Bilingual word production

Jana Klaus & Herbert Schriefers

Radboud University

1.0 Introduction

Language production is a complex cognitive process which requires the selection of an appropriate lexical entry and its phonological properties in order to articulate a to-be-expressed concept (e.g., Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992). Despite this enormous computational effort, humans produce words, ideally conjoined to conceptually, syntactically, grammatically, and phonologically correct sentences, in a fairly automatic way, at a high speed, and without unreasonable amounts of errors. On top of this, about half of the world is bilingual (Bialystok, Craik, & Luk, 2012; Grosjean, 2010), meaning that next to their native language, they use one (or more) languages on a more or less regular basis. Intuitively, this increases the cognitive demands imposed during speaking exponentially. However, bi- and multilinguals master this task remarkably well: Language intrusions happen rarely (Poulisse, 2000; Poulisse & Bongaerts, 1994), and bi- and multilinguals are able to switch between languages without much effort. Nevertheless, how two languages work together on the behavioral, neurophysiological, and hemodynamic level has not been fully understood.

Before turning to the neuroscientific aspect of how two languages are represented in a speaker's brain, we should point out that even on a behavioural level, it is not clear to what extent a speaker's two languages operate separately from each other (or not). While it is assumed that the conceptual system is shared by both languages (e.g., Costa, 2004; De Bot, 1992), there is still debate about whether lexical and phonological features work in a distinct or combined fashion. In other words, the question of whether lexical selection and subsequent processes are target language-specific (i.e., both languages are activated, but the non-target language does not compete for selection) or target language non-specific (i.e., both languages compete for selection) in bilinguals has not been satisfactorily settled yet. The majority of behavioural studies has provided evidence for the hypothesis that lexical entries from both languages compete for selection, a problem solved by active inhibition of the language which is not to be used in the present context (Green, 1998; Hermans, Bongaerts, De Bot, & Schreuder, 1998; Jacobs, Fricke, & Kroll, 2016; Kroll, Bobb, Misra, & Guo, 2008; Misra, Guo, Bobb, & Kroll, 2012; Spalek, Hoshino, Wu, Damian, & Thierry, 2014). However, other studies also identified situations in which no selection for competition occurred, arguing for target language-specific lexical selection (Colomé, 2001; Costa & Caramazza, 1999; Costa, Miozzo, & Caramazza, 1999). More recently, lexical selection has been discussed as a dynamic process which can act both target language-specific and non-specific (e.g., Boukadi, Davies, & Wilson, 2015; Costa et al., 2003; Hermans, Ormel, van Besselaar, & van Hell, 2011; Kroll, Bobb, & Wodniecka, 2006). The constant improvement of methodologies to investigate the neural substrates underlying successful bilingual speech production has helped to provide additional insights into its mechanisms. In this chapter, we will review studies that investigated the functional and neuroanatomical representations of first and second languages during single-word production in healthy bilingual speakers. We will emphasize three related aspects: (1) are the languages of bilinguals underpinned by common or distinct brain regions during production, (2) what are the neurophysiological signatures of timing aspects of bilingual word production, and (3) to what extent are potential dissociations modulated by moderating factors such as age of acquisition, proficiency, and immersion of the L2.

2.0 Bilingual word production: common or distinct neural signatures?

In the last two decades, a substantial amount of studies investigated to what extent a speaker's first and second languages are represented by the same neural substrates. The majority of these studies have provided evidence for the single-network hypothesis (Abutalebi & Green, 2007*)*, which has been continuously updated in recent years to account for new findings (Abutalebi & Green, 2016; Green & Abutalebi, 2013). At its core, this framework assumes largely shared neural representations of L1 and L2. Specifically, it purports that both cortical and evolutionarily older subcortical structures are involved in language control and lexical selection. Further, speaking in any one language creates a competition process at both the cortical and the subcortical levels which is resolved by inhibition processes. Critically, in the case of bilinguals, managing separate linguistic systems (i.e., with respect to syntax, phonology, and prosody) requires an adaptation process, which results in similar, but nevertheless distinct processing of L1 and L2. In other words, the theory suggests that bilingual speakers (a) process their L1 differently from their L2, and (b) also process their L1 differently compared to purely monolingual speakers. In this section, we will provide an overview of studies that look at neural activations during bilingual word production. This has been addressed both in single-language and mixed-language contexts, which we will present in the two following subsections. We will mainly focus on within-speaker comparisons which allow us to evaluate directly to what extent native production differs from non-native production.

2.1. Production in a single-language context

One of the first studies that investigated hemodynamic differences between L1 and L2 word production was conducted by Klein, Milner, Zatorre, Meyer, & Evans (1995). In their PET study, four English-French bilinguals were asked to generate a rhyme word or a synonym in response to an auditorily presented word, either in their L1 or their L2. Behaviourally, participants were slower and less accurate during L2 synonym generation, while there were no differences in the rhyme word generation task. On the hemodynamic level, higher activation was observed in left inferior frontal cortex (BAs 45 and 47) and dorsolateral prefrontal cortex (BAs 8, 9, and 46) compared to a baseline task (word repetition). However, no differences were found between L1 and L2 production, suggesting that word generation is subserved by identical neural substrates in L1 and L2. In a similar study, Klein, Milner, Zatorre, Zhao, and Nikelski (1999) asked seven Mandarin-English bilinguals to generate nouns and verbs in response to an auditory cue, either in their L1 or L2. On the behavioural level, generating words in L2 took longer and resulted in more naming errors. Again, however, this processing deficit was not paralleled in the hemodynamic response: Both L1 and L2 naming caused higher activity in the left inferior cortex compared to a baseline task, but there were no significant differences between L1 and L2 naming. Thus, although participants were late learners of L2, there was no evidence for distinct neuronal signatures, in line with the single-network hypothesis.

