Research Article

The influence of inhibitory skill on phonological representations in production and perception ☆

Shiri Lev-Ari [a,b,*], Sharon Peperkamp [b]

[a] Max Planck Institute for Psycholinguistics, 1 Wundtlaan, 6525 XD, Nijmegen, The Netherlands
[b] Laboratoire de Sciences Cognitives et Psycholinguistique, Département d'Etudes Cognitives, Ecole Normale Supérieure – PSL Research University / EHESS / CNRS, 29, rue d'Ulm, Paris 75005, France

ABSTRACT

Inhibition is known to play a role in speech perception and has been hypothesized to likewise influence speech production. In this paper we test whether individual differences in inhibitory skill can lead to individual differences in phonological representations in perception and production. We further examine whether the type of inhibition that influences phonological representation is domain-specific or domain-general. Native French speakers read aloud sentences with words containing a voiced stop that either have a voicing neighbor (target) or not (control). The duration of pre-voicing was measured. Participants similarly performed a lexical decision task on versions of these target and matched control words whose pre-voicing duration was manipulated. Lastly, participants performed linguistic and non-linguistic inhibition tasks. Results indicate that the lower speakers' linguistic or non-linguistic inhibition is, the easier it is for them to recognize words with a voiceless neighbor when these words have a shorter, intermediate, pre-voicing rather than a longer one. Inhibitory skill did not predict recognition time for control words, indicating that the effect was due to the greater activation of the voiceless neighbor. Inhibition did not predict pre-voicing duration in production. These results indicate that individual differences in cognitive skills can influence phonological representations in speech perception.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Spoken language is characterized by inter- and intra-individual variability. For example, one English speaker's /ɔ/ is another English speaker's /ɑ/ (Peterson & Barney, 1952). Furthermore, the same speaker might produce the same feature differently in different words or occasions depending on linguistic and articulatory constraints, among others. Thus, the voice onset time of English voiceless stops is shorter when the stop is in medial position before an unstressed vowel (Lisker, 1977). Even in the same word, VOT can vary along contextual factors, such as in convergence with the VOT values produced by a recent speaker (Shockley, Sabadini, & Fowler, 2004).

Language comprehension similarly shows both inter- and intra-individual differences. For example, people differ in their ease of lexical retrieval, and their ease of retrieval of a word will depend on contextual factors, among others, leading to greater ease in some contexts than in others. For instance, people rely on the semantic context during language processing, and the more semantically constraining a context is, the easier it is to process the predicted word, yet the ability to rely on this context depends on cognitive resources (Federmeier & Kutas, 2005). People also differ in their sensitivity to gradation and contrasts of different linguistic cues (Kong & Edwards, 2011; Law, Fung, & Kung, 2013) as well as in their ability or tendency to integrate or disregard certain types of cues (Repp, 1981; Stewart & Ota, 2008; Yu, 2010). Few studies have examined the source of such individual differences. One recent emerging pattern, however, suggests that differences in "autistic" traits might influence integration of cues. For example, people with higher yet still typical values of autistic traits are less influenced by lexical factors during phoneme classification (Stewart & Ota, 2008).

Here we examine whether individual differences in a cognitive skill – inhibition – can also lead to inter- and intra-individual variability in the production and perception of phonetic features. We focus on voicing in French stops, and test whether there is a

relationship between French speakers' inhibitory skill and their production and perception of voiced stops in words that have a voicing neighbor. Specifically, we examine whether lower inhibitory skill leads to the production of intermediate voicing values in such words, but not in control words, because of the simultaneous activation of the competing voicing feature in the voicing neighbor. Similarly, we examine whether lower inhibitory skills facilitates the perception of such words when they are produced with intermediate voicing values because of their greater match with the activated voicing values.

1.1. Inhibition in speech perception

Speech perception is often described as the mapping of an acoustic signal to a lexical item via forward spreading activation from the activated phonemes to lexical items in which they occur, with gradual narrowing of activation as more information is received. Importantly, most speech perception models (e.g., Luce & Pisoni, 1998; McClelland & Elman, 1986; Norris, 1994 but see Gaskell & Marslen-Wilson, 1997; Norris & McQueen, 2008) posit that phonetic neighbors compete during this process and hypothesize inhibitory mechanisms to deal with this competition. Indeed, empirical evidence supports the notion of such competition by showing that speech perception depends on neighborhood density, such that words from denser neighborhoods are recognized more slowly and less accurately (Luce & Pisoni, 1998). Similarly, direct comparison of response times to words that have a neighbor differing in voicing to those that do not, show that those without a neighbor are recognized faster (Andruski, Blumstein, & Burton, 1994).

1.2. Inhibition in speech production

Models of speech production similarly propose that phonetic neighbors are activated during speech planning, and some of these models hypothesize that inhibition is exercised to overcome the competition from these neighbors (Meyer & Gordon, 1985; Stemberger, 1985). In accordance with this hypothesis, Wheeldon (2003) showed that picture naming is delayed if participants just produced a phonetically related word rather than a phonetically unrelated word in response to a definition.

At the same time, other models of speech production (Dell, 1986; Roelofs, 1997) do not assume any inhibitory connections between competing items. Indeed, many studies not only fail to find effects of inhibition between phonetically related words, but, in contrast, find that such competition facilitates production. For example, Meyer & Schriefers (1991) showed that picture naming is facilitated if the speaker has just heard a phonetically related compared with an unrelated word.

