than monolinguals (e.g., Gollan & Silverberg, 2001; Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007; Gollan, Montoya, Cera, & Sandoval, 2008; Gollan & Goldrick, 2012; Guion, Flege, Liu, & Yeni-Komshian, 2000; Kohnert, Hernandez, & Bates, 1998; Mackay & Flege, 2004; Roberts, Garcia, Desrochers, & Hernandez, 2002). Remarkably, these costs surface not only in a bilingual's non-dominant language (L2), but even in his or her first learnt and dominant language (L1; e.g., Ivanova & Costa, 2008; Sadat, Martin, Alario, & Costa, 2012). Given that costs have been observed in the two languages of a bilingual, it is assumed that a common mechanism proper to the fact of mastering two languages is at the source of such costs.

Since speaking in several languages has become an increasingly common mode of communication, it is important to understand the linguistic cost that is associated with it. Despite a growing body of research on bilingual language production, our knowledge about the origin of this phenomenon remains rather limited. Many studies have transferred knowledge about the main variables found to determine monolingual speech production to bilingual language processing. Yet the extent to which these variables influence bilingual speech production or how exactly they relate to a bilingual cost is still a matter of debate. In the current article, we will explore and manipulate the most prominent variables thought to influence bilingual speech performance, namely lexical frequency and phonological similarity across translations (also known as cognate status). These two variables are the most studied variables in the context of bilingual speech processing. Nevertheless it still remains to be established how they contribute to the bilingual cost in language production. Thus, in the present study, we aimed at tracking down the bilingual cost with respect to the most important variables that influence language production in bilinguals.

Over the years, three explanations have emerged to account for a bilingual cost in speech production. However, the prevalence of each of these accounts still remains vague, since more empirical evidence on the bilingual cost is needed to evaluate them. In the present study, we will consider the phenomenon of a bilingual cost and investigate how the key factors proposed to be important for bilingual language performance relate to it. This approach will give more or less weight to each of the frameworks that have been put forward to explain the bilingual cost. In what follows we will outline the three explanations of the bilingual cost in language production. We will then describe the approach of the present study to break down the phenomenon of the bilingual cost into the most important variables that are suggested to contribute to it.

The first account of the bilingual cost relies on the general principle of frequency effects in speech production. It assumes that higher frequency usage strengthens the links between a concept and its lexical representation which in turn leads to faster word retrieval for production. Since bilinguals use each of their two languages less frequently than monolinguals use their only language, bilinguals will have weaker links in each of their languages ("weaker links hypothesis", Gollan et al., 2008). As a result of this frequency lag, bilinguals will be slower in retrieving a word for production relative to monolinguals. The weaker links hypothesis makes clear predictions with respect to the frequency effect and modulations of the bilingual cost, claiming that the size of the frequency effect depends on language usage. This means that the frequency effect in language production should decrease with an increasing usage of that language. More specifically, the weaker links account predicts that low-frequency words suffer most from a reduced usage. Due to the logarithmic shape of the frequency effect, a low frequency word like “kite” should suffer more from not being used than a high frequency word like “house”. This means that reduced language usage will have less of an impact on high than low frequency words. Thus, the weaker links hypothesis makes two testable predictions: First, the frequency effect should increase from monolingual speakers to bilinguals speaking in L1 to bilinguals speaking in L2. Second, the bilingual cost should be maximal in the case of low frequency words and non-existent or pretty much reduced in the case of high frequency words, at least for L1 bilinguals. Hence, one of the important variables that we will consider in the present study is lexical frequency.

A second explanation claims that delays in bilingual language processing stem from processing stages posterior to lexical access (Indefrey, 2006; Hanulová, Davidson, & Indefrey, 2011). This post-lexical account mainly stems from of a review of the literature on studies comparing naming performance between non-native and native speech. In their review, Hanulová and colleagues state that electrophysiological and hemodynamic evidence on non-native language processing suggest a locus of the bilingual cost after lexical word form retrieval. The authors propose that the bilingual cost results from more demanding processes at level of phonological and phonetic encoding, syllabification, and/or articulation. To address predictions of the post-lexical account, variables capturing information from phonological or articulatory processing should be considered when investigating the bilingual cost. Note, however, that it is difficult to specify variables that exclusively index processing at post-lexical stages. Most variables thought to influence late stages of production such as lexical frequency also affect speech processing at earlier levels (e.g., Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Gahl, 2008). Altogether, it is difficult to establish clear predictions for this account, because the post-lexical account commits to the locus from where the bilingual cost should originate, but it is less specific about the mechanisms underlying them.

Finally, an alternative explanation postulates that part of the bilingual cost originates from language control processes (executive control account, Green, 1998). When speaking, bilinguals constantly need to resolve potential competition between their two language systems. This competition process is thought to slow down bilingual lexical access compared to monolinguals. Recent bilingual research has linked the ability to resolve conflict between languages to a more general type of conflict resolution. Following this idea, it seems important to consider the contributions of general executive control differences between monolinguals and bilinguals when investigating speech production. Thus, the ability of non-linguistic conflict resolution will be included as a covariate in the present study when accounting for speech performance differences between monolinguals and bilinguals.

The Present Study

From the above explanations on the bilingual cost, it is primarily the variable lexical frequency that can be retained to explore performance differences between mono- and bilinguals. However, there is one additional item-related variable that is specific to bilingual language production and that has been shown to be very influential, namely cognate status. Cognate words are translations sharing a high phonological overlap across languages like for example “tomato” (English) and “tomate” (Spanish) as compared to words that do not share many sounds ("orange"

[English] and "naranja" [Spanish]). Several studies have reported that cognate words show a clear processing benefit over non-cognates in production (i.e., cognate effects, see Costa, Caramazza, & Sebastián-Gallés, 2000). Thus, the amount of phonological overlap across translations is a highly influential variable in bilingual speech production. Nevertheless, the effect of this variable has been largely neglected in the context of explaining the bilingual cost. Thus, we hypothesize that one important aspect to explain difficulties during bilingual speech is to consider the extent to which a word bears more or less phonological similarity with its translation.

In this vein, we will also focus on the influence of phonological similarity beyond translations. One may suppose that the number of phonological similar words of the unspoken language would influence the production of the intended word in bilinguals (similar as for phonological neighbors in monolingual speech). It is widely assumed that during bilingual speech production the two languages interact at the level of phonology (Colomé, 2001; Colomé & Miozzo, 2010; Costa et al., 2000). Thus, we used a variable that measured phonological neighbors of the unspoken language as an additional predictor to assess phonologically similarity effects beyond translations in bilingual speech production.

In the present study, we tested 60 bilinguals (30 L1 and 30 L2 speakers) in a large-scale picture naming experiment and compared their performance to a group of 30 monolingual speakers from a previous study (Sadat, Martin, Costa, & Alario, submitted). Our rational was to use the monolingual group described in Sadat et al. as a control group for the bilinguals tested in the current study. Therefore, we used identical materials, design and procedure as in the previous study. We included the most relevant item-related predictor variables to evaluate how each of these variables influences mono- and bilingual speech production. Given that we focused on variables capturing the properties of the words to produce (i.e., item-related variables), we tested a large set of 533 items. This allows us to go beyond dichotomic measures and to account for effects over a continuous range of values. Thus, we included item-related variables such as lexical frequency and translation overlap as continuous measures to better account for naming performance along the continuum of values. In addition, we also included several participant-related predictors like individual measures of executive control (as suggested by the executive control account), vocabulary size (e.g., Bialystok, Craik, & Luk, 2008a; Luo, Luk, & Bialystok, 2010), and socio-economic status (e.g., Morton & Harper, 2007) as control variables across speaker groups to account for possible group differences due to these variables.

The naming data of mono- and bilinguals was analyzed by means of linear mixed regression modeling (Baayen, Davidson, & Bates, 2008). This method applies analysis at a single trail level and thus offers more insights on individual effects and variations. Therefore one part of the analysis explored the effects of frequency and translation overlap at the individual level. This approach is further motivated by one possible confound of the weaker links account. Increased frequency effects are not only observed in participants with decreased language usage, but are general to slower reaction times at which participants perform a task (e.g., see Balota & Ferraro, 1993; Balota & Ferraro, 1996; Cerella, 1985; Spieler & Balota, 2000, for showing a larger frequency effect in slower response times of older adults). In assessing individual effects, we can provide more detailed information on the effects of lexical frequency and how they relate to individual response speed. It may be that larger frequency effects in bilinguals compared to monolinguals are due to the mere fact that bilinguals show slower naming times. In addition, we will also analyze the effect of translation overlap at the individual level and compare it to the effect of lexical frequency. An effect of translation overlap clearly depends on the status of the speaker (i.e., mono- versus bilingual or L1 versus L2 bilingual), but less obviously on individual response speed (unlike lexical frequency). The analyses at the individual level will help clarifying the origin of differences in the size of lexical frequency effects across speaker groups.

Method

Participants

Sixty Spanish-Catalan bilinguals were recruited. Participants were undergraduate students at the University Pompeu Fabra, Barcelona, Spain. They were all highly proficient in Spanish and

Catalan (see Appendix B for a description of the bilingual community in Catalonia). Importantly, all bilinguals were early highly proficient, but unbalanced speakers: They were all tested in Spanish which was the dominant language for half of them (L1 bilinguals) and the non-dominant language for the other half (L2 bilinguals; see Table 3 for language history and proficiency ratings). None of the participants were fluent in any other language. All participants had normal or corrected-to-normal vision and were matched on age. They received 20 Euros for participating in the experiment.

Materials

The stimuli were 533 black-and-white line drawings of common objects (identical to those used in Sadat et al., submitted). They had black outlines and white surfaces and were presented 300 pixels wide x 300 pixels high on a white rectangle with a monitor resolution of 800 x 600 pixels.

For the multiple regression analysis, the following item-related predictors important for speech production were collected from Sadat et al. (submitted):

- Word form frequency with values for written lexical frequency (range [0.07, 2.80] log occurrences per million, M = 0.98, SD = 0.54). A logarithmic transformation was applied to avoid the undue influence of extreme values in the regression.

Table 3: Means (*M*) and standard deviations (*SD*) for participant self-report ratings.

	30 Spanish monolinguals		30 Spanish-Catalan bilinguals		30 Spanish-Catalan bilinguals	
	M	*SD*	*M*	*SD*	*M*	*SD*
Age	22	2	21	2	22	3
Percent daily use of Spanish	99	4	63	18	30	19
Age exposed to Spanish	0	0	0	0	1	1
Age exposed to Catalan	-	-	1	2	0	1
Spanish proficiency	4.8	0.5	4.9	0.3	4.9	0.4
Catalan proficiency	-	-	4.6	0.6	5.0	0.2

Note: The measure "Percent daily use of Spanish" was obtained by asking participants to estimate their daily language use of Spanish, Catalan, and any other language with the constraint that their sum equals 100 percent. "Age exposed to Spanish/Catalan" refers to the mean age at which participants were continuously exposed to these languages. Proficiency ratings are on a 1–4 scale, where 1 indicates ''very little knowledge of the language'' and 4 indicates ''native proficiency". Proficiency values represent the average of the participants' responses in four domains (speech comprehension, speech production, reading, and writing).

- Name agreement (range [10, 100] %, M = 85, SD = 18).
- Subjective estimates of age-of-acquisition (AoA) from adult ratings (range [2.3, 10.3] years; M = 4.6, SD = 1.4).
- Phonological neighborhood density in Spanish (PhND; range [0, 37] number of neighbor words, M = 5, SD = 7). It refers to the number of words in Spanish that can be formed from a given word by substituting, adding or deleting one phoneme (Luce, 1986).
- Word length measured in phonemes (range [2, 11] number of phonemes, M = 6, SD = 2).
- First syllable frequency (range [0.08, 4.61] log occurrences per million, M = 3.33, SD = 0.91). A logarithmic transformation was applied to avoid the undue influence of extreme values in the regression.

In addition, the following item-related predictors specific to bilingual language production were collected:

- Translation overlap measured by Levenshtein editing distance (range [0.6, 100] %, M = 26, SD = 13). It calculates how many phonemes of a word have to be changed to transform it into its translation and captures the amount of editing difference between two words (Schepens, Dijkstra, & Grootjen, 2012). This measure was standardized and expressed in percentages.

- Translation overlap measured by ALINE (Kondrak, 2000; range [0, 86] % of overlap, M = 54, SD = 19). It aligns the phonetic sequences to be compared and assigns similarity values to the common phonemes across translations. To find the best match of strings, it uses the phonetic similarity of surface forms (e.g., “alcachofa” [Spanish for artichoke] needs to be aligned 2 steps to the right to best match “carxofa” [Catalan]). In addition, ALINE associates different weights to each phoneme pair according to its saliency. This salience constraint leads for example to higher weight assignments in the case of identical consonant sounds over identical vowel sounds. This measure was standardized and expressed in percentages.
- Phonological neighbors in Catalan (range [1, 126] number of words, M = 15, SD = 17). It is defined as the number of Catalan lemmas that can be formed for each of the Spanish stimuli words by the substitution, addition or deletion of a single phoneme at any position within the Spanish word. The number of Catalan lemma neighbors was estimated on the basis of a corpus of 137,028 Catalan words.

In addition to the item-related predictors, several participant-related predictors were obtained (these predictors were also available for the monolingual participants of Sadat et al., submitted):

- Executive control measures and impulsivity speed. Participants performed an adaptation of the Simon task (Simon & Wolf, 1963). In this task, participants were asked to manually respond as fast and as accurate as possible to the color of a visually presented cue on the screen. The color cue appeared either on the same side as the required response button (congruent trials), in the center of the screen (neutral trials), or on the opposite side as the required response button (incongruent trials). Interference effects were calculated by subtracting average performance in incongruent trials from congruent trials (Monolinguals: *M* = -46 ms, *SD* = 21; L1 bilinguals: *M* = -42 ms, *SD* = 21; L2 bilinguals: *M* = -36 ms, *SD* = 20). Impulsivity speed was assessed by averaging the reaction times on the neutral trials (Monolinguals: *M* = 423 ms, *SD* = 39; L1 bilinguals: *M* = 407 ms, *SD* = 46; L2 bilinguals: *M* = 416 ms, *SD* = 47).
- Socio-economic status. Participants completed a questionnaire on their socio-economic status with eleven questions (see Appendix C; results ranging from 14 – 35 points; Monolinguals: *M* = 22, *SD* = 4; L1 bilinguals: *M* = 22, *SD* = 5; L2 bilinguals: *M* = 26, *SD* = 4).

- Vocabulary size. Participants completed a Spanish vocabulary-size test (WAIS-III vocabulary subtest with 33 definitions; Wechsler, 1997). A Spanish speaker evaluated the participants' answers according to the test instructions with zero to two points per definition (Monolinguals: *M* = 43, *SD* = 4; L1 bilinguals: *M* = 45, *SD* = 3; L2 bilinguals: *M* = 43, *SD* = 4).

Design and Procedure

As in Sadat et al. (submitted), participants were randomly assigned to one of six experimental lists and were tested in a sound-proof room. Stimulus presentation and the software voice-key were controlled via DMDX (Forster & Forster, 2003). The sensitivity of the voice-key was adjusted for each participant. Each trial started with a fixation cross displayed at the centre of the computer screen for 500 ms. After a 300 ms blank screen, the picture of the object to name was displayed. The picture remained on the screen until either the voice key detected the response or a 2500 ms deadline was reached without any overt response detected. The next trial began 700 ms after the recording period finished.