Chee, Tan, and Thiel (1999) compared fifteen early and nine late Chinese-English bilinguals in an fMRI study using a covert word completion task. Participants were visually presented with either the beginning or the end of a word and instructed to produce the full word (e.g., “cou” to generate a word like “couple” or “ter” to generate a word like “water”), or they were shown a single Mandarin character and instructed to generate a compound. Compared to fixation as the control task, word generation increased activity in prefrontal, temporal, and parietal regions as well as the supplementary motor area. However, no significant differences were found between both L1 and L2 word generation, and between early and late bilinguals, again suggesting that L1 and L2 word retrieval recruit the same neural regions.

De Bleser et al. (2003), in a PET study, tested Flemish-French bilinguals in a covert picture naming task investigating the production of cognates. In bilingual speakers, cognates (i.e., words which have a similar form in the two languages) are typically named faster than non-cognates (Costa, Caramazza, & Sebastian-Galles, 2000; Costa, Santesteban, & Caño, 2005). Behaviourally (in an overt production task prior to the scanning session), this effect was indeed obtained: Naming in L2 was overall slower compared to L1, but naming non-cognates in L2 additionally slowed naming latencies compared to cognates. With respect to the hemodynamic response, activation was comparable between L1 cognates, L1 non-cognates, and non-cognates of the much later acquired L2. Only the presumably most difficult condition (i.e., L2 non-cognates) elicited higher activation in left inferior frontal and temporo-parietal regions (BAs 44, 47, and 20), again suggesting a role of these region in effortful lexical retrieval during L2 production as reported by Hernandez et al. (2000, 2001).

Perani et al. (2003) tested eleven highly proficient Spanish-Catalan bilinguals in a phonemic verbal fluency task. In the experiment, participants were asked to covertly generate as many words as possible beginning with a given letter in one of the two languages. Note that all had acquired their L2 (either Catalan or Spanish) at the age of three and showed no differences in proficiency. The fMRI results, however, showed a dissociation between the L1 and the L2 which was acquired in early childhood: Generating words in a speaker’s L2 resulted in an increased neural response compared to L1. Interestingly, this pattern differed depending on whether speakers were Catalan-Spanish or Spanish-Catalan speakers. While both groups showed higher activation in the left inferior and middle frontal gyrus and in the insula, Catalan-born speakers recruited additional resources in the left premotor cortex, left inferior parietal lobule, left caudate nucleus, and right inferior frontal gyrus. Importantly, the Catalan-born group was exposed less to their L2, so increases in activation compared to the Spanish-born group were interpreted as reflecting more effortful lexical processing. Overall, this study provides evidence that even with an early age of acquisition, producing words in L2 requires more resources, which is additionally modulated by exposure to the L2.

Extending the investigations to trilingual speakers, Vingerhoets et al. (2003) tested twelve native Dutch speakers with varying degrees of proficiency in their L2 and L3 (French and English) in a covert picture naming task and a phonemic verbal fluency task. Behavioural performance prior to the fMRI session indicated that fluency scores differed between all languages, and picture naming performance (measured as percent correct) was best for the L1, but comparable between the L2 and the L3. Imaging during the fluency task showed activation of left frontal and medial frontal/cingulate areas as well as of the left superior parietal lobule and a small right (anterior) temporal-(inferior) frontal region for all languages. When the task was carried out in the L2 and L3, respectively, an additional activation of left posterior temporal regions was found (BA 21). Note that Perani et al. (2003) had not obtained increased posterior temporal activation in a comparable task. However, their participants were highly proficient and immersed in both their L1 and L2, potentially resulting in a more automated (i.e., less effortful) use of the non-native language. In the picture naming task, all languages recruited medial/superior frontal and occipital (BAs 18 and 19) regions bilaterally. Covert naming in L2 and L3 additionally showed activation in left frontal superior parietal and temporal-occipital areas. When compared across languages, the right cuneate area (BA 19) was significantly more activated during L1 naming, whereas L2 and L3 showed additional predominantly left hemispheric medial frontal and inferior lateral frontal activation. Vingerhoets et al. (2003) concluded that speech production in a foreign language relies on largely overlapping areas, but that a proficiency deficit necessitates the recruitment of additional neural resources, predominantly in left frontal and posterior temporal regions. These findings, again, converge with those reported in previous studies attributing a supporting role of frontal and temporal areas to more effortful word production.

Liu, Hu, Guo, and Peng (2010) investigated neural differences during blocked L1 and L2 picture naming in 24 Chinese-English speakers. Compared to L1 naming, L2 naming resulted in higher error rates and elicited higher activation in left inferior frontal, precentral, and lingual gyri, the left cuneus, and supplementary motor areas, putamen, globus pallidus, caudate nucleus and cerebellum bilaterally. In line with previous findings, this increased activation across widespread areas was attributed to the recruitment of additional neural resources during the less automatic L2 production process, specifically with respect to lexical retrieval, articulatory processing, and cognitive control.

Another fMRI study on multilingual speakers was conducted by Videsott et al. (2010). Twenty native speakers of Ladin, a language spoken by a small community in South Tyrol, who were also proficient in Italian (L2), German (L3), and English (L4), were tested in an overt picture naming task that required the production of Ladin, Italian, and English nouns in separate blocks. Pre-experimental tests revealed that English was spoken with the lowest proficiency, which was confirmed by the highest error rates in this language compared to Ladin and Italian. Compared to L1 production, English picture naming resulted in higher activation in the left inferior frontal gyrus (BAs 6 and 44) and the cerebellum; compared to L2 production, only the cerebellum showed a stronger signal during English naming. Furthermore, composite proficiency scores for English and Italian were correlated with the neural response in the right dorsolateral prefrontal cortex (BA 45), such that higher proficiency was associated with higher activation in this region. This is at odds with other findings reporting lower cortical activity with increasing proficiency (Abutalebi & Green, 2007). However, in line with the results reported by Vingerhoets et al. (2003) on trilingual speakers, this study found an important role of the cerebellum in low-proficiency word production. Recall that Vingerhoets et al. (2003) only found increased activity in frontal and temporal regions during non-native word generation in trilingual speakers. Videsott et al. (2010) thus show that, at relatively low proficiency and in the presence of three competitor languages, the cerebellum may play a coordinating role for successful production.