Studies examining the role of neighborhood density have similarly exhibited conflicting results. Some studies found neighborhood density to inhibit production (e.g., Sadat, Martin, Costa, & Alario, 2014; Vitevich & Stamer, 2006), while others found that words from denser neighborhoods are produced more quickly than words from sparser neighborhoods (e.g., Vitevitch, 2002). Some factors, such as frequency, and the position of the competing segment (onset vs. rhyme) have been proposed to account for some of these conflicting results (e.g., Colombo, 1986; Wheeldon, 2003). Others have argued that differences are due to a lack of control of other factors in the stimuli, or to the simultaneous facilitatory priming links and inhibitory selection links between lexical and phonological units (Chen & Mirman, 2012; Sadat et al., 2014). For now, the issue of whether and when neighbors inhibit production has yet to be fully resolved.

1.3. Individual differences in inhibitory skill and phonological representation

Considering the role that inhibition might play in speech production and perception, one might wonder how individual differences in inhibitory skill might influence production and perception abilities, and more importantly, whether individual differences in inhibitory skill can even lead to individual differences in linguistic representations.

Some preliminary evidence attests to the influence of inhibitory skill on perception ability. In particular, when embedded in noise, words of denser neighborhoods are recognized less accurately, and, crucially, this effect of neighborhood density is larger for those who have lower inhibitory score, as measured by a lexical Stroop task (Taler, Aaron, Steinmetz, & Pisoni, 2010). Similar evidence has been found at other linguistic levels with bilinguals. In general, bilinguals must inhibit the language not in use during language processing and production (Green, 1998), but this inhibition is rarely fully successful (Spivey & Marian, 1999). Importantly, it has been shown that success at inhibition depends, among other things, on individuals' level of inhibitory skill (Gollan, Sandoval, & Salmon, 2011; Linck, Schwieter, & Sunderman, 2012).

Further evidence shows that individual differences in inhibitory skill during language processing can lead to individual differences in representation. Looking at bilinguals, Lev-Ari and Peperkamp (2013) showed that individual differences in inhibition of competing languages lead late bilinguals with lower inhibitory skill to produce voiceless stops in their native and dominant language with Voice Onset Times (VOT) that are closer to the canonical values of their second language. Additionally, lower inhibitory skill leads bilinguals to discriminate between native languagephonetic neighbors differing in their VOT in a manner more similar to the way monolingual speakers of their second language discriminate them.

The goal of the current study is to examine whether individual differences in inhibitory skill can similarly influence phonological representations among monolinguals. Specifically, this study tests whether, in cases where words have a neighbor differing in voicing, the co-activation of this neighbor leads individuals to form representations with intermediate VOTs and, consequently, be faster at recognizing the words when they have intermediate rather than canonical voicing values. The logic is that if individuals with lower inhibitory skill habitually experience greater activation of the competing neighbor, and therefore the competing voicing feature, then their representation of the word would have an intermediate voicing value, as in the case of the intermediate VOTs of English–French bilinguals with relatively low inhibitory skill. We assume here, in line with the common phonological tradition, that negative and

positive VOTs lie on a single continuum, as VOT is a feature that differentiates stops in many languages, the boundary can occur at different points on the continuum (Lisker & Abramson, 1964), and people are sensitive to gradation in VOT (McMurray, Tanenhaus, & Aslin, 2002). Consequently, word tokens with an intermediate voicing value would have greater average similarity to the activated words among individuals with lower inhibitory skill, and would therefore activate the stored representation more efficiently (Andruski et al., 1994). Individuals with better inhibitory skill, in contrast, would better inhibit the competing neighbor, including its competing voicing feature. Therefore, their word representation will not have intermediate voicing values, leading their representations to be more dissimilar from token words with intermediate voicing values, and therefore be less efficiently primed by them.

Additionally, we test whether individuals with lower inhibitory skill would produce more intermediate voicing values for words that have neighbors differing in voicing. As with perception, lower inhibitory skill could lead to greater activation of the voicing neighbor, and co-activation of two competing voicing features could lead to production that is influenced by both features, i.e., showing intermediate values. At the same time, as some previous research suggests that production does not involve inhibition of competing neighbors (Dell, 1986; Roelofs, 1997), it is possible that the level of co-activation that competing neighbors enjoy does not depend on individual differences in inhibitory skill. Moreover, as neighbors do not seem to inhibit one another in production, it might be the case that having a neighbor differing in voicing does not influence production values at all, independently of inhibitory skill.

1.4. The nature of inhibition

Another question that we explore in this study regards the nature of the inhibition that individuals use during language production and perception. Some previous research has suggested that while bilinguals rely on the general executive function for controlling their languages, monolinguals rely on domain-specific, i.e., linguistic, inhibition to control competition during language processing. For example, Blumenfeld and Marian (2011) found that performance on a non-linguistic inhibition task predicted how quickly bilinguals resolved lexical competition but not how quickly monolinguals did. Another suggestive finding is that bilinguals activate different brain regions and networks than monolinguals when performing conflict tasks, even when their behavioral results are similar (Bialystok et al., 2005; Garbin et al., 2010; Luk, Anderson, Craik, Grady, & Bialystok, 2010). These regions and networks, however, do not clearly map onto domain-general and specifically language-related networks, respectively. Therefore, our final goal in this study is to examine whether it is linguistic or non-linguistic inhibition that relates to phonological representation in monolinguals.