The experiment consisted of a short training followed by two sessions that were separated by a break of 15 minutes. First, participants were asked to name eight practice black-and-white pictures similar to the materials used in the experiment. They were instructed to name the pictures as fast and as accurately as possible using single nouns. After that, in the first session, they had to name the whole set of 533 object pictures divided into eight short blocks. The responses were monitored by the experimenter and if participants gave another name for the picture than the intended one, they were corrected by the experimenter at the end of the first session. In the second session, the same 533 pictures were presented in the same way as in the first session, but in a different order. Participants' responses were automatically recorded by the computer as digitized sound files, and errors were noted online by the experimenter. Each session lasted about 45 minutes. In total, the experiment including breaks lasted about two hours.

Data analyses

All 63,960 vocal responses and speech onset markers (533 pictures x 2 presentations x 60 participants) were visually checked offline with the software Check-Vocal (Protopapas, 2007) and corrected if necessary. Responses other than the intended one were classified as errors and excluded from onset latency analyses.

We performed several preliminary examinations on the predictors to check for importance of variables in predicting bilingual naming latencies and to reduce multicollinearity. We first ran random forest analyses (e.g., Breiman, 2001) using the package party in R and the function cforest (Hothorn, Bühlmann, Dudoit, Molinaro, & Van Der Laan, 2006; Strobl, Boulesteix, Zeileis, & Hothorn, 2007; Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008). A random forest is a collection of classification trees providing a single measure of importance for each predictor. Regarding inter-language similarity, the random forest analysis clearly rated the ALINE measure as being of higher importance in predicting bilingual naming latencies than the Levenshtein distance. The measure of phonological neighbors across languages was rated as least important toward predicting naming latencies.[12] All continuous predictor variables that were involved in interaction terms were centered to avoid multicollinearity. Due to the high level of collinearity between PhND and phoneme word length (even after centering), the latter was orthogonalized on the former by running a linear model in which phoneme word length was predicted by PhND. The residuals of this linear model were entered as fixed effect in the linear mixed-effects model. The same procedure was applied to first syllable frequency by replacing it with the residuals obtained from a linear model in which first syllable frequency was predicted by PhND.

Onset latencies and accuracy rates were analyzed by mixed regression models at the single trial level (Baayen et al., 2008). In addition to fixed predictors considered in simple linear regressions, linear mixed-effects models account for random variation induced by specific words or speakers. All statistical analyses were run with the statistical software R (R Development Core Team, 2012) and

[12] We ran several linear mixed effect models testing for the influence of cross-language phonological neighborhood in the bilingual naming latencies. Results showed no influence of this predictor on bilingual naming latencies. We will further discuss this null result in the Discussion section.

linear mixed-effects models were computed with the package lme4 in R (Bates, Maechler, & Bolker, 2012). P-values were validated by Markov chain Monte Carlo simulations using the function pvals.fnc in the package "languageR" (Baayen, 2011) whenever possible (i.e., in the case of simple by random intercept models). The Box-Cox test (using the function boxcox in the package MASS in R, Venables & Ripley, 2002) indicated that the reciprocal transformation of the latencies was the most appropriate transformation for the data to reduce skewness and approximate a normal distribution. We used -1000/RT as an order preserving transformation to facilitate the interpretation of our results.

In the first linear mixed-effects model, three variables were included as predictors. Variables accounting for general fatigue and/or training effects in the naming latencies by coding for trial order presentation (from 1 to 533) and session (first or second), and the main factor speaker group (monolinguals vs. L1 speaker vs. L2 speaker) to test for the bilingual naming cost. Thereafter, participant-related variables like socio-economic status, vocabulary size, executive control measure, and impulsivity speed were included as fixed effects in the model. This is to ensure that the speaker group difference would still be present beyond the inclusion of these control variables. Finally, predictors motivated by previous naming studies on monolingual naming performance were entered as fixed effects (name agreement, AoA, PhND, first syllable frequency, phoneme length). Importantly, the variables of interest to explain the bilingual cost being lexical frequency and inter-language overlap were included in the model. Variables significantly contributing to the model’s fit were retained and different models were compared by means of log-likelihood tests. We followed Baayen (2008) procedure of model criticism in which trials whose standardized residual value is above 2.5 were removed and the model was recomputed.

Equation 1 describes the structure of the final model in which naming latencies were predicted as a linear function of intercept (b_0), the control variables ($b_{1\text{-}3}$; i.e., participant-related variables, trial order, and session), the item-related variables ($b_{4\text{-}8}$), the speaker group factor ($b_{9\text{-}10}$), and all interactions of the group factor with the item-related variables ($b_{11\text{-}20}$). The item-related variables of interest to explain the bilingual cost (i.e., frequency, inter-language similarity) were entered with reference points at (1) high frequency values in Model 3 and (2) high translation overlap values in Model 4 to estimate their contributions to the speaker group effect. Table 4 presents an overview of all models. In what follows, a sequence of increasingly complex models is reported.

$$RT = b_0 +$$

$$b_1\,\text{Trial order}\begin{bmatrix}533\\ \vdots\\ 268\\ \vdots\\ 1\end{bmatrix} + b_2\,\text{Session}\begin{bmatrix}2\\ 1\end{bmatrix} + b_3\,\text{Impulsivity speed}\begin{bmatrix}539\\ \vdots\\ 408\\ \vdots\\ 327\end{bmatrix} +$$

$$b_4\,\text{NameAgr}\begin{bmatrix}0.8\\ \vdots\\ 0\\ \vdots\\ -4.8\end{bmatrix} + b_5\,\text{AoA}\begin{bmatrix}3.2\\ \vdots\\ 0\\ \vdots\\ -2.3\end{bmatrix} + b_6\,\text{PhND}\begin{bmatrix}4.6\\ \vdots\\ 0\\ \vdots\\ -0.7\end{bmatrix} +$$

$$b_7\,\text{Freq}\begin{bmatrix}1.8\\ \vdots\\ 0\\ \vdots\\ -0.9\end{bmatrix} + b_8\,\text{TrOverlap}\begin{bmatrix}0.3\\ \vdots\\ 0\\ \vdots\\ -0.5\end{bmatrix} +$$

$$b_9\,\text{Group}\begin{bmatrix}0\\ 1\\ 0\end{bmatrix} + b_{10}\,\text{Group}\begin{bmatrix}0\\ 0\\ 1\end{bmatrix} +$$

$$b_{11}\,\text{Group}\begin{bmatrix}0\\ 1\\ 0\end{bmatrix} * \text{Freq}\begin{bmatrix}1.8\\ \vdots\\ 0\\ \vdots\\ -0.9\end{bmatrix} + b_{12}\,\text{Group}\begin{bmatrix}0\\ 0\\ 1\end{bmatrix} * \text{Freq}\begin{bmatrix}1.8\\ \vdots\\ 0\\ \vdots\\ -0.9\end{bmatrix} + \ldots +$$

$$b_{19}\,\text{Group}\begin{bmatrix}0\\ 1\\ 0\end{bmatrix} * \text{TrOverlap}\begin{bmatrix}0.3\\ \vdots\\ 0\\ \vdots\\ -0.5\end{bmatrix} + b_{20}\,\text{Group}\begin{bmatrix}0\\ 0\\ 1\end{bmatrix} * \text{TrOverlap}\begin{bmatrix}0.3\\ \vdots\\ 0\\ \vdots\\ -0.5\end{bmatrix}$$

Equation 1: Structure of the model tested in the present analyses. It includes control variables, the relevant item-related predictors, speaker group, and the interactions between the item-related predictors and speaker group. The bilingual cost was estimated using centered predictors except for the predictors of interest lexical frequency and translation overlap that were entered with maximal values in model 3 and 4 respectively. Freq = lexical frequency; TrOverlap= translation overlap; NameAgr = name agreement; AoA = age-of-acquisition; PhND = phonological neighborhood density.

Results

Analysis of naming latencies

After removing errors and non-target responses (i.e., responses that did not match the intended target name of the picture), 55,249 responses remained for analyses. We added the 28,046 error-free and offline checked responses from the monolingual speakers tested in Sadat et al. (submitted) to be used as a control group. We used a linear mixed effect model with by-item random slopes for speaker group. After outlier removal, the basic model was recomputed on 82,101 responses (see Model 1 in Table 4). The average naming latency was 910 ms (SD = 118) for monolinguals, 947 ms (SD = 92) for L1 bilinguals, and 963 ms for L2 bilinguals (SD = 93). Trial order and Session were significant, showing that responses to pictures became slower with increasing trial order and that responses in the second session were faster than in the first. There were significant differences in naming latencies between monolinguals and L1 and between monolinguals and L2 bilinguals. In contrast, the difference between L1 and L2 bilinguals was not significant. This establishes that both L1 and L2 groups showed the bilingual cost which is at stake in this article.

After adding the participant-related predictors, the bilingual cost was unchanged. The measures of executive control and socio-economic status were not significant. The measures impulsivity speed and vocabulary size were significant (see Model 2 in Table 4). Participants who were fast responders in a button-press task were also faster in naming pictures. Participants with higher vocabulary scores in Spanish had faster naming latencies. However, as indicated by log-likelihood tests, vocabulary size did not significantly improve the model's fit, and thus was not retained in the next model.

All item-related predictors were added into the following model (Model 3 in Table 4). Based on the predictions of the weaker links hypothesis on the bilingual cost, lexical frequency was entered with a reference at its highest value (i.e., 2.8 log occurrences per million). Results showed that the difference between monolinguals and L1 and L2 bilinguals remained significant (see Figure 5). The difference between L1 and L2 bilinguals was not significant. As expected, there were significant effects of AoA, name agreement, lexical frequency, and PhND in the monolingual speakers: earlier learned words were named faster than later learned ones, and words with high percentages of name agreement were named faster than words with lower percentages. Naming latencies increased with lower frequency values of the words and higher numbers of phonological neighbors. These item-related predictors interacted significantly with speaker group. Both bilingual groups showed a smaller effect than monolinguals for name agreement and AoA. L1 bilinguals showed a smaller effect of PhND. Regarding the effect of lexical frequency, L1 bilinguals did not show a significant difference in the effect of lexical frequency when compared to monolinguals, but L2 bilinguals showed a significantly larger frequency effect than monolinguals and L1 bilinguals ($\beta = -1.84 \times 10^{-2}$, $SE = 4.22 \times 10^{-3}$, $t(82069) = 4.35$, $p = < 0.001$; VIF = 1.9). The effects of word length measured in phonemes and first syllable frequency were not significant.

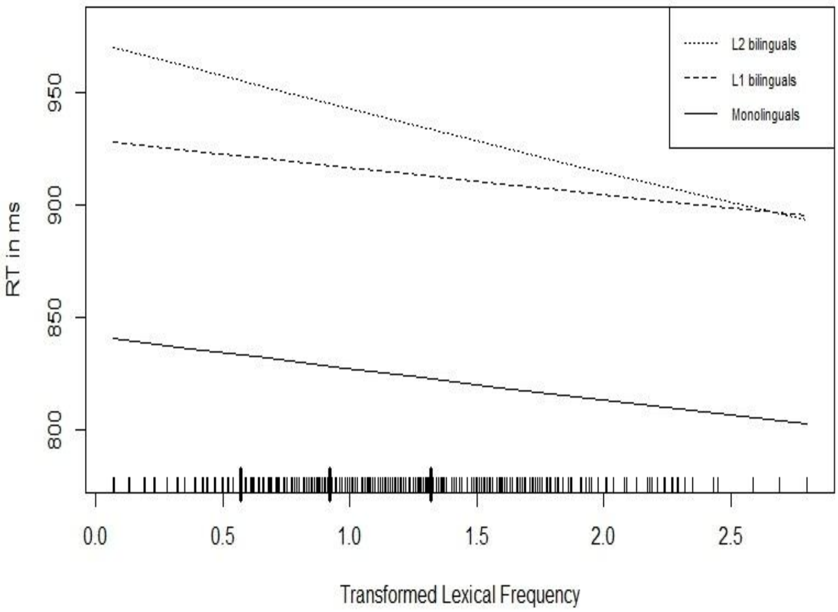

Figure 5: Estimation of speaker group effects when the model was referenced at highest frequency values. The difference between mono– and bilinguals was significant along the continuum of the lexical frequency variable. The distribution of the variables is depicted along the x-axis with the bold lines indicating the 1st, 2nd, and 3rd quartiles.

In the next step of the analyses, we added translation overlap to the model with the reference point at its highest value (i.e., 86% translation overlap; see Model 4 in Table 4). Importantly, L1 and L2 bilinguals were not significantly different from monolinguals (see Figure 6). Neither was the difference between L1 and L2 bilinguals. As expected, translation overlap interacted significantly with speaker group. There was no translation overlap effect in monolinguals (since they do not have translations), but L1 and L2 speakers both showed significant effects. Bilingual latencies were faster for words with more translation overlap. This effect was significantly stronger in L2 than L1 bilinguals ($\beta = 4.63 \times 10^{-2}$, $SE = 1.06 \times 10^{-2}$, $t(82075) = 4.39$, $p = < 0.001$; VIF = 1.4).

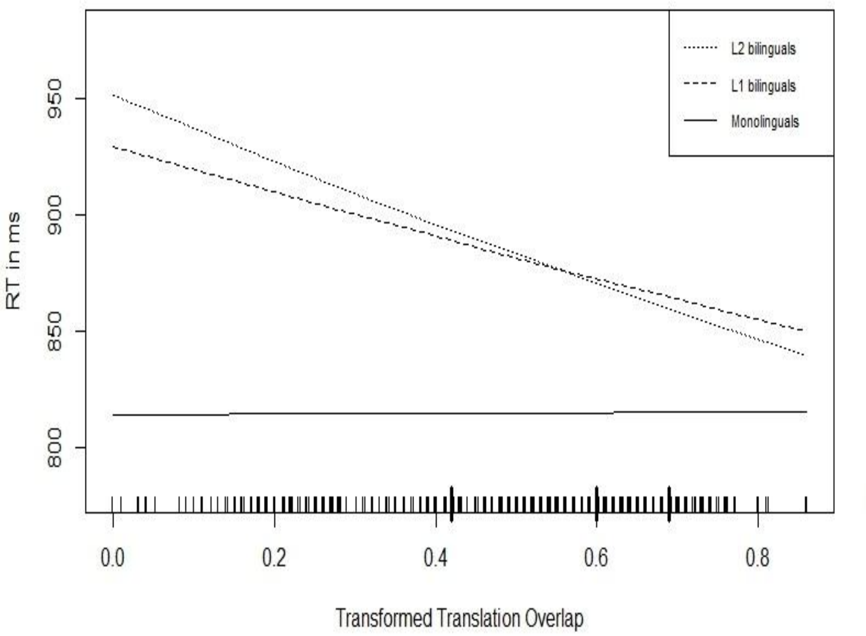

Figure 6: Estimation of speaker group effects when the model was referenced at highest translation overlap values. The speaker group difference was significant when the model was referenced at the lower end or at the center of the translation overlap variable. In contrast, the speaker group difference was non-significant when estimated for translation overlap referenced at higher values. The distribution of the variables is depicted along the x-axis with the bold lines indicating the 1st, 2nd, and 3rd quartiles.