Parker Jones et al. (2012) investigated speech production in 31 L2 speakers of English using a variety of production tasks. Depending on the visual display, participants, who had different first languages (Greek or European languages with a Latin-based script), were required to name objects, read out words, or respond “1, 2, 3” to a string of non-objects or Greek symbols, either in their L2 (English), or for the Greek participants, also in their L1. Behaviourally, the bilinguals performed equally well in both languages and did not differ from a group of 36 matched monolingual English speakers. The hemodynamic response, however, did show significant differences between the mono- and bilingual group. During picture naming and word reading, bilinguals showed higher activation than monolinguals in the left central sulcus, the planum temporale, superior temporal gyrus, inferior frontal gyrus and anterior insula. Interestingly, this increase was found even when bilingual participants responded in their L1. Furthermore, when the bilinguals were divided into two groups depending on how many languages they spoke overall (i.e., two vs. at least three), it was found that those who spoke three or more languages showed higher activation in the planum temporale, pars triangularis, and insula. The authors concluded that areas related to more effortful monolingual production underlie bilingual word production. For instance, effects of lexical frequency, which affects production performance in monolinguals, have been localised in the pars triangularis (Graves, Grabowski, Mehta, & Gordon, 2007). Speaking several languages, similarly, is considered to make lexical selection more difficult, recruiting resources from the same region.

2.2. Production in a mixed-language context

Hernandez, Martinez, and Kohnert (2000) conducted an fMRI study in which they tested six early highly proficient Spanish-English bilinguals in a cued switching task. Participants were asked to name pictures in response to a printed naming instruction (“say” or “diga” on each trial). Both L1 and L2 naming showed comparable activity in the dorsolateral prefrontal cortex (BA 9 and 46), the supramarginal gyrus (BA 40), inferior frontal gyrus (BAs 44 and 45), and superior temporal gyrus (BA 22). However, when comparing activity between blocks that required naming in only one language to blocks in which participants had to switch constantly between the two languages, the dorsolateral prefrontal cortex was activated more in the latter. Overall, these findings provided evidence for the involvement of the left dorsolateral prefrontal cortex in executive functioning. In a similar study, Hernandez, Dapretto, Mazziotta, and Bookheimer (2001) tested six early Spanish-English bilinguals who were more dominant in their L2 (English) than in their L1. The results replicated those from

Hernandez et al. (2000), such that activation patterns in left frontal regions were comparable between languages in the blocked naming condition, and the dorsolateral prefrontal cortex showed increased activity in the mixed naming condition. This again supports the notion that producing words in one's L1 and L2 recruits similar neural regions (in single-language contexts, at least), while the dorsolateral prefrontal cortex serves as a control structure in situations where rapid switches between languages are required.

Rodriguez-Fornells et al. (2005) employed a different approach to investigate word production in bilinguals. In a tacit picture-naming task to measure phonological competition in Spanish-German bilinguals, participants were asked to give a speeded button response to a picture. Specifically, a button press was required if the name of a given object (either in L1 or L2, depending on block) started with a consonant, but no response was required if it started with a vowel. To investigate cross-language activation, the objects were selected such that for half of the trials, both the L1 and the L2 name would require the same response (e.g., pressing a button in response to the picture of a syringe, which is "jeringuilla" in Spanish and "Spritze" in German, as opposed to pliers, which translates to "alicates" in Spanish and "Zange" in German). Eleven early Spanish-German bilinguals who at the time of testing reported to be more fluent in their L2 carried out this task in an fMRI session. Compared to a German monolingual control group, bilinguals were slower both when naming in their L1 and L2. Furthermore, reaction times were faster when both the L1 target word and the L2 translation required the same response, compared to a condition in which the target word began with a vowel, but the translation with a consonant, or vice-versa. On the hemodynamic level, this difference was manifested in higher activity when the onset of the target word and the translation were incongruent (i.e., the target word started with a vowel but the translation with a consonant, or vice versa) in left frontal regions (BAs 9 and 46) and the supplementary motor area (BA 6), as well as the anterior cingulate cortex (BAs 24 and 32). These regions were thus interpreted as playing a crucial role in language selection. Note that the finding of higher activity in frontal regions coincides with increased activity in dorsolateral prefrontal cortex in language switching (Hernandez et al., 2000, 2001). This provides evidence that greater involvement of the dorsolateral prefrontal cortex when language selection is difficult can be demonstrated with different language switching tasks.

Continuing investigations on the hemodynamics of language switching, Abutalebi et al. (2008) conducted an fMRI study in which they contrasted naming in L1 and L2 in both single-language and mixed-language contexts. Twelve German-French bilinguals completed three tasks, among which there was an object naming task that required switching between L1 and L2. In this condition, no latency differences were found between L1 and L2 naming, but error rates were higher for L2 naming. Comparing the hemodynamic response between L1 and L2 naming in this mixed-language context revealed higher activation during L2 naming predominantly in the left cingulate cortex, but also left inferior frontal gyrus (BAs 44, 45, and 47), middle frontal gyrus (BAs 10 and 46), precentral gyrus, right dorsal frontal gyrus (BA 9), right anterior cingulate cortex and putamen, and the caudate nucleus bilaterally. This increased activity was interpreted as reflecting higher processing demands during lexical selection in the non-dominant language and additionally provided further evidence for the involvement of frontal, but also subcortical regions during L2 production.