2. Study

This study tests whether individual differences in inhibitory skill lead to individual differences in the production and perception of voicing in words that have a voicing neighbor among native French speakers. French is a short lag language that is characterized by pre-voicing in the production of voiced stops, and by short positive Voice Onset Times for voiceless stops (Lisker & Abramson, 1964). As voicing is contrastive, some words in French have a neighbor differing in voicing. For instance, *cod*é ('coded') has a voicing neighbor *cot*é ('listed'). Consequently, whenever the word *cod*é is produced or heard, its neighbor *cot*é is activated.

We used the existence of words with a voicing neighbor to test our hypotheses. Specifically, we manipulated the amount of pre-voicing in pairs of matched words that differed in whether they had a voiceless neighbor, and measured participants' response time to these words in a lexical decision task. As previous research has shown, response time in lexical decision tasks depends on how representative the target word's VOT is (Andruski et al., 1994). Additionally, we embedded those pairs of words in matched pairs of sentences. We asked participants to read these sentences aloud, and measured their amount of pre-voicing in the paired words. Finally, we measured participants' inhibitory skill using both linguistic and non-linguistic tasks, and examined the relation between performance on the inhibition tasks and performance on the speech perception and production tasks.

2.1. Method

2.1.1. Participants

Thirty-seven native French speakers (24 women; Age: 19–30, $M = 23$) participated in the experiment. The sentence reading task was not recorded for one participant by error, and the lexical decision data was not collected from another participant by error. Therefore, both production and perception analyses are based on the data of 36 participants. The analysis of the link between perception and production is based on the 35 participants who contributed data to both tasks.

2.2. Stimuli and procedure

2.2.1. Sentence reading task

Twenty pairs of words were selected such that each pair consisted of disyllabic or trisyllabic words that have a voiced stop in either the second or third syllable, but with only one of the words having a neighbor differing in voicing for that position, e.g., *cod*é (neighbor: *cot*é) – *brod*é (no voicing neighbor) (see Appendix A for a complete list). The target and control words were always of the same part of speech and had the same number of syllables. They were additionally matched in terms of frequency and neighborhood density using Lexique (New, Pallier, Ferrand, & Matos, 2001) ($t(38) < 1$ and $t(38) = 1.56$, $p > 0.05$, respectively). The voicing competitors were

always at least as frequent as the target words to ensure that they pose sufficient competition. These paired words were then embedded in pairs of sentences. Paired sentences were similar in length and structure. For example:

(a) Le message a été codé par quelqu'un sans beaucoup d'expérience.
   'The message was coded by someone without a lot of experience'.
(b) La robe a été brodée par quelqu'un sans beaucoup d'expérience.
   'The dress was embroidered by someone without a lot of experience'.

Twenty additional pairs of sentences similar in structure and length were generated and served as fillers. Sentences were then organized in one quasi-random order, ensuring that the first two sentences were filler sentences and that paired sentences were separated by at least two other sentences. Participants were asked to read aloud the sentences at their own pace. If participants were disfluent during their reading of one of the sentences, they were asked to re-read it. Once participants finished reading the list of sentences, they were asked to read the list once more. Participants' production was recorded.

2.2.2. Lexical decision

The 20 word pairs from the sentence reading task were used in the lexical decision task. A native French speaker read aloud a list of these 40 words, as well as 45 filler words and 85 nonwords. Seven additional native French speakers read the critical 40 words in order to ensure that the speaker's pre-voicing values fall within the range of average speakers. When the speaker's values were higher or lower than all others, they were trimmed or extended such that they would fall within the other speakers' range. In total, the pre-voicing of 23 words (11 target words, and 12 control words) had to be trimmed, and that of 2 words (1 target and 1 control) had to be extended. Then, another version of all target and control words was created by trimming the pre-voicing of all target voiced stops to 67% of their original duration. Even though individuals might be sensitive to even finer differences in voice onset time (McMurray et al., 2002, but see Andruski et al., 1994), we decided to take a difference that is relatively large, so it could distinguish individuals with different levels of inhibition without rendering the words ambiguous and while ensuring that the items are still natural-sounding. The average pre-voicing duration of the "long" versions was 82.1 ms, and the average pre-voicing duration of the "short" versions was 54.6 ms (see Fig. 1 for an example).

The lexical decision task was carried out in two blocks. Each block contained one version of each of the target and control words, such that the shorter version was included for half of the words, and the longer version for the other half. Forty filler words and 80 nonwords were included as well (the same ones in both blocks). Words were presented in a random order. The order of the two blocks was counterbalanced across participants, and for all participants, 10 practice trials with five additional filler words and five additional nonwords preceded the first block.

On each trial, a word or a nonword was presented auditorily. Participants' task was to respond as quickly as possible whether it was a word. A delay of 1500 ms separated each response from the following item. The dependent measure was the difference between response time to the long and short pre-voicing versions of each word (after a subtraction of the difference in duration between the two versions).