Table 4: Overview of increasingly complex linear mixed models estimating speaker group effects in naming latencies

	Model 1			Model 2			Model 3			Model 4		
Predictor	**raw β**	**SE β**	**sign**	**raw β**	**SE β**	**sign**	**raw β**	**SE β**	**sign**	**raw β**	**SE β**	**sign**
Intercept	-1.18	$2.42x10^{-2}$	***	-1.36	$1.95x10^{-1}$	***	-1.63	$1.25x10^{-1}$	***	-1.63	$1.26x10^{-1}$	***
Session	$-9.89x10^{-2}$	$1.61x10^{-3}$	***	$-9.79x10^{-2}$	$1.62x10^{-3}$	***	$-9.83x10^{-2}$	$1.62x10^{-3}$	***	$-9.87x10^{-2}$	$1.62x10^{-3}$	***
Trial order	$1.21x10^{-4}$	$5.20x10^{-6}$	***	$1.21x10^{-4}$	$5.25x10^{-6}$	***	$1.21x10^{-4}$	$5.23x10^{-6}$	***	$1.21x10^{-4}$	$5.23x10^{-6}$	***
L1	$5.53x10^{-2}$	$3.22x10^{-2}$	**	$8.82x10^{-2}$	$3.21x10^{-2}$	**	$8.22x10^{-2}$	$3.17x10^{-2}$	**	$4.61x10^{-2}$	$3.19x10^{-2}$	
L2	$7.75x10^{-2}$	$3.22x10^{-2}$	**	$7.57x10^{-2}$	$3.25x10^{-2}$	**	$6.28x10^{-2}$	$3.14x10^{-2}$	**	$1.24x10^{-2}$	$3.16x10^{-2}$	
SES				$2.03x10^{-3}$	$2.93x10^{-3}$							
Simon				$6.88x10^{-4}$	$6.18x10^{-4}$							
Impulsivity				$9.82x10^{-4}$	$2.86x10^{-4}$	***	$9.35x10^{-4}$	$2.87x10^{-4}$	***	$9.35x10^{-4}$	$2.88x10^{-4}$	***
Vocabulary				$-5.87x10^{-3}$	$3.41x10^{-3}$	**						
PhonLength							$7.11x10^{-3}$	$6.53x10^{-3}$				
SylFreq							$-1.11x10^{-2}$	$1.31x10^{-2}$				
AoA							$5.25x10^{-2}$	$6.40x10^{-3}$	***	$5.35x10^{-2}$	$6.26x10^{-3}$	***
L1 * AoA							$-2.21x10^{-2}$	$2.38x10^{-3}$	***	$-1.91x10^{-2}$	$2.40x10^{-3}$	***
L2 * AoA							$-2.27x10^{-2}$	$2.40x10^{-3}$	***	$-1.83x10^{-2}$	$2.41x10^{-3}$	***
Nagr							$-1.01x10^{-1}$	$5.43x10^{-3}$	***	$-9.98x10^{-2}$	$5.34x10^{-3}$	***
L1 * Nagr							$1.63x10^{-2}$	$2.20x10^{-3}$	***	$1.78x10^{-2}$	$2.20x10^{-3}$	***
L2 * Nagr							$2.53x10^{-2}$	$2.23x10^{-3}$	***	$2.74x10^{-2}$	$2.23x10^{-3}$	***
PhND							$1.39x10^{-2}$	$5.51x10^{-3}$	**	$1.42x10^{-2}$	$5.43x10^{-3}$	**
L1 * PhND							$-4.73x10^{-3}$	$2.05x10^{-3}$	**	$-5.63x10^{-3}$	$2.05x10^{-3}$	**
L2 * PhND							$-2.27x10^{-3}$	$2.06x10^{-3}$		$-3.45x10^{-3}$	$2.06x10^{-3}$	*
FreqHF							$2.01x10^{-2}$	$1.13x10^{-2}$	**	$2.07x10^{-2}$	$1.11x10^{-2}$	**
L1 * FreqHF							$-5.03x10^{-3}$	$4.19x10^{-3}$		$-6.38x10^{-3}$	$4.18x10^{-3}$	
L2 * FreqHF							$1.33x10^{-2}$	$4.23x10^{-3}$	**	$1.19x10^{-2}$	$4.22x10^{-3}$	**
TrOlap										$-2.02x10^{-3}$	$2.70x10^{-2}$	
L1 * TrOlap										$-1.20x10^{-1}$	$1.04x10^{-2}$	**
L2 * TrOlap										$-1.66x10^{-1}$	$1.05x10^{-2}$	**

Note: SE=standard error; sig=significance level (*** is p<0.001, ** is p<0.055, * is p<0.101); L1 = L1 bilinguals; L2= L2 bilinguals; SES= socio-economic status; AoA=age of acquisition; Nagr = name agreement; PhND =phonological neighborhood density; FreqHF=transformed lexical frequency; TrOlap= translation overlap; PhonLength=phoneme length; SylFreq=first syllable frequency; TrOlapC=transformed translation overlap.

In the last step of the analysis, we focused on individual variations in the effects of frequency and translation overlap and their relation to speed of response. To do so, we run a linear mixed effect model with random intercepts and random slopes for frequency and translation overlap effects for each participant. In this way we obtained estimates of individual intercepts and individual effects for lexical frequency and translation overlap. In this model, the factor speaker group was omitted to allow the model to fit individual effects independent of any speaker group assignment. Furthermore, we estimated the individual effects in a model in which lexical frequency and translation overlap were referenced on their lowest values (i.e., low frequency and non-cognate words) to observe a maximal the bilingual cost.[13] We then run a linear regression model in which these individual frequency and translation overlap effects were predicted by the individual speed of the participants (i.e., the previously calculated individual intercept) and the factor speaker group (see Figure 7 and 8).

[13] The pattern of results regarding the speaker group effects remained unchanged when individual effects were calculated from models with centered or highest values as reference points. However, the effect of individual speed was significant in a model with centered or highest values of translation overlap as reference. Moreover, no effects were significant in a model with centered frequency values as reference point. These variations in the influence of individual speed relate to the shape of the underlying effects, being nonlinear for both variables.

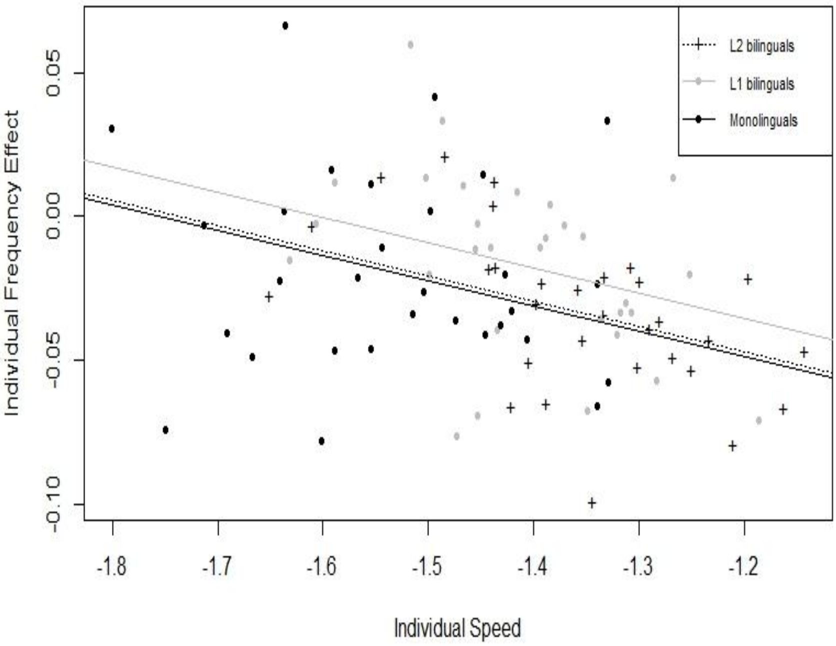

Figure 7: Results of the linear regression models on the individual effects of lexical frequency. The graph shows the relationship between magnitude of the effect and response speed estimated at the individual level. Each point represents a participant. Larger values in the x-axis indicate slower responding participants. Negative values in the y-axis indicate a negative coefficient for the effect (i.e. facilitatory effect). Note that the lexical frequency effect decreased for faster participants and that there was no difference between mono- and bilinguals.

The results of the first regression model on individual frequency effects showed a significant effect of individual response speed, but not of speaker group (see Table 5). The model accounted for 12% of the variance ($F[3,86]=5.10$, $p = 0.003$). Results showed that individual frequency effects are best accounted for by individual speed measures. The faster a participant's responses the smaller was the corresponding lexical frequency effect. The opposite pattern of results was shown by the second regression model on individual translation overlap effects: speaker group was a significant predictor, but not individual response speed (see Table 5). The model accounted for 54% of the variance ($F[3,86]=35.26$, $p < 0.001$). Results showed that translation overlap effects did not vary with individual response speed, but depended on the speaker group (monolingual vs. L1 vs. L2 bilingual).

Table 5 : Linear regression model on individual effects of lexical frequency and translation overlap.

	Lexical frequency			Translation overlap		
	Raw b	SE b	p	Raw b	SE b	p
Intercept	-0.15	0.04	< 0.001	-0.03	0.08	0.734
Individual Speed	-0.09	0.03	0.002	-0.01	0.05	0.881
L1 bilinguals	0.01	0.01	0.116	-0.10	0.02	< 0.001
L2 bilinguals	< 0.01	0.01	0.844	-0.14	0.02	< 0.001

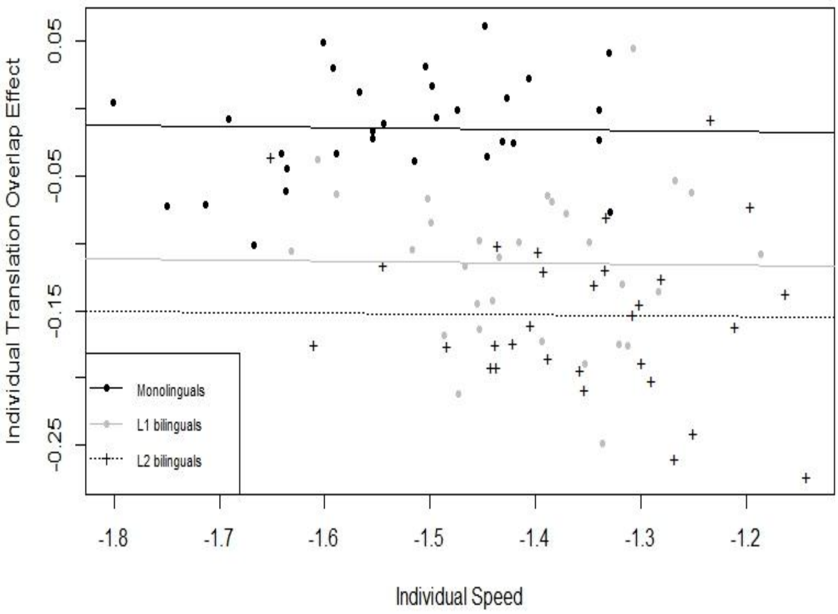

Figure 8: Results of the linear regression models on the individual effect of translation overlap. The graph shows the relationship between magnitude of the effect and response speed estimated at the individual level. Each point represents a participant. Larger values in the x-axis indicate slower responding participants. Negative values in the y-axis indicate a negative coefficient for the effect (i.e. facilitatory effect). Note that the effect of translation overlap was facilitatory for the bilingual speakers only.

Analysis of accuracy rates

Responses containing speech errors (Monolinguals: 1,681 trials; L1 bilinguals: 1,484 trials; L2 bilinguals: 2,048 trials) were contrasted with error-free responses (Monolinguals: 28,046 trials;

L1 bilinguals: 28,024 trials; L2 bilinguals: 27,225 trials) to accurately predict the probability of an error-free response. The generalized linear mixed-effects model included the predictors session, speaker group, all relevant item-related centered predictors (i.e., name agreement, lexical frequency, AoA, translation overlap), and their interactions with speaker group. After outlier removal, the model was recomputed on 86,268 responses. The odds of error were significantly larger for L2 bilinguals than L1 bilinguals (β = 1.02, SE = 0.39, z[86250] = 2.61, p = 0.009; VIF = 1.4) and monolinguals (β = 1.10, SE = 0.39, z[86250] = 2.83, p = 0.005; VIF = 1.4). The difference between monolinguals and L1 bilinguals was not significant. The odds of error for monolinguals were smaller for words that had higher name agreement (β = -1.65, SE = 0.09, z[86250] = -17.70, p = < 0.001; VIF = 1.4), higher lexical frequency (β = -0.45, SE = 0.12, z[86250] = -3.70, p = < 0.001; VIF = 1.0), and were learned earlier (β = 1.13, SE = 0.12, z[86250] = 9.18, p = < 0.001; VIF = 1.6). The interactions between these item-related predictors and speaker group were significant. Bilinguals showed reduced effects for name agreement (for L1: β = 0.22, SE = 0.06, z[86250] = 3.71, p = <0.001; VIF = 1.9; for L2: β = 0.58, SE = 0.06, z[86250] = 10.48, p = <0.001; VIF = 1.9) and AoA (for L1: β = -0.41, SE = 0.08, z[86250] = -4.88, p = <0.001; VIF = 2.0; for L2: β = -0.25, SE = 0.08, z[86250] = -3.24, p = 0.001; VIF = 2.1). The interaction between speaker group and lexical frequency was significant, showing that the odds of errors for L1 and L2 bilinguals were less than for monolinguals with increasing frequency of the words (for L1: β = -0.39, SE = 0.09, z[86250] = -4.11, p = <0.001;

VIF = 1.3; for L2: β = -0.42, SE = 0.08, z[86250] = -4.97, p = <0.001; VIF = 1.7). There was no significant difference regarding the frequency effect between the two bilingual groups. As expected, translation overlap did not had an effect on the odds of error in monolinguals, but L1 and L2 bilinguals were less probable to make errors with increasing translation overlap of the words (for L1: β = -0.56, SE = 0.06, z[86250] = -9.31, p = <0.001; VIF = 1.5; for L2: β = -0.88, SE = 0.06, z[86250] = -15.41, p = <0.001; VIF = 1.6). The difference between the two bilingual groups was significant, showing that the odds of error were smaller for L2 bilinguals than L1 bilinguals with increasing translation overlap (β = -0.31, SE = 0.06, z[86250] = -5.69, p = <0.001; VIF = 1.5). The effect of session was significant (β = -2.17, SE = 0.06, z[86250] = -37.65, p = <0.001; VIF = 1.0). There were no significant effects of word length, first syllable frequency, and PhND.

Discussion

In the present study, we compared naming performance of mono- and bilingual speakers in a large-scale picture naming study. We aimed at better describing the variables that may underlay the linguistic cost previously reported in bilingual language production (e.g., Gollan & Silverberg, 2001; Gollan et al., 2007; Gollan et al., 2008; Gollan & Goldrick, 2012; Guion et al., 2000; Kohnert et al., 1998; Mackay & Flege, 2004; Roberts et al., 2002). We included item- and participant-related variables to predict bilingual naming performance and explored which of these variables modulated the bilingual cost. Specifically, our analyses focused on the contributions of the item-related predictors lexical frequency and translation overlap, at the group and individual level. This is important because of the highly influential role of these variables on naming performance in bilinguals and because it is widely undefined how they relate to the bilingual cost.