Addressing inhibition mechanisms in bilingual production, Guo, Liu, Misra, and Kroll (2011) examined inhibitory processes during L1 and L2 production in 24 Chinese-English bilinguals. Specifically, the authors investigated how global inhibition (i.e., complete suppression of the non-target language) and local inhibition (i.e., restricted suppression of a specific set of lexical representations) are manifested in bilinguals. Participants first were asked to name pictures solely in their L1 or L2, and afterwards in a mixed block which required continuous switching between languages. In a purely behavioral session following the fMRI session the authors found that globally (i.e., when L1 and L2 naming occurred in separate blocks), participants were slower and less accurate when they named pictures in their L2 compared to their L1. However, whether participants had started with their L1 or L2 did not affect naming performance. As for the local inhibition effects, which were operationalised by comparing performance between blocked and mixed naming, L1 mixed naming was slower and elicited more errors than L1 blocked naming, while L2 mixed naming was faster than L2 blocked naming. In the fMRI study, an increased activity during L1 naming following L2 naming was found in left dorsolateral prefrontal cortex, left temporal and parietal regions, and right postcentral gyrus, which were interpreted in terms of interference suppression and enhanced articulatory and visual control. By contrast, L2 naming following L1 naming only increased activity in the right cuneus and precuneus, which was ascribed to participants' awareness of picture presentation, given that these regions are linked to visual processing. Relative to blocked naming, mixed naming activated frontal regions including the dorsal anterior cingulate gyri bilaterally, supplementary motor area, precuneus, and pre- and postcentral gyri. Importantly, no differences were found between L1 and L2 naming in the mixed compared to the blocked naming condition. Especially the dorsal anterior cingulate cortex, precuneus, and supplementary motor area were interpreted as exerting control, inhibition, and selection functions during continuous language switching, which ascribes an important role of frontal and parietal regions in bilingual language control.

Finally, Abutalebi et al. (2013) reported evidence for the involvement of subcortical structures in multilingual word production. Fourteen native speakers of German who had acquired their L2 (Italian) at kindergarten age and their L3 (English) at school were tested in a mixed naming task. The authors found similar activation for L1, L2, and L3 naming, with increasingly higher activation of the pars opercularis and anterior cingulate cortex when naming in the later acquired language. Additionally, L3 naming increased activation in the left putamen. Subsequent voxel-based morphometry analysis showed that compared to a matched group of monolinguals, the trilingual speakers displayed increased grey matter density in the left putamen. This study thus provided striking evidence for the role of the left putamen in multilinguals, which has been interpreted as a control region for articulation in a less proficient language. In accordance with this finding, Burgaleta, Sanjuán, Ventura-Campos, Sebastian-Galles, and Ávila (2016) recently demonstrated expanded subcortical structures, specifically the bilateral putamen and thalamus, left globus pallidus, and right caudate nucleus, in bilingual speakers (see also Pliatsikas, DeLuca, Moschopoulou, & Saddy, 2017, for complementary findings). This finding supports the notion that especially subcortical areas are susceptible to anatomical changes in response to bilingualism (but see Mechelli et al., 2004, for evidence for increased grey matter density in the left inferior parietal cortex in bilinguals).

In summary, findings from PET and fMRI studies have provided ample evidence that word production in L1 and L2 is largely subserved by the same neural structures, namely the well-established fronto-temporal language network with contributions from superior parietal regions and the supplementary motor area, both in single- and mixed-language contexts. Moreover, subcortical structures have been related to production beyond the native language, which, next to control regions like the dorsolateral prefrontal cortex and the anterior cingulate cortex, may play a crucial role in language selection in multilinguals in situations of language switching. Critically, comparable activation for L1 and L2 naming, as well as higher activation during L2 naming has been reported. This dissociation can partly be explained by a difference in L2 proficiency and immersion levels of the tested speakers. Typically, higher proficiency and immersion result in more native-like cortical signatures (see also section 4 of this chapter), so discrepancies between studies may have been caused by examining different study populations (i.e., early vs. late bilinguals). A most interesting endeavour for future research would be to directly compare otherwise “matched” individuals from different proficiency and immersion spectra to get a more detailed picture of how these factors contribute to different cortical responses during production.

Critically, with the current state of affairs, most studies differ on too many aspects to allow for solid direct comparisons, thus hindering further theoretical advances which would allow to account for discrepancies between studies. These aspects concern, among others, sample size and statistical power, types of populations investigated, and proficiency measures used. It would be very useful to have at least some “standardization” of these aspects in future research.

3.0The neurophysiology of timing aspects of bilingual word production

Bilingualism often comes at a cost in terms of speech production performance. That is, bilinguals usually name pictures more slowly in their L2 than in their L1 (for an overview, see Hanulová, Davidson, & Indefrey, 2011), but on top of this, they also often name pictures more slowly in their L1 (Ivanova & Costa, 2008; Sadat, Martin, Alario, & Costa, 2012), experience more tip-of-the-tongue states (Gollan & Acenas, 2004), and score lower on verbal fluency tasks compared to monolingual speakers (Gollan, Montoya, & Werner, 2002; Portocarrero, Burright, & Donovick, 2007). While this may in part be explained by a greater cognitive control demand or linguistic intrusions from the L2 on the L1 (Costa & Sebastián-Gallés, 2014), the weaker links hypothesis (Gollan, Montoya, Cera, & Sandoval, 2008; Gollan, Montoya, Fennema-Notestine, & Morris, 2005) has addressed this phenomenon in terms of relative frequency. Contrary to monolingual speakers, bilinguals can refer to a to-be-expressed concept with two words instead of one, decreasing the frequency of the individual words. Put differently, the links between semantics and phonology in L1 and L2 are weaker, rendering this "bilingual disadvantage" ultimately the result of a functional frequency effect (see also Poulisse & Bongaerts, 1994).

One way of testing this hypothesis experimentally is by comparing lexical frequency effects in monolingual and bilingual speakers. In L1 production, high-frequency words are produced faster than low-frequency words (Oldfield & Wingfield, 1965; Wingfield, 1968). In a behavioral experiment, Gollan, Montoya, Cera, and Sandoval (2008) investigated whether this L1 finding translates into L2 production and whether bilinguals differ from monolinguals. 57 Spanish-English bilinguals who were more dominant in their L2 (English) produced high- and low-frequency words in response to pictures either in their L1, their L2, or in the language that first came to mind. Compared to matched English-speaking monolinguals, bilinguals named pictures more slowly in their L2, but also showed a larger frequency effect. Additionally, comparing bilinguals' naming latencies in L1 and L2 showed that the frequency effect was larger for the L1, which was the non-dominant language. The authors concluded that a decrease in use (i.e., the non-dominance of the L1) weakens the activation of the lexical entries of that language, which is especially pronounced for low-frequent words. In a similar vein, Ivanova and Costa (2008) showed that compared to Spanish monolinguals, Spanish-Catalan speakers who dominantly spoke Spanish showed a larger frequency effect, while the size of this effect did not differ between monolinguals and Catalan-Spanish bilinguals whose dominant language was Catalan.