2.2.3. Stroop task

We used a manual version of the Stroop task (e.g., Logan, Zbrodoff, & Williamson, 1984). The Stroop task measures interference from an irrelevant but automatically activated aspect of the stimuli that needs to be inhibited. We selected this task as it is a well-established and
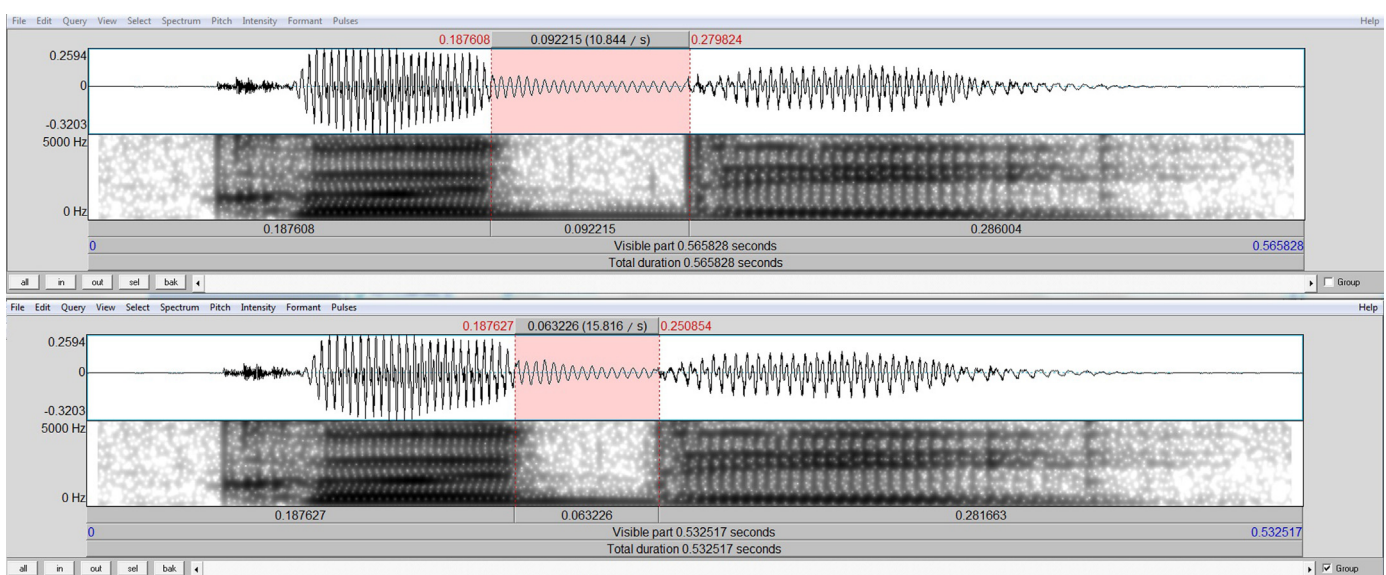

Fig. 1. The long and short versions of the token codé ('coded').

reliable measure of a type of inhibition that is similar to the inhibition that might be used to inhibit competing lexical items in language processing. As in the traditional Stroop task, participants have to indicate the color of visually presented words. Importantly, the words are often color words, and in some trials, there is a mismatch between the ink color and the semantics of the word. In the manual version, participants indicate the color by pressing a key. There are four conditions: (1) congruent, in which the semantics of the color word matches the ink color; (2) incongruent, in which the semantics of the color word does not match the ink color; (3) neutral, in which the symbol string @@@@ appears instead of a color word; (4) reading, in which the color word is written in gray, and participants' task is to read the word and indicate which color is written rather than indicate the color of the ink. The last condition is more common in the manual version of the task, and is intended to increase interference.

Our task used four colors – blue, red, green and yellow. It started with a learning phase, in which participants learned the association between the colors and their respective keys (100 trials). In this phase, participants did not see words but a colored symbol string @@@@. After completing the learning phase, participants performed 40 practice trails of the task itself, followed by two blocks of the task, each containing 72 trials.

2.2.4. Simon task

The Simon task measures interference from an irrelevant aspect of a stimulus (Craft & Simon, 1970). It is conceived of as a spatial, non-linguistic, version of the Stroop task. In this task, participants are asked to indicate the direction in which an arrow is pointing. In the version we used, participants indicated the direction of the arrow by pressing a key with their right index finger to indicate that the arrow is pointing right, and pressing a key with their left index finger to indicate that it is pointing left. The arrow appeared in the right half of the screen on a third of the trials, the left half of the screen in another third of the trials, and in the middle of the screen in the rest of the trials. When the position of the arrow (right vs. left half of the screen) does not match the direction that the arrow is pointing at, there is a conflict and response time increases. The efficiency of inhibition is therefore measured by comparing response time in conflict trials to response time in neutral trials. The version of the task that we used included 6 practice trials and 120 critical trials (40 congruent, 40 incongruent and 40 neutral trials) in two blocks.

All participants performed the tasks in the following order: Sentence Reading task, Stroop task, the first half of the Lexical Decision task, Simon task, and the second half of the Lexical Decision task.[1] The sentence reading task preceded the lexical decision tasks to prevent convergence in production to the pre-voicing values of the manipulated stimuli in the lexical decision task (Shockley et al., 2004). The lexical decision task was cut into two parts separated by an inhibition task, because each participant responded to two versions of the same word (one with a short pre-voicing and one with a long pre-voicing), and this way priming from the first presentation to the second one was minimized (Forster & Davis, 1984).