First of all, the general observation of faster naming latencies for monolinguals than bilinguals was replicated. These results are in line with various picture naming studies showing the bilingual cost for speakers of L2 (e.g., Gollan et al., 2008; Ivanova & Costa, 2008; Kohnert et al., 1998) and even L1 (Ivanova & Costa, 2008; Sadat et al., 2012). This result provided the basis for further investigations of the key variables underlying the bilingual cost.

Regarding lexical frequency, our findings did not reveal a modulation of the bilingual cost depending on this variable. That is, the bilingual cost for L1 and L2 remained significant along the continuum of the lexical frequency variable, and even when frequency was referenced to the highest frequency value (see Figure 5). According to the weaker links account, the bilingual cost should be substantially reduced in the case of high frequency words. Moreover, our results did not show slower naming responses for L2 than L1 bilinguals as could be expected by the weaker links account. However, we observed an interaction of speaker group by lexical frequency showing that the size of the lexical frequency effect was larger in L2 than in L1 bilinguals and monolinguals. L2 speakers were slower in naming lower frequency words which is predicted by the weaker links account.

The analysis on individual effects provided further insights on the role of lexical frequency for the bilingual cost. In a separate analysis, we aimed at disentangling the contributions of individual response speed and lexical frequency for each participant. Our results showed that speaker status (mono- vs. L1 vs. L2 bilingual) had no influence on the individual frequency effect, further supporting the idea that lexical frequency was not responsible for the bilingual cost. Instead, individual response speed was identified as the most important predictor towards the individual frequency effect. In other words, the overall speed of a participant's responses to the pictures determined the size of the lexical frequency effect. The effect of lexical frequency was independent of whether the participant was monolingual or bilingual. Thus, lexical frequency clearly influenced naming behavior in mono- and bilinguals, but it does not seem to be a key factor for the bilingual cost in speech production. Interestingly, a recent study by Diependaele, Lemhöfer, & Brysbaert (2012) showed that the most important predictor to explain a larger frequency effect in bilingual word recognition was vocabulary size, unlike speaker group (i.e., mono- vs. bilingual). Note that the measure vocabulary size was not included in our model in which individual lexical frequency and translation overlap effects were calculated. Thus, it is possible that the significant effect of individual response speed on the individual frequency effect may be accounted for by the vocabulary scores of the participants.[14]

Importantly, our results point to a highly influential role of phonological overlap across translations to explain the bilingual cost. A significant effect of speaker group was present in all the models tested here, except when we included a predictor that coded for phonological overlap across translations. Note that cross-linguistic phonological overlap was referenced on the highest overlap value (i.e., cognates) to estimate the bilingual cost. In this case, L1 and L2 bilinguals were not significantly different from monolinguals. This means that in the case of highly overlapping translations, no bilingual cost was observed and bilinguals performed similar to monolinguals. The bilingual cost then increased significantly as translation overlap decreased (see Figure 6). This finding points to translation overlap as one of the main sources to the bilingual cost.

Previous literature on the relation between the bilingual cost and cognate status of the words is sparse. In a post-hoc analysis, Ivanova and Costa (2008) assessed the effects of cognate status on the bilingual cost and observed that the bilingual cost were similar for cognates and non-cognates (in their study, cognates: 32 ms; non-

---

[14] We ran an additional analysis of the individual frequency effects in which our estimate of vocabulary size was included. The result showed no significant effect of vocabulary size. However, the lack of an effect might be clearly related to the coarse nature of our vocabulary estimates as compared to Diependaele et al. (2012).

cognates: 35 ms). This result contrasts with our present findings that showed that the bilingual cost disappeared in the case of highly overlapping translations (i.e., cognate words; see also Christoffels, Firk, & Schiller, 2007, for a reduced, but still significant, bilingual cost for non-cognates). However, as stated by Ivanova and Costa, their results should be interpreted with caution, since they were assessed post-hoc on a small and unbalanced set of stimuli. An important role of cognate status in bilingual naming performance has been reported in tasks relying on active word finding strategies. For example, in a verbal fluency study by Sandoval et al. (2010), bilinguals produced fewer exemplars of members of a semantic category than monolinguals. Interestingly, the bilingual cost was entirely driven by the fact that bilinguals named fewer non-cognates than monolinguals (see Figure 5, Sandoval et al.). Furthermore, several studies eliciting tip-of-the-tongue states reported no bilingual cost in tip-of-the-tongue states for cognate words or proper names (Gollan & Silverberg, 2001; Gollan & Acenas, 2004; Gollan, Montoya, & Bonanni, 2005). These findings show that high phonetic overlap between words (or identity as in the case of proper names) equalizes naming performance of mono- and bilinguals in tasks requiring active word search strategies. More evidence on the relation between phonetic effects and the bilingual cost comes from a recent study by Gollan and Goldrick (2012). They asked English monolinguals, early highly proficient Mandarin-English bilinguals, and early highly proficient Spanish-English bilinguals to repeat aloud tongue twisters consisting of non-words and words (tongue twisters are combinations of phonetically similar segments). Their results showed that overall both bilingual groups produced more twist errors than monolinguals. Importantly, bilinguals also showed more errors in the case of non-word twisters, suggesting a sub-lexical locus of the bilingual cost. This means that the bilingual cost is present in situations even when no word retrieval is involved and only phonetic processing is required. These results are in line with the present study showing that experience with an additional sound structure in bilinguals can hamper language production compared to monolinguals.

In addition to previous studies, we also explored the influence of participant-related variables on bilingual naming performance. The only significant participant-related variables were vocabulary size and impulsivity speed. Contrary to some studies claiming influencing effects of executive control ability, vocabulary size, and socio-economic status (e.g., Bialystok et al., 2008; Luo et al., 2010; Morton & Harper, 2007), no clear significant effects of these variables were observed. This might be due to the number of participants tested in the present study. As the focus of the present study was to explore the effect of item-related predictors by using a large set of stimuli, a very large number of participants would also be needed to explore the effect of participant-related predictors in more detail. Another explanation might be the rather coarse nature of estimating participant related measures in the present study. The usage of more adapted and elaborated methods will be of advantage to better determine the most important participant-related variables of language processing in bilinguals (see e.g., Diependaele et al, 2012, for detailed measures of vocabulary size).

Our findings are also in line with a post-lexical account of the bilingual cost (Indefrey, 2006; Hanulová et al., 2011), since they indicate costs deriving from phonological processing. However, our results do not speak to whether the bilingual cost originates from the level of phonology or not. In what follows, we will discuss how phonological similarity across translations is suggested to affect language production in bilinguals and describe the possible origins of the effect.

On the origin of the effect of phonological similarity across translations in speech production of bilinguals

The prevalent explanation in the literature on cognate effects attributes them to sublexical effects and processes occurring during lexical retrieval (Costa et al., 2000; Costa, Santesteban, & Caño, 2005; but see Sánchez-Casas & García-Albea, 2005, or Van Hell & De Groot, 1998, for explanations at the morphological and conceptual level). Due to interactivity in the speech production system (Dell, 1986), high phonological overlap would facilitate lexical retrieval of the word associated with the shared phonemes when compared to words that do not overlap in phonemes. Our findings are in line with this view, showing faster bilingual naming latencies with increasing phonological translation overlap. In addition, our results suggest that the amount of phonological overlap across translations is essential to explain the bilingual cost. Thus, bilinguals do not benefit from translation overlap relative to monolinguals, but rather suffer from non-overlapping translations. This may suggest two sources of the bilingual cost: On the one hand, it is possible that non-cognate words trigger interference at the lexical level, and thus slow down speech in bilinguals compared to monolinguals. For now and with the current data, we cannot unequivocally discard this possibility. On the other hand, it is possible that the slowing of non-cognates is due to the internal properties of the language system. Lexical representations of non-cognates are overall more widespread than cognates due to less phonological similarity. This relies on the general idea that more similar words are represented in a closer network in the mental lexicon. We suggest that the "compactness" of representations for highly overlapping translations in bilinguals is similar to words represented in a monolingual system. Only in this case, selection and retrieval of the lexical item would be accomplished without an additional cost.

Recent findings on monolingual speech production have shown that phonological similar words within one language slowed down lexical retrieval of an intended word (e.g., Gordon, 2011; Sadat et al., submitted). Applying this logic to bilinguals, it would mean that a translation with high phonological overlap would be detrimental to bilingual speech production, since it would strongly compete. Thus, for translation overlap effects to be facilitatory, one has to assume that the lexical representation of the translation does not compete for lexical selection (e.g., Costa & Caramazza, 1999). Moreover, it is not clear whether processing benefits in bilinguals could result exclusively from the converging activation at the phonological level, and independent of conceptual overlap. Surprisingly, our results on bilingual latencies showed no effect of purely phonological inter-language influences as assessed by the number of phonological neighbors in the unspoken language (see also De Groot, Borgwaldt, Bos, & van den Eijnden, 2002; Lemhöfer et al., 2008, for no effect of orthographic inter-language neighbors). When bilinguals produced a word, there was no effect of the number of phonologically similar words of the unspoken language. Thus, inter-language influences of phonological similarity seem to be restricted to translations (but see Costa, Hartsuiker, & Roelstraete, 2006). This missing influence from phonological similar words in the unintended language suggests that translation overlap effects cannot solely derive from phonological influences across languages. Instead, explanations for translation overlap effects should consider additional assumptions relying on shared conceptual or lexical representations.

One alternative explanation to cognate effects is to understand them as frequency effects in disguise (Sandoval et al., 2010; Strijkers, Costa, & Thierry, 2010). Following this idea, each time a cognate word is used, irrespective of the language of response, it would augment the frequency of use of this word. This implies that lexical frequency values for bilinguals would need to incorporate an extra dimension and to adjust for cognate status. The lexical frequency value of a non-cognate should be estimated even lower than based purely on bilingual language usage, whereas the value of a cognate should be estimated equally high as for monolinguals. In this sense, bilinguals would not show a cost in the case of phonological highly overlapping words when compared to monolinguals, because they would have the same frequency values. Costs would emerge when translations do not match and frequency values do not sum up as it is the case for non-cognates. Evidence in favor of such a view of translation overlap effects comes from a study by Strijkers et al. (2010). In their study using event-related potentials, the authors found divergent amplitudes at a similar point for the effects of lexical frequency, translation overlap, and speaking in L1 or L2. The authors suggested that these three effects may stem from similar processes during lexical access of bilinguals. This view of the cognate effect is supported by our results because in the case of cognates no difference between mono- and bilinguals was found, contrary to non-cognates for which a bilingual cost was present.

Altogether, our results provide detailed information on language production in bilinguals with respect to important item-related variables such as lexical frequency and translation overlap. Several studies have reported poorer performance of L1 and L2 bilinguals compared to monolinguals, but evidence on the influence of the underlying variables was unclear. The present study identified at least one main variable of the bilingual cost, that is phonological similarity across translations. The results of the current study have important implications for different types of bilingualism. It would mean that bilinguals speaking phonologically more similar languages would overall experience less of a cost than bilinguals speaking two phonologically more distant languages compared to monolinguals. Indeed, if we compare the studies that have previously assessed the bilingual cost in different populations of bilinguals, differences can be observed in the sizes of the effect (see Table 1 in Hanulová et al., 2011). As can been seen in the review by Hanulová et al., the bilingual cost is at a minimum for Spanish-Catalan bilinguals, whereas more distant language combinations (e.g. Spanish-English) showed a larger cost. One caveat here is that the studies also investigated different populations of bilinguals. Most studies tested successive dominant bilinguals (i.e., they learned one language after the other with a large time delay between the two; e.g., Christoffels et al., 2007) or switched-dominant bilinguals (i.e., their fist learnt and dominant language became the non-dominant language over time, Gollan et al., 2005; Gollan et al., 2008), whereas the present study assessed performance in dominant, but early and relatively simultaneous bilinguals. The results of the study by Gollan and Goldrick (2012) in which tongue twister production was compared across Spanish-English, Mandarin-English and English only groups, indicate that there were differences in the error patterns at sub-lexical processing between the two bilingual groups. Mandarin-English bilinguals showed a more consistent cost compared to monolinguals over all tongue twister conditions than did Spanish-English bilinguals. This finding is fits the proposal of the present study, suggesting that the more phonologically dissimilar the languages of a bilingual, the larger should be the bilingual cost.

One limitation of the present study is that we do not provide information on the extent to which positional effects of translation overlap are important. Similarly as in the literature on PhND effects in monolinguals, it can be suspected that similarity effects across translations differ depending on the position of the overlap (e.g., beginning vs. end). For example it may be the case that a highly overlapping translation would only induce a cost when the overlap occurred at the beginning as opposed to the end. The use of our translation overlap measure (ALINE, Kondrak, 2000) provided positional alignment between translations. This approach to similarity assignments is superior to string-edit approaches like the Levenshtein edit distance, because it searches optimal alignment between phonetic strings (and because it considers phonetic features). Thus, we suspect that positional aspects of overlap (among other things) seem to be important when assessing the influence of phonological similarity across translations. However, further detailed investigations are needed on how translation overlap effects may vary according to their position of the overlap.

Conclusion

The present study establishes a direct relation between two important phenomena of bilingual speech production: the cognate facilitation effect and a linguistic cost observed in bilingual speech production compared to monolinguals. We argue that the amount of phonological overlap across translations determines the bilingual cost, up to a disappearance of the cost in the case of phonological highly similar translations. This provides a new view on the involvement of phonological similarity in the bilingual cost and suggests an important role of sub-lexical features in the organization of the bilingual mental lexicon.

2.4 Characterizing the bilingual disadvantage in noun-phrase production

Sadat, J., Martin, C.D., Alario, F.-X., & Costa, A. (2012). Characterizing the bilingual disadvantage in noun phrase production. *Journal of Psycholinguistic Research, 41*(3), 159-179.

3. GENERAL DISCUSSION

The goal of the present dissertation was to provide a more detailed description of the bilingual cost and its underlying processes. In particular, we were interested in the effects observed at phonological and articulatory levels of speech production and how they would relate to the bilingual cost. This approach was motivated by the observation of strong inter-language influences at these late levels of speech processing (see also Hanulová et al, 2011; Indefrey, 2006). Our predictions were that phonological and articulatory processes would provide important insights about the origin and the scope of bilingual costs. We expected the bilingual cost to vary depending on the amount of phonological similarity across translations, and to extend over a range of performance measures and speech contexts. The experimental section of this dissertation presented three studies addressing these issues. Monolingual and bilingual naming performances were assessed by measuring onset latencies (in the first two studies) or onset latencies as well as articulatory durations (in the last study). Special attention was paid to the effects of variables indexing phonological similarity within and across languages and to the articulatory durations of the utterances. The findings of the present dissertation can be summarized as follows:

- The first study showed that increased phonological similarity within one language slowed down speech production. This result contrasts with the facilitatory effect of phonological similarity across translations in the speech of bilinguals. We conclude that explanations for the effect of phonological similarity within one language cannot by themselves account for the effect of phonological similarity across translations.

- The second study showed that phonological similarity across translations was a major determinant of the bilingual cost. That is, the bilingual cost decreased with increasing phonological overlap across translations. We conclude that phonological similarity across translations helps bilinguals to overcome the bilingual cost.