Building on this rationale, Strijkers, Costa, and Thierry (2010) conducted an ERP study in which highly proficient Spanish-Catalan (*N* = 16) and Catalan-Spanish (*N* = 16) speakers named pictures manipulated for lexical frequency in Spanish (i.e., either the L1 or the L2). Behaviourally, a similarly sized frequency effect was found for both groups, implying that high proficiency in L2 abolishes frequency effect differences observed as a function of language dominance (Gollan et al., 2008; Ivanova & Costa, 2008). ERP amplitudes including the P2, P3, and N3 component, were larger for low-frequent words starting around 180ms after picture onset for both groups of speakers, suggesting that bilinguals rapidly access lexical properties of a target word, regardless of whether it is named in their first or second language. The items in this study additionally were chosen such that cognate status could be investigated as well. As with word frequency, non-cognates elicited larger ERP amplitudes (P2, P3, and N3) starting at around 190 ms after picture onset, with no differences between the two speaker groups. However, in L1 speakers, the frequency effect preceded the cognate effect, whereas they co-occurred in L2 speakers. Furthermore, overall P2 amplitudes were larger in L2 naming. It should be noted, however, that these two differences were small and statistically not robust, which may at least partly be caused by the fact that the authors employed a between-speaker design. To account for this caveat, Strijkers, Baus, Runnqvist, FitzPatrick, and Costa (2013) contrasted the temporal signatures of word frequency in L1 and L2 naming *within* 40 early Spanish-Catalan bilinguals. Unlike previous findings with bilinguals who are highly proficient in both languages, the behavioral frequency effect was larger for blocked naming in L2 compared to L1. ERP modulations were found as a function of word frequency (low vs. high) and response language (L1 vs. L2) from 140 ms after picture onset onwards, with low-frequent and L2 words eliciting a stronger response than high-frequent and L1 words. The latter effect was restricted to the participant group that started naming in L1. According to the authors, participants who started speaking in L2 had a stronger need to suppress L1 in this block, subsequently making L1 less accessible in the second block. Overall, the results of this study were taken as evidence for cortical differences between L1 and L2 naming already at the onset of lexical access.

Christoffels, Firk, and Schiller (2007) investigated the cognate facilitation effect in a group of 24 unbalanced German-Dutch bilinguals. Participants named cognate and non-cognate words either in a single-language (L1 or L2) or a mixed-language context which required switching between L1 and L2. In the single-language context, a larger behavioral cognate facilitation effect was observed for L2 compared to L1 naming, accompanied by larger negative amplitudes during cognate naming starting as early as 275 ms after picture onset. Surprisingly, in the mixed-language condition, cognates were named faster than non-cognates in L1 compared to L2 naming. Given that the cognate facilitation effect indexes to what extent the non-target language is activated, this finding was interpreted as a more effortful suppression of the more dominant non-target language (L1). With respect to switching between L1 and L2, non-switch trials caused shorter naming latencies and larger negative amplitudes than switch trials, with the ERP effect being present for L1 naming in the early time window and for both languages in the later time window. Overall, the authors concluded that in mixed language conditions (i.e., when switching between L1 and L2), language control is exerted primarily by actively inhibiting the more proficient and more dominant language rather than by modulating activation levels of the weaker L2.

In order to assess both spatial and temporal signatures of bilingual word production, Pang and MacDonald (2012) conducted a magnetoencephalography (MEG) study. Twelve bilingual participants with varying first and second languages were asked to produce action names in response to pictures, either in their L1 or their L2. Those who performed the task in their native language showed activation of the left inferior frontal gyrus (BA 47) up to 400ms after picture onset, as well as activation in the left frontal cortex (BAs 9, 13, and 47). By contrast, L2 production was characterised by sustained right insula (BA 13) activation and an early involvement (50 – 200 ms) of the right anterior cingulate cortex (BA 32). Based on previous findings from semantic judgment tasks, the authors related the insular desynchrony to L2 proficiency, and activation of the right anterior cingulate cortex to conflict and error monitoring.

In summary, ERP studies have provided behavioral support for the weaker links hypothesis, such that languages which are spoken less frequently or with lower proficiency elicit larger word frequency and cognate effects (cf. Strijkers et al., 2013). On the ERP (and MEG) level, however, diverging results have been reported. Importantly, the experimental procedure of the reported studies varied substantially (e.g., with respect to between/within-participant comparisons, mixed vs. blocked naming, L1 vs. L2 naming). Future research should therefore attempt to further clarify which factors influence the emergence of different brain potentials in bilinguals. A comparably simple way to achieve this is by replicating existing studies to see whether previously obtained results are robust (Button et al., 2013).

Moreover, it seems reasonable to focus future research on mixed-language contexts as this is a scenario more common to bilingual speakers and a better candidate to uncover differences in the electrophysiological response.

4.0 Factors moderating neural representations of bilingual word production

In the previous sections we purposefully interspersed details about the language background of speakers examined in the respective studies, because intuitively it makes a lot of sense that differences in L2 usage and proficiency may shape the representation and word production of bilinguals differentially. This intuition has received much empirical support. In the following, we will provide an overview of how age of acquisition, proficiency, and immersion may modulate bilingual word production in behavioural, electrophysiological, and hemodynamic measures.