2.3. Results

Participants' Stroop score, calculated as the difference between performance on the incongruent trials and the control trials, ranged from −6 ms to 169 ms ($M=43$, $SE=37$). Participants' Simon score, calculated as the difference between performance on the incongruent trials and performance on the control trials ranged from −102 ms to 713 ms ($M=247$, $SE=189$). Performance on the Stroop and Simon tasks only marginally correlated ($r=0.32$, $p<0.06$). To minimize coding-related correlations among factors, the Stroop and Simon scores were centered for the analyses.

2.3.1. Perception

To test whether inhibitory skill influences perception of pre-voicing in words with a voicing-neighbor, we tested whether performance on the inhibition tasks predicted the difference in response time to the long vs. the short version of each word, after subtracting the raw difference in duration between the two versions. All errors were excluded at this point (5%). The difference in response time ranged from (−)2528 to 1348 ms ($M=-4$, $SE=316$). The data from the sentence reading task revealed that many of the participants pronounced the word *audit* /odit/ ('audit') as /odi/, thus rendering it a non-neighbor of *otite* ('otitis'). We therefore excluded the responses for *audit* and its control word, *taudis* ('slum'), for all these subjects ($N=21$). We also excluded responses that were more than 2 SDs from the mean (11%). Overall, response times to the target and control words were similar (long version: $t<1$; short version: $t<1$), as were response times to the short and long versions of the words ($t<1$; Target: long: $M=1103$, $SE=291$, short: $M=1082$, $SE=283$; Control: long: $M=1126$, $SE=286$, short: $M=1104$, $SE=298$). These data suggest that target and control words were well matched, and that the shorter and longer versions did not differ in naturalness. At the same time, the variability in participants' responses within each word type was large. These items are therefore suitable for testing how individual differences in inhibition influence the advantage for the short vs. long version in words with and without a voicing neighbor.

2.4. The relation between performance on the Stroop task and pre-voicing advantage in perception

We tested whether performance on the Stroop task was related to facilitation of the shorter pre-voicing values in word recognition with a mixed model with Participants and Items as random variables, and Type (target, control), Stroop Performance, Order, and their

---

[1] The Stroop task was followed by a Retrieval-Induced Inhibition task (Anderson, Bjork, & Bjork, 1994) not reported in here. This task measures inhibition used during lexical access. Due to an error in the script, half of the participants were only exposed to some of the words to be memorized, rendering their inhibition score less reliable. When statistical analyses are run on the full data set, no relation between inhibition score and either production or perception was found. It is unclear, however, whether the type of inhibition that this task measures is not related to pre-voicing advantage in perception or whether the unreliable version of the task that half of the participants performed prevented us from detecting this relation.
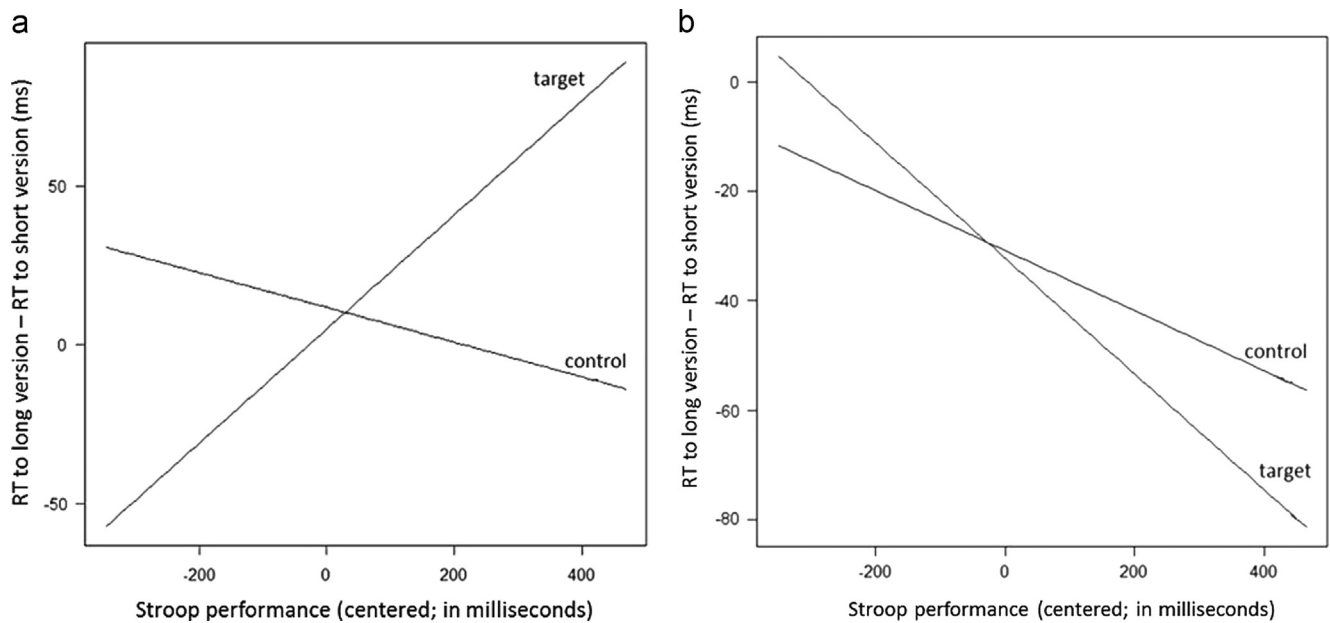
a

b

Fig. 2. Advantage for shorter pre-voicing as dependent on performance in the Stroop task and word type. Higher values in the Stroop task indicate poorer inhibition. Panel (a) shows cases in which the version with the longer pre-voicing was presented first. Panel (b) shows the cases in which the version with the shorter pre-voicing was presented first.