- The third study showed that the bilingual cost was apparent in the articulatory durations of L1 bilinguals and generalized to noun phrase production. We conclude that the bilingual cost is apparent in a variety of performance measures and speech contexts, and represents a non-negligible factor in speech production.

3.1 The influence of phonological properties on speech production in mono- and bilinguals

The first and the second study explored the influence of phonological properties on speech production in mono- and bilinguals respectively. The first study assessed the effect of phonological neighborhood density in monolingual speech. This approach was motivated by previous proposals of the supposedly similar role of phonological neighborhood density and translation overlap in both groups of speakers. Effects of both variables have been explained by shared phonological representations feeding back on lexical representations (Costa et al., 2005; Runnqvist, FitzPatrick et al., 2012). A detailed investigation of the effect of phonological neighborhood density was also motivated by the fact that previous research in monolinguals reported conflictive evidence (facilitatory vs. inhibitory effects of phonological neighborhood density). The results of our study showed that the increasing phonological similarity of a given word relative to other words within one language slowed down naming latencies. This result was confirmed by several analyses on additional datasets of monolingual speech production. The finding of increased naming latencies for words with increasing phonological similarity was explained by lexical competition among words sharing sounds with the word to utter. Moreover, we pointed out that the effect of phonological similarity depended on the task at hand or on the way the mental lexicon was accessed. This means that although phonological similar words slow down naming latencies within a given language, they nevertheless facilitate word retrieval in tasks that emphasize accuracy.

The second study investigated the role of phonological similarity across translations in speakers producing words either in L1 or L2 as compared to monolinguals. We showed that increasing translation overlap substantially reduced and eliminated the bilingual cost. L1 and L2 speakers showed a cost for naming less phonological similar translations compared to monolinguals. In contrast, there was no bilingual cost when naming highly overlapping translations. This result indicated that translation overlap is one of the major determinants to the bilingual cost in speech production.

The relationship between the effect of phonological properties within and across languages deserves further attention. In our first study, we showed a clear inhibitory effect of phonological similarity on onset latencies within one language. At first sight this finding contradicts the results of the second study in which we showed that increasing phonological similarity of translations helped bilinguals to overcome the bilingual cost. According to the results of our first study, co-activation of a highly similar sounding translation should be detrimental for bilingual naming. This is because the two lexical representations of translations would be highly activated (due to phonological and semantic overlap), and so they should compete during word selection. Thus, for a translation overlap effect to be facilitatory one basic assumption is that the lexical representation of the translation does not compete for lexical selection in bilinguals. This idea is in line with accounts claiming that lexical selection is language-specific (Costa & Caramazza, 1999; Costa et al, 1999).

One explanation to reconcile the different effects of phonological similarity across the two studies is based on the fact that translation overlap relies on both semantic and phonological overlap, whereas phonological neighborhood density only relies on the latter. Thus, it seems that the effect of phonological similar translations is mainly carried by the additional semantic overlap, that is absent in phonological neighbors.

Phonological similarity is essential regarding the way in which the mental lexicon is organized. As we pointed out in the first study, phonological similarity within one language clearly shows facilitatory effects in tasks emphasizing the accuracy of word retrieval. This implies that form similarity plays an important role for word retrieval from the mental lexicon. More similar words come from highly inter-connected networks and access to them is facilitated through the items with which they are connected. The same organization principle can be applied to the bilingual mental lexicon. More similar words within and across languages would be highly inter-connected and lexical access would be facilitated when compared to less similar ones. Phonological similarity across translations would help retrieving a word, because of highly inter-connected representations between the two language systems. The idea of increased connectivity for more similar items relates to a recent explanation of the effect of translation overlap as a frequency effect in disguise (Sandoval et al., 2010; Strijkers et al., 2010). Accordingly, phonologically similar translations would have increased activation levels, because over time they would receive more feedback from their shared phonological representations compared to non-overlapping translations. As a result, they would be accessed and retrieved faster compared to translations that do not share sounds.

One important caveat regarding the claim that pure phonological overlap would facilitate lexical access in bilinguals comes from study two. Our results showed no effect of the number of phonological neighbors in the unspoken language of bilinguals. Thus, inter-language influences of phonological similarity seems to be restricted to translations in our data. This finding contrasts with the results of a study by Costa, Hartsuiker, Roelstraete (2006). In a speech error inducing paradigm, bilingual participants were more likely to commit an error that resulted in a real word either in the spoken or unspoken language (i.e., lexical bias effect). The authors suggested that phonological activation was feeding back on the lexical representations of the two languages, thereby causing a lexical bias effect independent of language. Although we are cautious with the interpretation of the null result in our data, it suggests that pure form similarity across languages may not be enough to influence speech onsets in the intended language (see Costa, La Heij, & Navarrete, 2006, for a discussion on how feedback from the phonological level of the unintended language may influence speech production in bilinguals). However, given that form similarity influences the way lexical representations are accessed, it may nevertheless influence speech performance when different strategies of word retrieval are required (similar as the reported task dependent effect of phonological similarity within one language). This would mean that for example phonological similarity of false friends (e.g., constipado [Spanish for having a cold] vs. constipation) would not be activated enough to influence onset latencies when naming in one language, but they may provide enough cues to resolve word retrieval failures when experiencing a tip-of-the-tongue state.

We conclude that phonological similarity is an important parameter for the structure of the mental lexicon and the retrieval of words. Moreover, our results suggest a special role of translations in the bilingual mental lexicon due to their conceptual and phonological overlap. We showed that speech production, and in particular, the bilingual cost depended on phonological properties of the word to utter. The complex relationship between phonological similarity effects within and across languages requires further detailed investigations.

3.2 Contributions of articulatory durations to the bilingual cost

The third study addressed the bilingual cost in speech articulation. Our results showed that a bilingual cost was present in the articulatory durations of L1 bilinguals. In addition, we extended findings of a bilingual cost to noun phrase production by assessing onset latencies and articulatory durations of speech. Thus, our findings suggest that the bilingual cost is non-negligible in language production and apparent in different performance measures and speech contexts.

The evidence that a bilingual cost is apparent in articulatory durations calls for further investigations regarding its origin. In study two, we showed that one of the main factors to explain the bilingual cost in onset latencies was translation overlap. Thus, by taking an exploratory approach to the data of study three, we tried to determine post-hoc the influence of translation overlap on articulatory durations. Given that there was no clear interaction with the variable translation overlap and the L1 bilingual group in study three, we included an additional dataset of 70 L2 bilinguals to our analyses (i.e., 35 in each production context). L2 speakers usually show clear traces and increased effects of translation overlap effects compared to L1 bilinguals (e.g., Costa et al., 2000; Ivanova & Costa, 2008). We followed the same logic and analysis as applied in study two. We used a continuous measure to quantify the amount of phonological overlap across translations for the picture names tested in study three (ALINE, Kondrak, 2000). In doing so, we aimed to explore how translation overlap would relate to the bilingual cost in articulatory durations of speech production.

This reanalysis on the data of study three confirmed the findings of the previously reported bilingual cost with a new analysis method. Bilingual articulatory durations showed a cost when compared to monolinguals, and this cost was especially apparent in noun phrase production. By applying the same logic as in study two, the bilingual cost in articulation was estimated with respect to the variable translation overlap (see Appendix E for detailed results). Importantly, L2 bilinguals showed clear interactions with translation overlap in both production contexts.

However, this interaction was of opposing directions in the two production contexts. In bare noun production, highly overlapping translations had longer articulatory durations compared to non-overlapping translations. In contrast, in noun phrase production highly overlapping translations had shorter articulatory durations compared to non-overlapping translations. Although we have no clear explanation what may have caused this opposing pattern of the effect of translation overlap across production contexts, it indicates that L2 bilingual articulatory durations were sensitive to the effect of translation overlap. It may be that the additional speech planning processes of noun phrase production are responsible for the changed effect of translation overlap on articulatory durations.

Interestingly, the bilingual cost in articulatory durations remained significant when estimated with translation overlap referenced at highest overlap. This suggests that the bilingual cost in articulatory processing arises over and above the effect of translation overlap. L2 bilinguals showed longer articulatory durations than monolinguals, independent of translation overlap. Although these results are assessed post-hoc and need to be interpreted with caution, they suggest that the cost in articulatory durations may result from an additional source compared to the cost in onset latencies. It may be that the cost in articulatory durations is driven by phonetic inter-language influences that are absent in monolinguals. This idea is in line with studies showing changed voice-onset-times during bilingual articulations when compared to monolinguals (e.g., Caramazza et al., 1973; Flege, 1987; Flege & Eefting, 1987; Fowler et al., 2008). Inter-language influences from the underlying phonetic representations may constitute the major part of the bilingual cost in articulation (e.g., Guion et al., 2000). We may tentatively conclude that there is an additional independent cost for bilingual speakers emerging at late stages of speech articulation. This would be in line with the proposal by Runnqvist et al. (2011), suggesting that the bilingual cost most probably emerges during lexical access, but also affects subsequent stages of speech processing. Exploring the origin behind the bilingual cost in articulatory durations deserves further detailed investigations.

3.3 Final remarks

The present dissertation addressed the question of how the fact of mastering two languages affects speech performance. We provided new evidence from studies of phonological and articulatory processes in single word and noun phrase production. Moreover, we highlighted the essential role of phonological properties in speech production in mono- and bilinguals.

Several studies have shown that phonological similarity of a word relative to other words in the lexicon helps vocabulary acquisition, both in mono- and bilinguals. In the case of monolinguals, it has been shown that children acquire new words faster when they resemble already known words (e.g., Storkel & Morrisette, 2002). Similarly, bilingual speakers seem to naturally exploit the information provided by form similarity in translations. This is suggested by studies showing that foreign language vocabulary acquisition is facilitated by a word’s form similarity to the L1 word (e.g., Hall, 2002). Moreover, tip-of-the-tongue studies demonstrate that bilinguals experience less word retrieval failures for phonologically similar words across languages when compared to non-similar words (e.g., Gollan & Acenas, 2004). Therefore, similarity relations among words should be considered in foreign language teaching contexts.

One of the sociolinguistic issues that the present dissertation addresses is how the bilingual cost observed experimentally might apply to daily speech. Does the fact that we found a bilingual cost repeatedly over experimental tasks imply that bilinguals are hampered in their communication abilities? It is apparent from the present dissertation that a bilingual cost emerges in various contexts, but it seems that their consequences (measured on a scale of milliseconds) are minimal. We showed that there are differences such as slower speech onsets and longer articulations for bilinguals compared to monolinguals. However, these differences stem from the interactions between the two languages in bilingual speakers. They result from the fact that the language system is a highly flexible and interactive system that is continuously evolving. Any changes or additional experience with languages will influence speech performance over time. Support for the idea of a continuously changing language system comes from studies reporting influences on language performance even after brief immersion periods (e.g., Linck et al., 2009; Sancier & Fowler, 1997). These studies suggest that any additional language experience is reflected in an individual's language performance.

Determining the factors that constrain the adaptivity of our language system will be a fruitful topic for future research.

References

Abu-Rabia, S., & Kehat, S. (2004). The critical period for second language pronunciation: Is there such a thing? *Educational Psychology*, *24*(1), 77-97.

Alario, F. X., Costa, A., & Caramazza, A. (2002a). Frequency effects in noun phrase production: Implications for models of lexical access. *Language and Cognitive Processes*, *17*(3), 299-319.

Alario, F. X., Costa, A., & Caramazza, A. (2002b). Hedging one's bets too much? A reply to Levelt(2002). *Language and Cognitive Processes*, *17*(6), 673-682.

Alario, F. X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, *36*(1), 140-155.

Alario, F. X., Goslin, J., Michel, V., & Laganaro, M. (2010). The functional origin of the foreign accent evidence from the syllable-frequency effect in bilingual speakers. *Psychological Science*, *21*(1), 15-20.

Alario, F. X., Schiller, N. O., Domoto-Reilly, K., & Caramazza, A. (2003). The role of phonological and orthographic information in lexical selection. *Brain and Language*, *84*(3), 372-398.

Amengual, M. (2012). Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition, 15*(3), 517-530.

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.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 234-254.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4*(4), 439-461.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/and English /l/and/r. *Journal of Phonetics, 32*(2), 233-250.

Arbesman, S., Strogatz, S. H., & Vitevitch, M. S. (2010). Comparative analysis of networks of phonologically similar words in English and Spanish. *Entropy, 12*(3), 327-337.

Arnold, H. S., Conture, E. G., & Ohde, R. N. (2005). Phonological neighborhood density in the picture naming of young children who stutter: Preliminary study. *Journal of Fluency Disorders, 30*(2), 125-148.

Ayora, P., & Alario, F. X. (2009). Sequential retrieval of lexical items during noun phrase production. *Poster Presented at the 50th Annual Meeting of the Psychonomic Society, Boston, Massachusetts, USA.*

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge University Press.

Baayen, R. H. (2011). languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics". R package version 1.4. http://CRAN.R-project.org/package=languageR

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.

Baayen, R. H., Piepenbrock, R., and Rijn, v. (1993). *The CELEX lexical data base on CD-ROM.* Linguistic Data Consortium, Philadelphia, PA.

Badecker, W., Miozzo, M, & Zanuttini, R. (1995). The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition*, *57*(2), 193-216.

Balota, D. A., & Ferraro, F. R. (1993). A dissociation of frequency and regularity effects in pronunciation performance across young adults, older adults, and individuals with senile dementia of the Alzheimer type. *Journal of Memory and Language*, 32, 573-573.

Balota, D. A., & Ferraro, F. R. (1996). Lexical, sublexical, and implicit memory processes in healthy young and healthy older adults and in individuals with dementia of the Alzheimer type. *Neuropsychology*, 10(1), 82.

Balota, D. A., Boland, J. E., & Shields, L. W. (1989). Priming in pronunciation: Beyond pattern recognition and onset latency. *Journal of Memory and Language, 28*(1), 14-36.

Barbón Gutiérrez, A., & Cuetos Vega, F. (2006). Efectos de la edad de adquisición en tareas de categorización semántica. *Psicológica: Revista De Metodología y Psicología Experimental, 27*(2), 207-224.

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50*(3), 560-585.

Bates, D., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-39. http://CRAN.R-project.org/package=lme4

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0. http://CRAN.R-project.org/package=lme4

Baus, C., Costa, A., & Carreiras, M. (2008). Neighbourhood density and frequency effects in speech production: A case for interactivity. *Language and Cognitive Processes, 23*(6), 866-888.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60(1), 92-111.

Ben-Zeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. *Child Development*, *48*(3), 1009-1018.

Bernstein Ratner, N., Newman, R., & Strekas, A. (2009). Effects of word frequency and phonological neighborhood characteristics on confrontation naming in children who stutter and normally fluent peers. *Journal of Fluency Disorders*, *34*(4), 225-241.

Best, W. (1995). A reverse length effect in dysphasic naming: When elephant is easier than ant. *Cortex*, *31*(4), 637-652.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008a). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics*, *21*(6), 522-538.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008b). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 859-873.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2009). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition, 13,* 525-531.

Bien, H., Baayen, R.H., & Levelt,W. J. M. (2011). Frequency effects in the production of Dutch deverbal adjectives and inflected verbs. *Language and Cognitive Processes*, 26 (4-6), 683-715.