Whether an early versus late acquisition of a second language affects the degree to which it is mastered is still a matter of ongoing debate. While it has originally been argued that there is a critical period of L2 acquisition, by now a bulk of research has found opposing evidence (Singleton, 2005; Vanhove, 2013). However, with respect to neural differences, it has been shown that in the domain of syntax, differences do exist, with higher connectivity between relevant language regions observed for highly proficient speakers compared to less proficient speakers during L2 sentence production (Abutalebi, 2008; Dodel et al., 2005). In the lexico-semantic domain, by contrast, no robust neural differences have been found (Indefrey, 2006; Perani & Abutalebi, 2005). A recent meta-analysis by Hengshuang Liu and Cao (2016), however, reports a crucial role of age of acquisition in mediating neural differences. That is, compared to L1 processing, L2 processing in late bilinguals was found to activate the left superior frontal gyrus (BAs 8 and 9) to a greater degree than in early bilinguals, arguing for a greater planning demand in late bilinguals. Additionally, the left fusiform gyrus (BA 37) of early bilinguals was activated more during L1 processing, which was interpreted as a parallel activation of the two orthographies. Note, however, that this meta-analysis did not focus exclusively on word production studies, but also included semantic and phonological judgment, comprehension, and repetition tasks. Thus, the direct influence of age of acquisition on neural differences in bilingual production cannot be derived from these results.

A less ambiguous factor modulating neural differences in bilingual word production is the proficiency level at which bilinguals master their second language. Notably, although there is no clear-cut standard as to how to quantify proficiency (scores used as regressors are often composites of highly variable and unstandardized measures), which may obscure between- and within-language comparisons, the evidence on proficiency effects on neural representations in bilinguals is fairly clear. The strongest finding is that less proficient bilinguals activate left prefrontal regions to a greater extent than highly proficient bilinguals. This finding has been related to increased cognitive control demands (Abutalebi, 2008). In line with the convergence hypothesis (Green, 2003), these differences diminish with increasing L2 proficiency. That is, as L2 production becomes more native-like, the need to exert cognitive control in juggling the two languages appears to decrease, and a speaker’s two languages activate identical areas (Perani & Abutalebi, 2005). Abutalebi and Green (2007) argued for three different possibilities determining less prefrontal activity with higher L2 proficiency. First, high proficiency likely results in stronger neural connections between a concept and its lexico-semantic and phonological properties, which in turn facilitates lexical retrieval. Second, high L2 proficiency may decrease the need to inhibit the L1 during L2 production. The more proficient speakers are in their L2, the more automatic L2 processing becomes, reducing interference from L1. Third, high proficiency invariably results in less cognitive effort as the L2 will have become more practised and retrieval thus more automatic. As a result, except for language switching situations, conflict resolution between L1 and L2 concepts may be less demanding, thus reducing the involvement of prefrontal regions. Note that cognitive control is inherent to all three possibilities. In sum, it can be assumed that high proficiency reduces the need for conscious control processes.

Directly related to the concept of proficiency is the degree to which bilinguals are exposed to their second language. Although constant practice undoubtedly has a beneficiary effect on mastering a second language, there is evidence that high immersion in an L2 environment influences neural representations on top of proficiency. For instance, Perani et al. (2003) reported that the exposure to a speaker's L2 differentially affects neural activations in picture naming. Despite comparable degrees of proficiency, less immersed speakers showed stronger activations in parietal and prefrontal regions including the left caudate nucleus. Analogous to experience and practice effects in monolingual naming (Thompson-Schill, D 'Esposito, & Kan, 1999), Perani and Abutalebi (2005) thus argued for immersion as an additional factor modulating the neurobiological underpinnings of bilingual word production. More direct evidence for this claim comes from a recent study by Pliatsikas, DeLuca, Moschopoulou, and Saddy (2017) who compared structural MRI scans of bilingual speakers currently living in an L2 environment to bilinguals without daily exposure to the L2. Compared to monolinguals, the highly immersed speakers showed a bilateral expansion of the putamen and globus pallidus, and right thalamus, while the bilateral thalamus and putamen were contracted in the less immersed speakers. Additionally, high immersion, quantified as the time living in the L2 environment, predicted the expansion of the right globus pallidus. These findings were taken as support for an exposure-related theory in which a bilingual's brain dynamically exhibits remarkable plasticity as a function of exposure to the second language.

5.0Conclusion

In this chapter, we have reviewed cognitive neuroscience studies investigating bilingual word production. While many studies reported comparable neurophysiological and hemodynamic signals for both L1 and L2 processing, there is also ample evidence that more effortful processing in L2 results in higher activation of left frontal and temporal areas, and that a number of areas (e.g., the dorsolateral prefrontal cortex, the anterior cingulate cortex, and the supplementary motor areas) appear to be specialised for the coordination of a speaker's two languages within a control network. Furthermore, speaker-specific factors such as L2 proficiency and exposure have been shown to influence neural representations. Many methodological advances have provided a clearer picture of the mechanisms underlying bilingual speech production, although they have not been immune to criticism (García-Pentón, Fernández García, Costello, Andoni Duñabeitia, & Carreiras, 2017). Finally, one should keep in mind that the vast majority of the studies have investigated bilingual production at the word-level, although, as we have mentioned in the introduction, this is a simplification of natural language production which usually occurs at the sentence and discourse level. Similar to the investigation of monolingual language production, it will therefore remain a conceptual and methodological challenge to further illuminate the neurobiological processes of bilingual language production at the sentence and discourse level.