interaction as fixed effects. Order refers to whether participants first heard the shorter or longer version of the word. We included this variable and its interaction with the other variables, because the first presentation of the word could prime and thus influence response to its second presentation (Forster & Davis, 1984). Such priming could facilitate response to the second presentation enough to wash out any effects, and it might be asymmetrical depending on the relative ease of processing each version. The model included intercepts for the random variables as well as a slope for Order for the Participants variable, and a slope for Stroop performance for the Items variable.[2] Results revealed an effect of Order ($\beta=-42$, $SE=17.4$, $t=-2.41$), such that participants showed less advantage for the shorter version if they encountered it first. In other words, hearing the shorter version first facilitated response to the longer version, thus minimizing the difference in response time between the two versions. Results also showed the anticipated Type x Stroop Performance interaction ($\beta=0.23$, $SE=0.1$, $t=2.22$), but it was moderated by a triple Type $\times$ Stroop Performance $\times$ Order interaction ($\beta=-0.27$, $SE=0.12$, $t=-2.15$; see Table B1). We therefore ran separate follow-up analyses for the cases where the shorter version was presented first and those where the longer version was presented first.

We ran a mixed model analysis with Participants and Items as random variables, and Type, Stroop Performance and their interaction as fixed effects for the cases where the longer version was presented first. The model included a slope for Stroop Performance for the Items variable. The analysis revealed the predicted Type $\times$ Stroop Performance interaction ($\beta=0.23$, $SE=0.11$, $t=2.07$). The interaction shows that the worse participants performed on the Stroop task, the more they preferred the shorter pre-voicing values for the target words ($\beta=0.2$, $SE=0.06$, $t=3.26$), but not for the control words ($\beta=-0.04$, $SE=0.06$, $t=-0.62$; see Fig. 2 and Table B2).

Running the same mixed model analysis for the cases where the short version was presented first did not reveal any effects or interactions (all t's<1), potentially because of the facilitation of the longer version, which reduced the advantage for the shorter version.

To sum up, inhibition, as measured on the Stroop task, influences advantage for pre-voicing values in perception, though sufficient priming of a word can override its influence.

2.5. Relation between performance on the Simon task and pre-voicing advantage in perception

We next tested whether performance on the Simon task predicted advantage for shorter pre-voicing in word perception. As the performance of one participant in the Simon task was more than 2 SDs away from the performance of all other participants, that participant was excluded from the following analyses.

To test for the role of performance on the Simon task, we ran a mixed model analysis with Participants and Items as random variables and Simon Performance, Type, Order and their interaction as fixed effects. The model included a random slope for Simon Performance for the Items variable. The analysis revealed the predicted Type x Simon Performance interaction ($\beta=1.76$, $SE=0.65$, $t=2.71$), but it was moderated by a Type $\times$ Simon Performance $\times$ Order interaction ($\beta=-2.27$, $SE=0.8$, $t=-2.84$; see Table B3 for the full results). We therefore ran separate follow-up analyses for the cases where the shorter version was presented first and those where the longer version was presented first.

---

[2] In all analyses, we included all slopes that contributed to the significance of the model, even if only marginally, as determined by a Likelihood Ratio Test (Baayen, 2008, p. 275).
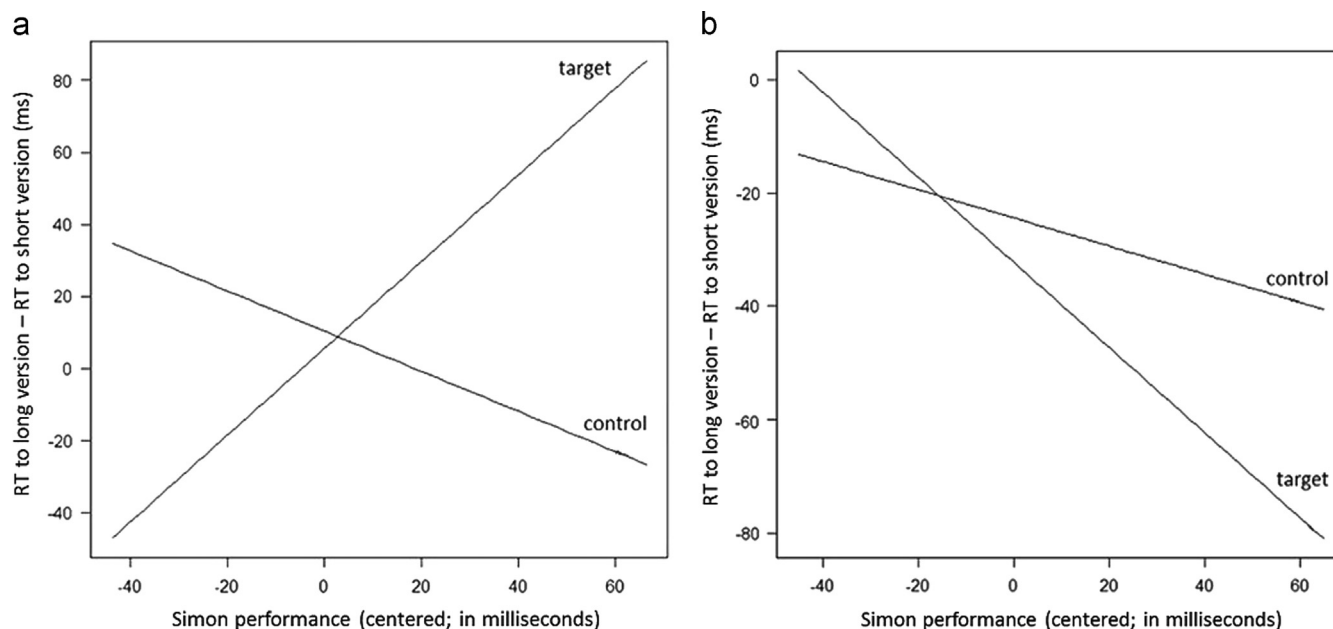
Fig. 3. Advantage for shorter pre-voicing as dependent on performance in the Simon task and word type. Higher values in the Simon task indicate poorer inhibition. Panel (a) shows cases in which the version with the longer pre-voicing was presented first. Panel (b) shows the cases in which the version with the shorter pre-voicing was presented first.