Bien, H., Levelt, W. J. M., & Baayen, H. (2005). Frequency effects in compound production. *Proceedings of the National Academy of Sciences, 102*(49), 17876-17881.

Bles, M., & Jansma, B. (2008). Phonological processing of ignored distractor pictures, an fMRI investigation. *BMC Neuroscience, 9*(1), 20.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. *Second Language Acquisition and the Critical Period Hypothesis,* , 133-159.

Bonin, P., Chalard, M., Meot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology, 93*(1), 89-114.

Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, 35*(1), 158-167.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32.

Brennen, T., Baguley, T., Bright, J., & Bruce, V. (1990). Resolving semantically induced tip-of-the-tongue states for proper nouns. *Memory & Cognition, 18*(4), 339-347.

Brown, R., & McNeill, D. (1966). The" tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 325-337.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30*(5), 542-579.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14*(1), 177-208.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 27(6), 1430-1450.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America, 54*(2), 421-428.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(4), 857–871.

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *The Quarterly Journal of Experimental Psychology*, *25*(1), 85-95.

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin, 98*(1), 67.

Chakraborty, R., Goffman, L., & Smith, A. (2008). Physiological indices of bilingualism: Oral-motor coordination and speech rate in Bengali-English speakers. *Journal of Speech, Language and Hearing Research, 51*(2), 321.

Chen, H. C., Cheung, H., & Lau, S. (1997). Examining and reexamining the structure of Chinese-English bilingual memory. *Psychological Research, 60*(4), 270-283.

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review, 119*(2), 417-430.

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, 1147, 192-208.

Coady, J. A., & Aslin, R. N. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of Child Language, 30*(2), 441-470.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428.

Colomé, À, & Miozzo, M. (2010). Which words are activated during bilingual word production? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 96.

Colomé, À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language, 45*(4), 721-736.

Costa, A. (2005). Lexical access in bilingual production. In J. F. Kroll & A. M. B. De Groot (Eds.). *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 308-325). New York: Oxford University Press.

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition, 2*(03), 231-244.

Costa, A., & Caramazza, A. (2002). The production of noun phrases in English and Spanish: Implications for the scope of phonological encoding in speech production. *Journal of Memory and Language, 46*(1), 178-198.

Costa, A., Calabria, M., Marne, P., Hernández, M., Juncadella, M., Gascón-Bayarri, J., Lleó, Ortiz-Gil, J., Ugas, L., Blesa, R., Reñé, R (2012). On the parallel deterioration of lexico-semantic processes in the bilinguals' two languages: Evidence from Alzheimer's disease. *Neuropsychologia, 50*(5), 740-753.

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283-1296.

Costa, A., La Heij, W., & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition, 9*(2), 137–151.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41*(3), 365-397.

Costa, A., Roelstraete, B., & Hartsuiker, R. (2006). The lexical bias effect in bilingual speech production: Evidence for feedback between lexical and sublexical levels across languages. *Psychonomic Bulletin & Review, 13*, 612–617.

Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language*, *94*(1), 94-103.

Cuetos, F., Ellis, A. W., & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, 31*(4), 650-658.

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology*, *65*(2), 171-237.

Damian, M. F. (2003). Articulatory duration in single-word speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(3), 416-431.

Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods*, *37*(4), 665-671.

De Bot, K. (1992). Applied linguistics. *Applied Linguistics, 13*(1), 1-24.

De Groot, A., Borgwaldt, S., Bos, M., & van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language*, 47(1), 91-124.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283-321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language, 27*(2), 124-142.

Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon: Friends or foes? In Schiller, N., & Meyer, A. (Eds.). *Phonetics and phonology in language comprehension and production: Differences and similarities.* Germany: Mouton de Gruyter.

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior, 20*(6), 611-629.

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science, 23*(4), 517-542.

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2012). The word frequency effect in first and second language word recognition: A lexical entrenchment account. *The Quarterly Journal of Experimental Psychology.*

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.

Dufour, S., & Frauenfelder, U. H. (2010). Phonological neighbourhood effects in French spoken-word recognition. *The Quarterly Journal of Experimental Psychology*, *63*(2), 226-238.

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. J. (2008). The frequency effect in second-language visual word recognition. *Psychonomic Bulletin & Review, 15*(4), 850.

Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(2), 515-523.

Elston-Güttler, K. E., Paulmann, S., & Kotz, S. A. (2005). Who's in control? proficiency and L1 influence on L2 processing. *Journal of Cognitive Neuroscience*, *17*(10), 1593-1610.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language, 46*(1), 57-84.

Ferreira, V. S., & Griffin, Z. M. (2003). Phonological influences on lexical (mis) selection. *Psychological Science*, *14*(1), 86-90.

Finkbeiner, M., & Caramazza, A. (2006). Now you see it, now you don't: On turning semantic interference into facilitation in a Stroop-like task. *Cortex, 42*(6), 790-796.

Finkbeiner, M., Nicol, J., Greth, D., & Nakamura, K. (2002). The role of language in memory for actions. *Journal of Psycholinguistic Research, 31*(5), 447-457.

Flege, J. E. (1987). The production of ‘new' and ‘similar' phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47-65.

Flege, J. E. (1995). Second-language Speech Learning: Theory, Findings, and Problems. In W. Strange (Ed), *Speech Perception and Linguistic Experience: Issues in Cross-language research* (pp. 229-273). Timonium, MD: York Press.

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.

Flege, J. E., & Hojen, A. (2004). Why are non-native (L2) utterances longer than native-produced utterances? *The Journal of the Acoustical Society of America, 116*, 2604.

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25*(2), 169-186.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995a). Effects of age of second-language learning on the production of English consonants. *Speech Communication, 16*(1), 1-26.

Flege, J.E., Munro, M. J., & MacKay, I. R. A. (1995b). Factors affecting strength of perceived foreign accent in a second language *Journal of the Acoustical Society of America, 97*(5), 3125-3134.

Flege, J.E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*(4), 467-491.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods Instruments and Computers, 35*(1), 116-124.

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 696-713.

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics, 36*(4), 649-663.

Frank, A. F., Tanenhaus, M. K., Aslin, R. N., Salverda, A. P. (2007). *Frequency, Neighborhood Density, and Phonological*

Similarity Effects in Picture Naming: An Artificial Lexicon Study. Proceedings of the 29nd Annual Conference of the Cognitive Science Society, 995-1000.

Gahl, S. (2008). Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language*, *84*(3), 474-496.

German, D. J., & Newman, R. S. (2004). The impact of lexical factors on children's word-finding errors. *Journal of Speech, Language, and Hearing Research, 47*(3), 624-636.

Gilhooly, K., & Gilhooly, M. (1979). Age-of-acquisition effects in lexical and episodic memory tasks. *Memory & Cognition, 7*(3), 214-223.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language, 28*(5), 501-518.

Goldrick, M., Folk, J. R., & Rapp, B. (2010). Mrs. Malaprop's neighborhood: Using word errors to reveal neighborhood structure. *Journal of Memory and Language, 62*(2), 113-134.

Gollan, T. H., & Acenas, L. 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.

Gollan, T. H., & Brown, A. S. (2006). From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology. General, 135*(3), 462-483.

Gollan, T. H., & Goldrick, M. (2012). Does bilingualism twist your tongue? *Cognition,125*(3), 491-497.

Gollan, T.H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition*, *4*(01), 63-83.

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston naming test performance. *Journal of the International Neuropsychological Society*, *13*(02), 197-208.

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology, 16*(4), 562-576.

Gollan, T. 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.

Gollan, T. 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.

Gollan, T. H., Montoya, R., & Bonanni, M. (2005). Proper names get stuck on bilingual and monolingual speakers' tip of the tongue equally often. *Neuropsychology, 19*(3), 278-287.

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency Drives Lexical Access in Reading but Not in Speaking: The Frequency-Lag Hypothesis. *Journal of Experimental Psychology: General, 140*(2),186-209.

Gordon, J. K. (2002). Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and Language, 82*(2), 113-145.

Gordon, J. K. (2011). The Aging Neighborhood: Phonological Density in Word Retrieval. *Procedia Social and Behavioral Sciences, 23*, 138–139.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29*(2), 228-244.

Grainger, J., & Seguí, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. *Attention, Perception, & Psychophysics, 47*(2), 191-198.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Seguí, J. (1989). On the role of competing word units in visual word recognition:

The neighborhood frequency effect. *Attention, Perception, & Psychophysics, 45*(3), 189-195.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Seguí, J. (1992). Neighborhood frequency effects and letter visibility in visual word recognition. *Attention, Perception, & Psychophysics, 51*(1), 49-56.

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language, 27*(2), 210-223.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*(2), 67-81.

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism* Harvard University Press.

Guion, S. G. (1995). Word frequency effects among homonyms. *Texas Linguistic Forum,* 35, 103-116.

Guion, S. G., Flege, J. E., & Loftin, J. D. (2000). The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics, 28*, 27-42.

Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics, 21*(02), 205-228.

Hall, C.J. (2002). The automatic cognate form assumption. Evidence for the parasitic model of vocabulary development. *IRAL 40*, 69-87.

Hanulová, J., Davidson, D. J., & Indefrey, P. (2011). Where does the delay in L2 picture naming come from? psycholinguistic and neurocognitive evidence on second language word production. *Language and Cognitive Processes, 26*(7), 902-934.

Harley, T. A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science, 8*(3), 191-219.

Harley, T. A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes, 8*(3), 291-309.

Harley, T. A., & Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology, 89*(1), 151-174.

Harrell, F. E. Jr., & with contributions from many other users. (2010). Hmisc: Harrell Miscellaneous. R package version 3.8-3. http://CRAN.R-project.org/package=Hmisc

Heiberger, R.M. (2009). HH: Statistical Analysis and Data Display: Heiberger and Holland. R package version 2.1-32. http://CRAN.R-project.org/package=HH

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition, 106*(1), 501-511.

Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A., & Van Der Laan, M. (2006). Survival Ensembles. *Biostatistics*, *7*(3), 355-373.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5*(1), 67-104.

Indefrey, P. (2006). A Meta-analysis of Hemodynamic Studies on First and Second Language Processing: Which Suggested Differences Can We Trust and What Do They Mean? *Language Learning*, *56*, 279–304.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277-288.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, *87*(1), B47-B57.

Jaeger, T. F., Furth, K., & Hilliard, C. (2012). Phonological overlap affects lexical selection during sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1378-1391.

Janssen N and Caramazza A (2011) Lexical selection in multi-word production. *Frontiers in Psychology*, 2:81.

Janssen, N., Schirm, W., Mahon, B. Z., & Caramazza, A. (2008). Semantic interference in a delayed naming task: Evidence for the response exclusion hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 34*(1), 249-256.

Jescheniak, J.D., & Levelt, W.J.M. (1994).Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 2*0(4), 824–843.

Jones, Ō P., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., Lee, H. L., Ramsden, S., Gazarian, K., Prejawa, S., Seghier, M. L., & Price, C. J. (2012). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cerebral Cortex, 22*(4), 892-902.

Kawamoto, A. H., Kello, C. T., Higareda, I., & Vu, J. V. Q. (1999). Parallel processing and initial phoneme criterion in naming words: Evidence from frequency effects on onset and rime duration, *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(2), 362-381.

Kawamoto, A. H., Kello, C. T., Jones, R., & Bame, K. (1998). Initial phoneme versus whole-word criterion to initiate pronunciation: Evidence based on response latency and initial phoneme duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(4), 862-885.

Kello, C. T. (2004). Control over the time course of cognition in the tempo-naming task. *Journal of Experimental Psychology, 30*(5), 942-955.

Kello, C. T., Plaut, D. C., & MacWhinney, B. (2000). The task dependence of staged versus cascaded processing: An empirical and computational study of Stroop interference in speech production. *Journal of Experimental Psychology: General, 129*(3), 340-360.

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology, 25*(4), 463-492.

Kohnert, K. (2004). Cognitive and cognate-based treatments for bilingual aphasia: A case study. *Brain and Language, 91*(3), 294-302.

Kohnert, K. J., Hernandez, A. E., & Bates, E. (1998). Bilingual performance on the Boston naming test: Preliminary norms in Spanish and English. *Brain and Language, 65*(3), 422-440.

Kondrak, G. (2000). A new algorithm for the alignment of phonetic sequences. Proceedings of the 1st North American Chapter of the Association for Computational Linguistics Conference, 288-295.

Kremin, H., Hamerel, M., Dordain, M., De Wilde, M., & Perrier, D. (2000). Age of acquisition and name agreement as predictors of mean response latencies in picture naming of French adults. *Brain and Cognition, 43*(1-3), 286-291.

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*, 149-149.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences, 97*(22), 11850-11857.

La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory & Cognition, 16*(5), 400-410.

Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. *Journal of Experimental Psychology, 99*(2), 199-208.

Lachman, R., Shaffer, J. P., & Hennrikus, D. (1974). Language and cognition: Effects of stimulus codability, name-word frequency, and age of acquisition on lexical reaction time. *Journal of Verbal Learning & Verbal Behavior, 13*(6), 613-625.

Landauer, T. K., & Streeter, L. A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior, 12*(2), 119-131.

Lavoie, L. (2002). Some influences on the realization of for and four in American English. *Journal of the International Phonetic Association, 32*(2), 175-202.

Lemhöfer, K., Dijkstra, T., & Michel, M. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes, 19*(5), 585-611.

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 12-31.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Levelt, W. J. M. (1989). *Speaking : From intention to articulation*. Cambridge, Mass.: MIT Press.

Levelt, W. J. M. (2002). Picture naming and word frequency: Comments on Alario, Costa and Caramazza. *Language and Cognitive Processes, 17*(6), 663-671.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98*(1), 122-142.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences, 22*(1), 1-38.

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.

Luce, P. A. (1986). Neighborhoods of words in the mental lexicon. Ph.D. dissertation, Indiana University.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*(1), 1-36.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*(1), 29-41.

Mack, M. (1989). Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Attention, Perception, & Psychophysics, 46*(2), 187-200.

Mackay, I. R. A., & Flege, J. E. (2004). Effects of the age of second language learning on the duration of first and second language sentences: The role of suppression. *Applied Psycholinguistics, 25*(03), 373-396.

MacKay, I. R., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *The Journal of the Acoustical Society of America, 110*(1), 516-528.

Mägiste, E. (1979). The competing language systems of the multilingual: A developmental study of decoding and encoding processes. *Journal of Verbal Learning and Verbal Behavior, 18*(1), 79-89.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33*(3), 503-535.

Marian, V., Blumenfeld, H. K. & Boukrina, O. V. (2008). Sensitivity to phonological similarity within and across languages. *Journal of Psycholinguistic Research*, *37*(3), 141–170.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29-63.

Martin, N., Gagnon, D. A., Schwartz, M. F., Dell, G. S., & Saffran, E. M. (1996). Phonological facilitation of semantic errors in normal and aphasic speakers. *Language and Cognitive Processes*, *11*(3), 257-282.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1-86.

Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture-word interference experiments. *Journal of Memory and Language*, *35*(4), 477-496.

Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, *20*(6), 715-726.

Miceli, G., Capasso, R., & Caramazza, A. (1999). Sublexical conversion procedures and the interaction of phonological and orthographic forms. *Cognitive Neuropsychology*, *16*(6), 557–572.