References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*, *128*(3), 466–478. http://doi.org/10.1016/j.actpsy.2008.03.014

Abutalebi, J., Annoni, J.-M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., … Khateb, A. (2008). Language Control and Lexical Competition in Bilinguals: An Event-Related fMRI Study. *Cerebral Cortex*, *18*, 1496–1505. http://doi.org/10.1093/cercor/bhm182

Abutalebi, J., & Green, D. W. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242–275. http://doi.org/10.1016/j.jneuroling.2006.10.003

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and Cognition*, *19*(4), 689–698. http://doi.org/10.1017/S1366728916000225

Abutalebi, J., Rosa, P. A. Della, Castro Gonzaga, A. K., Keim, R., Costa, A., & Perani, D. (2013). The role of the left putamen in multilingual language production. *Brain and Language*, *125*(3), 307–315. http://doi.org/10.1016/j.bandl.2012.03.009

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, *16*(4), 240–50. http://doi.org/10.1016/j.tics.2012.03.001

Boukadi, M., Davies, R., & Wilson, M. A. (2015). Bilingual lexical selection as a dynamic process: Evidence from Arabic-French bilinguals. *Canadian Journal of Experimental Psychology*, *69*(4), 297–313. http://doi.org/10.1037/cep0000063

Burgaleta, M., Sanjuán, A., Ventura-Campos, N., Sebastian-Galles, N., & Ávila, C. (2016). Bilingualism at the core of the brain. Structural differences between bilinguals and monolinguals revealed by subcortical shape analysis. *NeuroImage*, *125*, 437–445. http://doi.org/10.1016/j.neuroimage.2015.09.073

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. http://doi.org/10.1038/nrn3475

Chee, M. W. L., Tan, E. W. L., & Thiel, T. (1999). Mandarin and English Single Word Processing

Studied with Functional Magnetic Resonance Imaging. *Journal of Neuroscience*, *19*, 3050–3056.

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, *1147*, 192–208. http://doi.org/10.1016/j.brainres.2007.01.137

Colomé, À. (2001). Lexical Activation in Bilinguals' Speech Production: Language-Specific or Language-Independent? *Journal of Memory and Language*, *45*, 721–736. http://doi.org/10.1006

Costa, A. (2004). Speech Production in Bilinguals. In T. K. Bhatia & W. C. Ritchie (Eds.), *The Handbook of Bilingualism* (pp. 201–223). Oxford, UK: Blackwell Publishing Ltd. http://doi.org/10.1002/9780470756997.ch8

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*(3), 231–244.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The Cognate Facilitation Effect: Implications for Models of Lexical Access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283–1296. http://doi.org/10.1037/0278-7393.26.5.1283

Costa, A., Colomé, À., Gómez, O., & Sebastián-Gallés, N. (2003). Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. *Bilingualism: Language and Cognition*, *6*(3), 167–179. http://doi.org/10.1017/S1366728903001111

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. http://doi.org/10.1006/jmla.1999.2651

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. http://doi.org/10.1016/j.bandl.2004.12.002

Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience*, *15*, 336–345. http://doi.org/10.1038/nrn3709

de Bleser, R., Dupont, P., Postler, J., Bormans, G., Speelman, D., Mortelmans, L., & Debrock, M.

(2003). The organisation of the bilingual lexicon: a PET study. *Journal of Neurolinguistics*, *16*, 439–456.

De Bot, K. (1992). A Bilingual Production Model: Levelt's "Speaking" Model Adapted. *Applied Linguistics*, *13*, 1–24.

Dodel, S., Golestani, N., Pallier, C., ElKouby, V., Le Bihan, D., & Poline, J.-B. (2005). Condition-dependent functional connectivity: syntax networks in bilinguals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1457), 921–935. http://doi.org/10.1098/rstb.2005.1653

García-Pentón, L., Fernández García, Y., Costello, B., Andoni Duñabeitia, J., & Carreiras, M. (2017). The neuroanatomy of bilingualism: how to turn a hazy view into the full picture. *Language, Cognition and Neuroscience*, *31*(3), 303–327. http://doi.org/10.1080/23273798.2015.1068944

Gollan, T. H., & Acenas, L.-A. R. (2004). What Is a TOT? Cognate and Translation Effects on Tip-of-the-Tongue States in Spanish-English and Tagalog-English Bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 246–269. http://doi.org/10.1037/0278-7393.30.1.246

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. http://doi.org/10.1016/j.jml.2007.07.001

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), 1220–34.

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology*, *16*(4), 562–576. http://doi.org/10.1037/0894-4105.16.4.562

Graves, W. W., Grabowski, T. J., Mehta, S., & Gordon, J. K. (2007). A Neural Signature of Phonological Access: Distinguishing the Effects of Word Frequency from Familiarity and Length in Overt Picture Naming. *Journal of Cognitive Neuroscience*, *19*(4), 617–631. http://doi.org/10.1162/jocn.2007.19.4.617

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language* *and Cognition*, *1*, 67–81. http://doi.org/10.1017/S1366728998000133

Green, D. W. (2003). Neural basis of lexicon and grammar in L2 acquisition: The convergence hypothesis. In R. van Hout, A. Hulk, F. Kuiken, & R. Towell (Eds.), *The Lexicon-Syntax Interface in Second Language Acquisition* (pp. 197–218). Amsterdam: John Benjamins. Retrieved from http://discovery.ucl.ac.uk/164901/

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. http://doi.org/10.1080/20445911.2013.796377

Grosjean, F. (2010). *Bilingual: Life and Reality*. Cambridge, Massachusetts; London, England: Harvard University Press.

Guo, T., Liu, H., Misra, M., & Kroll, J. F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese-English bilinguals. *Neuroimage*, *56*(4), 2300–2309. http://doi.org/10.1016/j.neuroimage.2011.03.049

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. http://doi.org/10.1080/01690965.2010.509946

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, *1*(3), 213–229. http://doi.org/10.1017/S1366728998000364

Hermans, D., Ormel, E., van Besselaar, R., & van Hell, J. (2011). Lexical activation in bilinguals' speech production is dynamic: How language ambiguous words can affect cross-language activation. *Language and Cognitive Processes*, *26*(10), 1687–1709. http://doi.org/10.1080/01690965.2010.530411

Hernandez, A. E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language Switching and Language Representation in Spanish–English Bilinguals: An fMRI Study. *NeuroImage*, *14*(2), 510–520. http://doi.org/10.1006/nimg.2001.0810

Hernandez, A. E., Martinez, A., & Kohnert, K. (2000). In Search of the Language Switch: An fMRI

Study of Picture Naming in Spanish–English Bilinguals. *Brain and Language*, *73*(3), 421–431. http://doi.org/10.1006/brln.1999.2278

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*(s1), 279–304. http://doi.org/10.1111/j.1467-9922.2006.00365.x

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277–288. http://doi.org/10.1016/j.actpsy.2007.06.003

Jacobs, A., Fricke, M., & Kroll, J. F. (2016). Cross-Language Activation Begins During Speech Planning and Extends Into Second Language Speech. *Language Learning*, *66*(2), 324–353. http://doi.org/10.1111/lang.12148

Klein, D., Milner, B., Zatorre, R. J., Meyer, E., & Evans, A. C. (1995). The neural substrates underlying word generation: a bilingual functional-imaging study. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(7), 2899–903. http://doi.org/10.1073/PNAS.92.7.2899

Klein, D., Milner, B., Zatorre, R. J., Zhao, V., & Nikelski, J. (1999). Cerebral organization in bilinguals: A PET study of Chinese-English verb generation. *Neuroreport*, *10*, 2841–2846.