A mixed model analysis for the cases where the long version was presented first revealed the predicted Type x Simon Performance interaction ($\beta=1.76$, $SE=0.57$, $t=3.08$; see Table B4 for the full results), reflecting the fact that poorer performance on the Simon task predicted greater advantage for shorter pre-voicing for the target words ($\beta=1.21$, $SE=0.4$, $t=3.06$), but not for the control words ($\beta=-0.56$, $SE=0.41$, $t=-1.36$; see Fig. 3).

A mixed model analysis for the cases where the short version was presented first did not reveal any effects or interactions (all t's<1). As before, this might be due to priming from the first presentation of the word.

To sum up, non-linguistic inhibition, as measured on the Simon task, influences advantage for pre-voicing values in perception, though sufficient priming of a word can override its influence.

2.5.1. Production

To test whether inhibitory skill influences production of pre-voicing in words that have a voiceless neighbor, we tested whether performance on the inhibition tasks predict pre-voicing values in the sentence reading task. As for perception, we excluded the item *audit* and its matched control, *taudis,* for participants who produced the former as /odi/. Pre-voicing values ranged from 0 to 122 ms ($M=50$, $SD=17$). Mixed effects model analyses did not reveal any effects or interactions for any of the inhibition measures (all t's<2). Additionally, there was no overall difference in pre-voicing values for words with and without a neighbor ($t<1$).

2.5.2. Relation between perception and production

Next, we examined whether shorter pre-voicing values in production relate to advantage for shorter pre-voicing values in perception. A mixed model analysis over the difference in RTs to the long and short version of the word, with Participants and Items as random variables and Produced Pre-voicing, Type, Order and their interaction as fixed variables, revealed an effect of Type ($\beta=134.36$, $SE=54.96$, $t=2.44$), indicating that the advantage for shorter pre-voicing was stronger for words with a voiceless neighbor. The analysis also revealed a Produced Pre-voicing$\times$Type interaction ($\beta=-2.95$, $SE=1.06$, $t=-2.8$), a Produced Pre-voicing$\times$Order interaction ($\beta=-2.48$, $SE=1.03$, $t=-2.41$), a Type$\times$Order interaction ($\beta=-218.75$, $SE=77.18$, $t=-2.83$), and a triple Produced Pre-voicing$\times$Type$\times$Order interaction ($\beta=4.66$, $SE=1.48$, $t=3.16$; see Table B5 for the full results).

Follow-up analyses showed that when the long version was presented first, there was a main effect of Type ($\beta=137.59$, $SE=52.99$, $t=2.6$) which was moderated by a Produced Pre-voicing$\times$Type interaction ($\beta=-3.04$, $SE=1.02$, $t=-2.98$; see Table B6 for the full results). This interaction reflects the fact that shorter Produced Pre-voicing predicted greater advantage for the shorter version for the target words ($\beta=-1.97$, $SE=0.74$, $t=-2.66$) but not for the control words ($\beta=1.22$, $SE=0.71$, $t=1.72$). No relation between production and perception was found when the shorter version was presented first. Together with the results revealing a relation between inhibition and perception when the long version of the word was presented first but not when the shorter version was presented first, it seems likely that the difference in response times to the long and short versions is not indicative of one's representation when the short version is presented first.

3. Discussion

This study examines whether individual differences in a cognitive skill can influence which phonetic values are processed most easily and produced most frequently. Specifically, we tested whether lower inhibitory skill can lead to an advantage for shorter compared to longer pre-voicing values in the perception and production of words that have a voiceless neighbor. Additionally, we tested the role of both linguistic and non-linguistic inhibition in this advantage.

Results suggest that inhibitory skill can indeed influence perception advantage, as individuals with poorer inhibitory skill responded faster to words with shorter pre-voicing, and, importantly, only when those words had a voiceless neighbor. However, this relation between inhibitory skill and perception advantage only held for the cases when the version with the longer pre-voicing was presented first. This suggests that the influence of inhibitory skill on perception is not always strong enough to overcome the effect of long-term identity priming. It should be noted that, numerically, response times to the longer version were facilitated by the previous exposure to the short version ($M = 1122$ vs. 1110 ms for first and second encounter, respectively; difference $= 12$ ms) whereas responses to the short version were not facilitated by exposure to the long version ($M = 1092$ vs. 1096 ms for first and second encounter, respectively; difference $= -4$ ms), but this difference in priming was not significant.[3] Nevertheless, it seems that the presentation of the shorter version might have led to more long-term priming, which invalidated the representativeness of the response times to the words when encountered again. Additional evidence that response times in the lexical decision task were less indicative of participants' representation when the short version was encountered first comes from the finding that production was related to perception when the long version was encountered first, but not when the short version was encountered first. More research is needed to investigate why the effect of priming was asymmetrical.