Middleton, E. L., & Schwartz, M. F. (2010). Density pervades: An analysis of phonological neighbourhood density effects in aphasic speakers with different types of naming impairment. *Cognitive Neuropsychology, 27*(5), 401-427.

Mirman, D., Kittredge, A. K., & Dell, G. S. (2010). Effects of near and distant phonological neighbors on picture naming. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, 1447-1452.

Mitterer, H., & Ernestus, M. (2008). The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition, 109*(1), 168-173.

Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28*(3), 555-563.

Mortensen, L., Meyer, A. S., & Humphreys, G. W. (2008). Speech planning during multiple-object naming: Effects of ageing. *The Quarterly Journal of Experimental Psychology, 61*(8), 1217-1238.

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. Developmental Science, 10(6), 719-726.

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*(3), 289-306.

Munro, M. J., Flege, J. E., & MacKay, I. R. A. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics, 17*, 313-334.

Munson, B., Swenson, C. L., & Manthei, S. C. (2005). Lexical and phonological organization in children: Evidence from repetition tasks. *Journal of Speech, Language, and Hearing Research*, *48*(1), 108-124.

Navarrete, E., Basagni, B., Alario, F. X., & Costa, A. (2006). Does word frequency affect lexical selection in speech production? *Quarterly Journal of Experimental Psychology, 59*(10), 1681-1690.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE™//A lexical database for contemporary French: LEXIQUE™. *L'Année Psychologique, 101*(3), 447-462.

Newman, R. S., & Bernstein Ratner, N. (2007). The role of selected lexical factors on confrontation naming accuracy, speed, and fluency in adults who do and do not stutter. *Journal of Speech, Language, and Hearing Research, 50*(1), 196-213.

Newman, R. S., & German, D. J. (2002). Effects of lexical factors on lexical access among typical language-learning children and children with word-finding difficulties. *Language and Speech, 45*(3), 285-317.

Newman, R. S., & German, D. J. (2005). Life span effects of lexical factors on oral naming. *Language and Speech, 48*(2), 123-156.

Nissen, S. L., Dromey, C., & Wheeler, C. (2007). First and second language tongue movements in Spanish and Korean bilingual speakers. *Phonetica, 64*(4), 201-216.

Nooteboom, S. G. (2005). Listening to oneself: Monitoring speech production. In R. Hartsuiker, R. Bastiaanse, A. Postma, & F. Wijnen (Eds.). *Phonological encoding and monitoring in normal and pathological speech.* Hove, UK: Psychology Press.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*(3), 189-234.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17*(4), 273-281.

Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, *28*(2), 191-230.

O'Seaghdha, P. G., & Marin, J. W. (2000). Phonological competition and cooperation in form-related priming: Sequential and nonsequential processes in word production.

Journal of Experimental Psychology: Human Perception and Performance, 26(1), 57-73.

Paivio, A., & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology/Revue Canadienne De Psychologie, 34*(4), 388.

Pérez, M. Á. (2007). Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. *Quarterly Journal of Experimental Psychology, 60*(1), 32-42.

Pérez, M. Á., & Navalón, C. (2003). Normas españolas de 290 nuevos dibujos: Acuerdo en la denominación, concordancia de la imagen, familiaridad, complejidad visual y variabilidad de la imagen. *Psicológica,* (002), 215-241.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(3), 539.

Pind, J., & Tryggvadottir, H. B. (2002). Determinants of picture naming times in Icelandic. *Scandinavian Journal of Psychology, 43*(3), 221-226.

Pintner, R. (1932). The Influence of Language Background on Intelligence Tests. *Journal of Social Psychology, 3,* 325-240.

Piske, T., Flege, J. E., MacKay, I. R. A., & Meador, D. (2002). The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica*, *59*(1), 49-71.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*(2), 191-215.

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *22*(3), 415-422.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, *15*(1), 36-57.

Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859-862.

R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/.

R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*(3), 460-499.

Rapp, D.N., & Samuel, A.G. (2002). A reason to rhyme: Phonological and semantic influences on lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 564-571.

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston naming test. *Aphasiology, 16*(4-6), 635-645.

Rodriguez-Gonzalez, E. (2012). Processing of Spanish preterite regular and irregular verbs: The role of neighborhood density. *The Spanish Journal of Psychology, 15*(1), 35-47.

Roelofs, A. (2003). Shared phonological encoding processes and representations of languages in bilingual speakers. *Language and Cognitive Processes, 18*(2), 175-204.

Rosselli, M., & Ardila, A. (2002). A cross-linguistic comparison of verbal fluency tests. *International Journal of Neuroscience, 112*(6), 759-776.

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., et al. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology, 7*(1), 17-24.

Runnqvist, E., FitzPatrick, I., Strijkers, K., & Costa, A. (2012). In T. K. Bhatia & W. C. Ritchie (Eds.), *The Handbook of Bilingualism and Multilingualism, Second Edition* (pp. 244-265). John Wiley & Sons, Ltd.

Runnqvist, E., Strijkers, K., Alario, F.-X., & Costa, A. (2012). Cumulative semantic interference is blind to language: Implications for models of bilingual speech production. *Journal of Memory and Language, 66*(4), 850-869.

Runnqvist, E., Strijkers, K., Sadat, J., & Costa, A. (2011). On the temporal and functional origin of L2 disadvantages in speech production: A critical review. *Frontiers in Psychology, 2,* 379.

Saer, D. J. (1923). The Effect of Bilingualism on Intelligence. *British Journal of Psychology*, 25-38.

Sadat, J., Martin, C. D., Alario, F. X., & Costa, A. (2012). Characterizing the bilingual disadvantage in noun phrase production. *Journal of Psycholinguistic Research, 41*(3), 159-179.

Sadat, J., Martin, C. D., Costa, A., & Alario, F. X. (submitted). Reconciling phonological neighborhood effects in speech production through single trial analysis. *Journal of Memory and Language.*

Sampat, P. (2001). Last words: The dying of languages. *World Watch Magazine*, *14*(3), 34-40.

Sánchez-Casas, R., & García-Albea, J. E. (2005). The representation of cognate and noncognate words in bilingual memory In J. F. Kroll & A. M. B. De Groot (Eds.). *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 226-250). New York: Oxford University Press.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian-Portuguese and English. *Journal of Phonetics, 25*(4), 421-436.

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition, 13*, 231-252.

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. Bilingualism: Language and Cognition, 15(1), 157-166.

Schriefers, H., & Teruel, E. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology, 11*(1), 17-50.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29*(1), 86-102.

Schwartz, A. I., & Kroll, J. F. (2006). Language processing in bilingual speakers. In Matthew J. Traxler, & Morton A.

Gernsbacher (Eds.), *Handbook of psycholinguistics - second edition* (pp. 967-999). London: Academic Press.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*(4), 876-900.

Sevald, C. A., & Dell, G. S. (1994). The sequential cuing effect in speech production. *Cognition*, *53*(2), 91-127.

Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, *119*(2), 159–187.

Simmonds, A. J., Wise, R. J. S., Dhanjal, N. S., & Leech, R. (2011). A comparison of sensory-motor activity during speech in first and second languages. *Journal of Neurophysiology, 106*(1), 470-478.

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.

Slevc, L. R., & Ferreira, V. S. (2006). Halting in single word production: A test of the perceptual loop theory of speech monitoring. *Journal of Memory and Language*, *54*(4), 515-540.

Smith, M. E. (1939). Some Light on the Problem of Bilingualism as Found from a Study of the Progress in Mastery of English

Among Pre-school Children of Non-American Ancestry in Hawaii. *Genetic Psychology Monolographs*, *21*, 221-284.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174-215.

Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods*, *28*(4), 516-536.

Spieler, D. H., & Balota, D. A. (2000). Factors influencing word naming in younger and older adults. *Psychology and Aging*, *15*(2), 225.

Stemberger, J. P. (2004). Neighbourhood effects on error rates in speech production. *Brain and Language*, *90*(1-3), 413-422.

Storkel, H. L. (2004). Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. *Journal of Speech, Language, and Hearing Research*, *47*(6), 1454-1468.

Storkel, H. L., & Morrisette, M. L. (2002). The lexicon and phonology: Interactions in language acquisition. *Language, Speech, and Hearing Services in Schools*, *33*(1), 24.

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex, 20*(4), 912-928.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., (2008). Conditional Variable Importance for Random Forests. *BMC Bioinformatics*, *9*(307).

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T., (2007). Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. *BMC Bioinformatics*, *8*(25).

Sullivan, M. P., & Riffel, B. (1999). The nature of phonological encoding during spoken word retrieval. *Language and Cognitive Processes*, *14*(1), 15-45.

Székely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of picture naming. *Center for Research in Language Newsletter*, *12*(2). La Jolla: University of California, San Diego.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*(2), 247-250.

Tabak, W., Schreuder, R., & Baayen, R. H. (2010). Producing inflected verbs: A picture naming study. *The Mental Lexicon*. *5*(1), 22-46.

Taler, V., Aaron, G. P., Steinmetz, L. G., & Pisoni, D. B. (2010). Lexical neighborhood density effects on spoken word recognition and production in healthy aging. *Journal of Gerontology: Psychological Sciences*, 65B(5), 551–560.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? cognate effects in a sentence context. *Psychological Science, 20*(8), 923-927.

Van Hell, J. G., & De Groot, A. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1*(03), 193-211.

Venables, W. N. & Ripley, B. D. (2002). *Modern Applied Statistics with S*. Fourth Edition. Springer, New York.

Vitevitch, M. S. (1997). The neighborhood characteristics of malapropisms. *Language and Speech, 40*(3), 211-228.

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28*(4), 735-747.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science, 9*(4), 325-329.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*(3), 374-408.

Vitevitch, M. S., & Rodríguez, E. (2005). Neighborhood density effects in spoken word recognition in Spanish. *Journal of Multilingual Communication Disorders, 3*(1), 64-73.

Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory & Cognition*, *31*(4), 491-504.

Vitevitch, M. S., & Stamer, M. K. (2006). The curious case of competition in Spanish speech production. *Language and Cognitive Processes, 21*(6), 760-770.

Vitevitch, M. S., & Stamer, M. K. (2009). The influence of neighborhood density (and neighborhood frequency) in Spanish speech production: A follow-up report. *University of Kansas, Spoken Language Labor*atory *Technical Report, 1*, 1–6.

Vitevitch, M. S., Armbruster, J., & Chu, S. (2004). Sublexical and lexical representations in speech production: Effects of phonotactic probability and onset density. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 30*(2), 514-529.

Vitkovitch, M., & Tyrrell, L. (1995). Sources of disagreement in object naming. *The Quarterly Journal of Experimental Psychology, 48*(4), 822-848.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(2), 423-440.

Wechsler, D. (1997). Wechsler Adult Intelligence Scale - Third Edition. New York: The Psychological Corporation.

Wheeldon, L. (2003). Inhibitory form priming of spoken word production. *Language and Cognitive Processes, 18*(1), 81-109.

Wright, C. E. (1979). Duration differences between rare and common words and their implications for the interpretation of word frequency effects. *Memory & Cognition, 7*(6), 411-419.

Yaniv, I., Meyer, D. E., Gordon, P. C., Huff, C.A., & Sevald, C.A. (1990). Vowel similarity, connectionist models, and syllable structure in motor programming of speech. *Journal of Memory and Language, 29*(1), 1-26.

Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition, 3*(02), 131-149.

Ziegler, J. C., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language*, *48*(4), 779–793.

Ziegler, J.C., & Muneaux, M. (2007). Orthographic facilitation and phonological inhibition in spoken word recognition: a developmental study. *Psychonomic Bulletin & Review, 14(*1), 75–80.

Ziegler, J.C., & Perry, C. (1998). No more problems in Coltheart's neighborhood: Resolving neighborhood conflicts in the lexical decision task. *Cognition*, *68*(2), 53-62.

Appendix A

Beta coefficients, standard errors (SE), t- and p-values together with the variance inflation factor (VIF) for each of the predictors in the final model with centered predictors.

Predictor	raw β	SE β	t-value	p-value	VIF
Intercept	-1.14	0.03	-40.31	< 0.001	
Session	-0.13	0.01	-13.30	< 0.001	1.3
Trial order	0.01	<0.01	4.60	< 0.001	1.3
Name agreement	-0.11	0.01	-19.69	< 0.001	1.1
Age of acquisition	0.04	0.01	5.20	< 0.001	1.1
Lexical Frequency	-0.02	0.01	-3.58	< 0.001	1.1
Neighborhood Frequency	0.01	0.01	1.77	0.041	1.1
PhND	0.02	0.01	2.60	0.006	1.0

Note: PhND = phonological neighborhood density. The same residualization procedures as described for the main analysis had to be applied to the centered variables to reduce collinearity.

Appendix B

Description of the bilingual community in Catalonia:

In Catalonia, Catalan and Spanish are both official languages. In many families, both languages are spoken. In kindergarten (ages 4–5), special Catalan programs are offered to children from monolingual Spanish families. The current education system requires that at the end of the primary school (year 11/12), children are able to read, write, speak, and understand both Catalan and Spanish. In high school, some classes are taught in Catalan and others in Spanish. At the university, classes and tests can be in either language – quite often half of the test is in Catalan, other half in Spanish. Radio and television programs broadcast in Catalan and in Spanish. Furthermore, newspapers contain articles written in Catalan and Spanish. All the bilingual participants passed the Catalan–Spanish language proficiency exam that is required for enrollment at the university. In order to pass this exam students must be highly proficient in all aspects of the two languages (vocabulary, grammar, etc.).

Appendix C

Excerpt of the socioeconomic questionnaire:

1. Please specify the maximum studies that you have completed.
2. Please specify your current occupation.
3. Please specify the maximum studies that your father has completed.
4. Please specify the maximum studies that your mother has completed.
5. Please indicate if your family receives some kind of financial support.
6. Please indicate how many places of residence your family possesses.
7. Please indicate whether you have lived most of your lifetime in a rural or urban environment.
8. Please indicate whether your current place of residence is rented or owned.
9. Please indicate the average monthly income of your family and how many members of the family are currently working?
10. Please indicate with how many people you live at your current place of residence.
11. Please indicate what kind of amenities and services your current place of residence includes.

Appendix D

Picture material used in the experiments of study 3:

HF Picture names	English	LF Picture names	English
+ AVIÓN (Avió)	Airplane	+ BOMBILLA (Bombeta)	Light bulb
+ BALCÓN (Balcó)	Balcony	+ ESCOBA (Escombra)	Broom
+ BOLSA (Bossa)	Bag	+ GLOBO (Globus)	Balloon
+ ESCALERA (Escala)	Ladder	+ PATÍN (Patí)	Roller skate
+ FLOR (Flor)	Flower	+ PELUCA (Perruca)	Wig
+ PELOTA (Pilota)	Ball	+ PINCEL (Pinzell)	Paintbrush
+ PUERTA (Porta)	Door	+ REGADERA (Regadora)	Watering can
+ RELOJ (Rellotge)	Watch	+ SIERRA (Serra)	Saw
+ TELÉFONO (Telèfon)	Telephone	+ TABURETE (Tamboret)	Stool
+ VESTIDO (Vestit)	Dress	+ VIOLÍN (Violí)	Violin
CUCHILLO (Ganivet)	Knife	CALCETÍN (Mitjó)	Sock
ESPEJO (Mirall)	Mirror	CEPILLO (Raspall)	Brush
HILO (Fil)	Spool of thread	COLCHÓN (Matalas)	Matress
HOJA (Fulla)	Leaf	COLUMPIO (Gronxador)	Swing
HUEVO (Ou)	Egg	CORCHO (Suro)	Cork
MANCHA (Taca)	Stain	HACHA (Destral)	Axe
MESA (Taula)	Table	HOZ (Falç)	Sickle
SILLA (Cadira)	Chair	HUCHA (Guardiola)	Piggy bank
SOMBRERO (Barret)	Hat	MULETA (Crossa)	Crutch
VENTANA (Finestra)	Window	PEONZA (Baldufa)	Top

Note: HF = high frequency; LF = low frequency; + = cognate words. Catalan translations are given in brackets.