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychologica*, *128*(3), 416–30. http://doi.org/10.1016/j.actpsy.2008.02.001

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, *9*(2), 119–135. http://doi.org/10.1017/S1366728906002483

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1-38-75. http://doi.org/10.1017/S0140525X99001776

Liu, H., & Cao, F. (2016). L1 and L2 processing in the bilingual brain: A meta-analysis of neuroimaging studies. *Brain & Language*, *159*, 60–73. http://doi.org/10.1016/j.bandl.2016.05.013

Liu, H., Hu, Z., Guo, T., & Peng, D. (2010). Speaking words in two languages with one brain: Neural overlap and dissociation. *Brain Research*, *1316*, 75–82. http://doi.org/10.1016/j.brainres.2009.12.030

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Neurolinguistics: Structural plasticity in the bilingual brain. *Nature*, *431*(7010), 757–757. http://doi.org/10.1038/431757a

Misra, M., Guo, T., Bobb, S. C., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, *67*(1). http://doi.org/10.1016/j.jml.2012.05.001

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *The Quarterly Journal of Experimental Psychology*, *17*(4), 273–81.

Pang, E. W., & MacDonald, M. J. (2012). An MEG study of the spatiotemporal dynamics of bilingual verb generation. *Brain Research*, *1467*, 56–66. http://doi.org/10.1016/j.brainres.2012.05.054

Parker Jones, O., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., … 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. http://doi.org/10.1093/cercor/bhr161

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, *15*(2), 202–206. http://doi.org/10.1016/j.conb.2005.03.007

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping*, *19*(3), 170–182. http://doi.org/10.1002/hbm.10110

Pliatsikas, C., DeLuca, V., Moschopoulou, E., & Saddy, J. D. (2017). Immersive bilingualism reshapes the core of the brain. *Brain Structure and Function*, *222*(4), 1785–1795. http://doi.org/10.1007/s00429-016-1307-9

Portocarrero, J., Burright, R., & Donovick, P. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, *22*(3), 415–422. http://doi.org/10.1016/j.acn.2007.01.015

Poulisse, N. (2000). Slips of the tongue in first and second language production. *Studia Linguistica*, *54*(2), 136–149. http://doi.org/10.1111/1467-9582.00055

Poulisse, N., & Bongaerts, T. (1994). First Language Use in Second Language Production. *Applied Linguistics*, *15*(1), 36–57. http://doi.org/10.1093/applin/15.1.36

Rodriguez-Fornells, A., Lugt, A. van der, Rotte, M., Britti, B., Heinze, H.-J., & Münte, T. F. (2005). Second Language Interferes with Word Production in Fluent Bilinguals: Brain Potential and Functional Imaging Evidence. *Journal of Cognitive Neuroscience*, *17*(3), 422–433. http://doi.org/10.1162/0898929053279559

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1–3), 107–42.

Sadat, J., Martin, C. D., Alario, F. X., & Costa, A. (2012). Characterizing the Bilingual Disadvantage in Noun Phrase Production. *Journal of Psycholinguistic Research*, *41*, 159–179. http://doi.org/10.1007/s10936-011-9183-1

Singleton, D. (2005). The Critical Period Hypothesis: A coat of many colours. *IRAL - International Review of Applied Linguistics in Language Teaching*, *43*(4), 269–285. http://doi.org/10.1515/iral.2005.43.4.269

Spalek, K., Hoshino, N., Wu, Y. J., Damian, M. F., & Thierry, G. (2014). Speaking two languages at once: Unconscious native word form access in second language production. *Cognition*, *133*(1), 226–231. http://doi.org/10.1016/j.cognition.2014.06.016

Strijkers, K., Baus, C., Runnqvist, E., FitzPatrick, I., & Costa, A. (2013). The temporal dynamics of first versus second language production. *Brain and Language*, *127*(1), 6–11. http://doi.org/10.1016/j.bandl.2013.07.008

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. http://doi.org/10.1093/cercor/bhp153

Thompson-Schill, S. L., D ’Esposito, M., & Kan, I. P. (1999). Effects of Repetition and Competition on Activity in Left Prefrontal Cortex during Word Generation. *Neuron*, *23*, 513–522.

Vanhove, J. (2013). The Critical Period Hypothesis in Second Language Acquisition: A Statistical

Critique and a Reanalysis. *PLoS One*, *8*(7), e69172. http://doi.org/10.1371/journal.pone.0069172

Videsott, G., Herrnberger, B., Hoenig, K., Schilly, E., Grothe, J., Wiater, W., … Kiefer, M. (2010). Speaking in multiple languages: Neural correlates of language proficiency in multilingual word production. *Brain & Language*, *113*, 103–112. http://doi.org/10.1016/j.bandl.2010.01.006

Vingerhoets, G., Van Borsel, J., Tesink, C., van den Noort, M., Deblaere, K., Seurinck, R., … Achten, E. (2003). Multilingualism: an fMRI study. *NeuroImage*, *20*(4), 2181–2196. http://doi.org/10.1016/j.neuroimage.2003.07.029

Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *The American Journal of Psychology*, *81*(2), 226–34.