Appendix E

Results on articulatory durations in bare noun production:

Beta coefficients and standard errors (SE) of predictors in two models estimating speaker group effects. In model 1, the basic bilingual cost is estimated. In model 2, contributions of the effect of translation overlap referenced at its highest value to the speaker group effect were estimated. A square-root transformation was applied to the articulatory durations to approximate a normal distribution.

Predictor	**Model 1**			**Model 2**		
	raw β	**SE β**	**sign**	**raw β**	**SE β**	**sign**
Intercept	14.98	0.55	***	13.64	0.72	***
Phoneme length	0.96	0.07	***	1.02	0.07	***
Repetition	-0.07	0.01	***	-0.07	0.01	***
Lexical Frequency	-0.26	0.19	*	-0.23	0.18	*
L1 bilinguals	0.48	0.35	*	0.46	0.37	*
L2 bilinguals	0.79	0.36	**	1.04	0.39	**
Translation overlap				1.67	0.61	**
L1 bilinguals * Translation overlap				0.03	0.17	
L2 bilinguals * Translation overlap				-0.43	0.27	*

Note: sig=significance level (*** is $p<0.001$, ** is $p<0.05$, * is $p<0.11$).

Results on articulatory durations in noun phrase production:

Beta coefficients and standard errors (SE) of predictors in two models estimating speaker group effects. In model 1, the basic bilingual cost is estimated. In model 2, contributions of the effect of translation overlap referenced at its highest value to the speaker group effect were estimated. A logarithmic transformation was applied to the articulatory durations to approximate a normal distribution.

Predictor	Model 1			Model 2		
	raw β	SE β	sign	raw β	SE β	sign
Intercept	6.53	0.03	***	6.53	0.03	***
Phoneme length	0.03	<0.01	***	0.03	<0.01	***
Repetition	-0.01	<0.01	***	-0.01	<0.01	***
Lexical Frequency	-0.03	0.01	***	-0.03	0.01	***
L1 bilinguals	0.07	0.03	**	0.07	0.03	**
L2 bilinguals	0.08	0.03	***	1.00	0.03	***
Translation overlap				-0.01	0.02	
L1 bilinguals * Translation overlap				-0.01	0.01	
L2 bilinguals * Translation overlap				-0.04	0.01	***

Note: sig=significance level (*** is p<0.001, ** is p<0.05, * is p<0.10).

Résumé substantiel en français

Le coût bilingue dans la production de la parole: des études sur les processus phonologiques et articulatoires

Résumé

L'objectif principal de cette thèse est d'étudier les conséquences du bilinguisme sur la production de la parole. Des recherches antérieures ont montré que les locuteurs bilingues montrent un coût par rapport aux monolingues dans différentes expériences linguistiques. Nous avons étudié les origines du coût bilingue en explorant l'influence de variables spécifiques comme la similitude phonologique. En outre, nous avons examiné la portée du coût bilingue en évaluant des durées articulatoires de la parole et la production de groupes nominaux. Nous rapportons que l'augmentation de la similarité phonologique entre les mots d'une même langue ralentit la parole, alors que l'augmentation de la similarité phonologique entre les langues aide les bilingues à surmonter le coût bilingue. En outre, nos résultats ont montré que le coût bilingue se généralise aux phases articulatoires et à la production de groupes nominaux. Cette thèse fournit de nouvelles données sur le traitement de la parole pendant les stades phonologiques et articulatoires chez les locuteurs mono- et bilingues, et étend notre connaissance sur le coût bilingue en production de parole.

Le contexte de la recherche actuelle

Le bilinguisme est devenu un phénomène de plus en plus courant dans notre monde globalisé. Aujourd'hui, plus de 60 pour cent de la population mondiale grandit bilingue et il y a plus de locuteurs utilisant l'anglais comme seconde langue que comme première langue (Sampat, 2001). Ainsi, la recherche sur les conséquences du bilinguisme a connu une croissance considérable au cours des deux dernières décennies. Bien que les premières études sur les bilingues, datant de la première moitié du siècle dernier, ont presque uniformément rapporté des conséquences négatives sur les capacités cognitives (par exemple, Pintner, 1932; Saer, 1923; Smith, 1939), les travaux conduits ces dernières décennies ont fourni une image plus précise et différenciée des implications du bilinguisme. De nombreuses études récentes ont rapporté des effets bénéfiques, que ce soit dans les domaines non linguistiques ou méta-linguistiques, de la maîtrise de deux langues. Cependant, quelques études ont également constaté des conséquences négatives sur la performance linguistique des bilingues, désignant ce phénomène comme le coût bilingue ou le désavantage bilingue. Dans cette thèse, nous allons nous référer au coût bilingue lorsque la performance linguistique diffère entre locuteurs mono- et bilingues. Cependant, il est important de souligner que ce coût réfère à tous les comportements linguistiques différant entre ces deux groupes de locuteurs (mono- et bilingues), plutôt qu'à un vrai inconvénient pour la communication. Malgré les désavantages linguistiques associés au bilinguisme, il est évident que la connaissance de plusieurs langues représente un véritable avantage dans la capacité d'interagir avec d'autres personnes et de s'immerger dans des cultures différentes. Néanmoins, ces différences de performance nécessitent une enquête empirique détaillée pour fournir une description appropriée de la production de parole chez les locuteurs bilingues et de faire progresser nos connaissances sur ce phénomène omniprésent.

Dans la littérature actuelle, les bilingues sont définis comme des personnes qui utilisent deux langues dans leur vie quotidienne de façon active (compréhension et production) ou passive (seulement comprendre, par exemple, Grosjean, 1982). Cette définition détend la notion populaire d'un locuteur bilingue qui parle avec un niveau presque natif dans les deux langues. Dans cette thèse, cependant, nous allons utiliser cette dernière définition de bilingues étant donné la communauté à l'étude. En Catalogne, le catalan et l'espagnol sont les deux langues officielles et l'éducation dans les deux langues est proposée aux enfants à partir de la maternelle. Le système éducatif actuel exige que, à la fin de l'école primaire, les enfants soient capables de lire, écrire, parler et comprendre le catalan et l'espagnol. En outre, dans beaucoup de familles les deux langues sont parlées. Ainsi, grandir en Catalogne signifie généralement que les deux langues sont présentes très tôt dans la vie quotidienne, et que les enfants bilingues ont des niveaux élevés de compétence dans les deux langues. La plupart des études sur les effets du bilinguisme portent sur une seconde langue apprise plus tardivement. Cependant, peu d’attention a été portée sur la façon dont le bilinguisme influence la production de la parole lorsque les deux langues sont acquises presque simultanément et à un âge précoce. Cette thèse se concentre précisément sur la performance de la parole chez les locuteurs bilingues précoces et très compétents.

Une caractéristique particulière des personnes bilingues espagnol-catalan est qu'ils parlent deux langues romanes qui partagent de nombreuses propriétés fondamentales. Environ 70% du vocabulaire peut être considéré comme ayant des traductions proches phonétiquement. En outre, la grammaire et la syntaxe sont étroitement liées dans les deux langues. Cependant, on trouve des différences significatives entre le catalan et l'espagnol dans leurs répertoires phonologiques. L’espagnol a moins de voyelles que le catalan et certaines consonnes existent seulement en espagnol, tandis que d'autres sont spécifiques au catalan. Par conséquent, ces deux langues fournissent un modèle idéal pour étudier le rôle des processus phonologiques dans la production de langage des bilingues. Un aspect central de cette thèse est l'impact que peut avoir l’interaction des systèmes phonologiques d’un bilingue sur sa capacité à maîtriser deux langues.

De travail de recherche vise à étudier deux aspects importants du coût bilingue. Tout d'abord, nous allons nous intéresser aux origines du coût bilingue précédemment établi dans la production de mots uniques et notamment si ces origines peuvent être liées à des variables particulières, comme la similarité phonologique. Ensuite, nous allons étudier la portée du coût bilingue en évaluant les durées articulatoires et dans des contextes de parole plus larges que la production de mots uniques. L'objectif de cette thèse est de fournir une meilleure compréhension des processus phonologiques et articulatoires impliqués dans la production de parole de mono- et bilingues.

Partie expérimentale

Plusieurs études ont montré que les propriétés phonologiques du mot à prononcer ont un impact sur la récupération du mot à la fois chez les mono- et bilingues. Ainsi, un aspect important dans la compréhension de difficultés lors de la production de parole est de considérer la mesure dans laquelle un mot donné présente plus ou moins de similarité phonologique avec d'autres mots de la même langue (pour les monolingues) et d'une autre langue (pour les bilingues). Ainsi des variables telles que la densité du voisinage phonologique et le statut de cognat sont devenues des questions centrales. Dans des études précédentes, il a été montré que ces deux variables ont un effet facilitateur sur la récupération de mots de similarité croissante. Étonnamment, les origines de l'effet facilitateur des cognats chez les locuteurs bilingues ont été étroitement liées à l'effet facilitateur de la densité du voisinage phonologique (par exemple, Costa, Santesteban, & Caño, 2005; Runnqvist, FitzPatrick, Strijkers, & Costa 2012). Étant donné que les effets de cognats et de la densité du voisinage phonologique sont tous deux interprétés comme reposant sur le principe d'activation interactive en production de parole, il semble évident de proposer un mécanisme sous-jacent commun.

Cependant, des études récentes sur l'effet de la densité du voisinage phonologique en production de parole ont donné des résultats plus contradictoires: certaines études ont rapporté que les latences de dénomination sont plus longues pour les mots issus de voisinage phonologiques denses (par exemple, Arnold, Conture, et Ohde, 2005; Taler, Aaron, Steinmetz, & Pisoni, 2010; Vitevitch & Stamer, 2006, 2009), tandis que d'autres ont rapporté des latences de dénomination plus rapides (par exemple, Baus, Costa, & Carreiras, 2008; Perez, 2007). Plus particulièrement, des études font état de résultats contradictoires dans le cas de la production de parole en espagnol (Baus et al., 2008; Perez, 2007; Vitevitch & Stamer, 2006, 2009), la langue utilisée dans ce travail de recherche. Avant d'aborder les effets des propriétés phonologiques dans la production de langage chez des locuteurs bilingues, nous avons d'abord étudié comment la similarité phonologique entre les mots d'une même langue influence la production de parole. Nous l'avons fait en évaluant la performance de dénomination des monolingues dans une étude à grande échelle de dénomination de dessins en espagnol. Dans cette étude, nous avons confronté nos résultats à la littérature très contradictoire sur l'effet de la densité du voisinage phonologique en production de parole. Ainsi, nous avons proposé que les différences d'effets des propriétés phonologiques sur la récupération du mot obtenues chez les locuteurs monolingues dépendent des différentes tâches utilisées.

La deuxième étude de cette thèse porte sur l'influence de la similarité phonologique entre les langues (ie, le statut cognat) sur la production de parole des personnes bilingues. Les études chez les locuteurs bilingues ont toujours montré une meilleure performance dans le cas de traductions phonologiquement similaires (i.e., cognats) par rapport aux traductions non similaires (i.e., non-cognats). Ainsi, il semble intéressant d'étudier l'effet des propriétés phonologiques dans le contexte du coût bilingue et d'expliquer comment ces deux phénomènes peuvent être interconnectés. Cette approche a permis de clarifier le rôle jusque là plutôt négligé du traitement phonologique pour le coût bilingue et d'approfondir l'origine de l'effet de similarité phonologique dans la production de parole des locuteurs bilingues.

La troisième étude de cette thèse examine la portée du coût bilingue et vise à fournir de nouvelles données basées sur les durées articulatoires et la production de groupes nominaux. Plusieurs études ont déjà montré que les durées articulatoires des bilingues tardifs dans leur seconde langue sont plus longues que celles des locuteurs natifs. Ici, nous avons vérifié si le coût bilingue était également évident chez des locuteurs bilingues précoces et très compétents qui utilisent leur langue maternelle. En outre, nous avons également étudié comment le coût bilingue peut se généraliser de la production de mots isolés à la production de groupes nominaux. Cette approche permettra de fournir de nouvelles données sur les implications d'un coût bilingue sur les durées articulatoires et la production de groupes nominaux.

Résumé des résultats

L'objectif de cette thèse est de fournir une description plus détaillée du coût bilingue et de ses processus sous-jacents. Plus précisément, nous nous sommes intéressés aux influences des niveaux phonologique et articulatoire de la production de parole sur le coût bilingue. Cette approche a été motivée par l'observation de fortes influences entre les langues à des niveaux tardifs de traitement de la parole (voir Indefrey, 2006; Hanulová, Davidson, & Indefrey, 2011). Nos prédictions étaient que les processus phonologiques et articulatoires devraient fournir des informations importantes sur l'origine et l'ampleur des coûts bilingues. Ainsi, nous nous attendions à ce que le coût bilingue varie en fonction de la similarité phonologique et s'étende à la fois sur toute une gamme de mesures de la performance et dans différents contextes de parole. La partie expérimentale de cette thèse contient trois études abordant ces questions. Nous avons utilisé des tâches de dénomination monolingue et bilingue et mesuré les latences de dénomination (dans les deux premières études), ou les latences de dénomination ainsi que les durées articulatoires (dans la dernière étude). Une attention particulière a été accordée aux effets des variables d'indexation de similarité phonologique (dans une même langue et entre les langues) et aux durées articulatoires des énoncés. Ainsi, les résultats de ce travail de recherche peuvent se résumer de la façon suivante:

- La première étude a montré que l'augmentation de similarité phonologique entre les mots d’une même langue ralentit la production de la parole. Ce résultat contraste avec l'effet facilitateur de la similarité phonologique que l’on retrouve entre les langues chez les bilingues. Ainsi, l'effet facilitateur de similarité phonologique entre les mots d'une même langue ne peut pas être transféré à l’effet de similarité phonologique entre les langues.

- La deuxième étude a montré que la similarité phonologique entre les langues a une influence majeure sur le coût bilingue. En effet, nous avons observé que le coût bilingue diminuait lorsque la similarité phonologique entre les langues augmentait. Ainsi, l'effet de similarité phonologique aiderait les traducteurs bilingues à surmonter le coût bilingue.

- La troisième étude a montré que le coût bilingue se retrouve dans les durées articulatoires des locuteurs bilingues s'exprimant dans leur langue maternelle et peut se généraliser à la production de groupes nominaux. Nous avons conclu que le coût bilingue est présent dans une grande variété de mesures de la performance et de contextes de parole et représente donc un facteur non négligeable pour la production de parole.