The importance of this study, however, goes beyond pointing to the role of inhibition in perception of pre-voicing in words. By showing that individual differences influence the perception of this feature, it opens the door for future research on how individual differences in cognitive skills that play a role in processing could influence perception of phonetic features in general. That is, individual differences in inhibition, priming, information integration and so forth can influence the preferred variants of different phonetic features. Consequently, individual differences in cognitive skills that govern activation patterns can even influence to what degree different features are perceived to be similar to one another.

A couple of previous studies have tried to simulate how variability along different skills could account for differences in performance between typically developing and language-impaired individuals in a visual world paradigm that examines activation of a target and its cohort and rhyme competitors during language processing (McMurray, Samelson, Lee, & Tomblin, 2010; Mirman, Yee, Blumstein, & Magnuson, 2011). McMurray et al. (2010) have found in their simulations that properties such as general slowing and differences in lexical and phonemic decay accounted better (though not fully) for the group differences than inhibition did. Mirman and colleagues, who compared typically developing individuals to individuals with lesions in Broca and Wernicke areas, found that differences in lexical inhibition of competitors can explain the greater cohort competition among patients with lesion in Wernicke's area but not the greater rhyme competition among patients with lesions in Broca's area, while differences in response selection can account for both patterns. Importantly, the difference between typically developing and language impaired individuals is likely not (only) inhibition. In fact, the results of the typical and language-impaired groups in McMurray and colleagues differed more in the activation of the target than in the activation of the competitors, and Mirman and colleagues found that the activation of the rhyme and cohort competitors are negatively correlated among aphasic but not among typically-developing individuals. Therefore, those simulations cannot speak to the role of individual differences in inhibition in language processing in linguistically-unimpaired individuals, though they point in a promising direction that research can take to study individual differences.

This study failed to find an effect of individual differences in inhibition on production of pre-voicing. This might indicate that inhibition does not play the same role in production and perception. Indeed, some speech production studies show that the existence of neighbors facilitates production, which indicates neighbors do not necessarily compete during word production (e.g., Meyer & Schriefers, 1991; Vitevich, 2002). It might be the case that neighbors compete only in certain circumstances, depending on feature position in the word, the words' frequency, the feature's relative articulation difficulty etc., and therefore, that only in those circumstances, if at all, individual differences in inhibition play a part. At this point, it is worth mentioning a line of research that examines how the existence of neighbors influences phonetic articulation. Specifically, it has been found that words from denser neighborhoods or whose frequency relative to their neighborhood is low are more likely to be hyperarticulated than words from lower neighborhoods or whose frequency relative to their neighborhood is high (e.g., Scarborough, 2013; Wright, 2003). Furthermore, some evidence shows that the VOT of voiceless stops that have a voicing neighbor is longer than that of words that do not have such a neighbor (Baese-Berk & Goldrick, 2009), but this pattern is sometimes restricted to only one voicing member or to certain locations (Goldrick, Vaughn, & Murphy, 2013). The source of these differences in articulation, however, is debated, as some attribute it to accommodation to the listener (e.g., Wright, 2003) while others attribute it to the influence of competition on lexical access by the speaker (e.g., Baese-Berk & Goldrick, 2009). Our findings might shed some light on the source, since, despite not finding the existence of a neighbor to influence the duration of pre-voicing in production, we did find some relation between production and perception. The shorter the pre-voicing that was produced in words with a voicing neighbor, the greater the advantage was for the shorter pre-voiced version, in cases where the longer version was presented first. This was not the case for words with no voicing

---

[3] Note also that the shorter versions are 27.5 ms shorter, and therefore, the fact that response time to them is 30 ms shorter does not indicate that that they were easier to recognize in general.

neighbor, suggesting that pre-voicing varies more randomly in production of such words. This is in line with other research that shows that sensitivity to a phonetic feature is stronger when the feature is contrastive in the specific lexical item (Law et al., 2013). These findings are more in line with accounts that assume that hyper-articulated production is at least in part driven by speakers' ease of lexical access, although speakers might also assume their listeners' ease of access to be similar to their own, and adjust their production accordingly. Importantly, however, the relation between production and perception that we found suggests that low inhibitory skill leads to an advantage for shorter pre-voicing for words which have a neighbor - the opposite direction than that usually assumed (exaggeration of differences). Inhibitory skill, however, did not directly relate to production, potentially because other factors masked the influence of inhibition on production. For example, we used a single quasi-random order. Possibly, order-related effects, such as the identity of preceding sentences or level of fatigue or interest, influenced clarity or quality of production, thus masking the effect. Further studies are therefore needed in order to determine when and how individual differences in inhibition relate to production.

Another issue that this study investigated is whether the type of inhibition that is used for inhibiting competitors is domain-general or linguistic in nature. Results showed that performance on a linguistic (Stroop) and a non-linguistic (Simon) task predicted performance in the same manner, even though the correlation of the two tasks was only marginal. This finding suggests that the component that is used for inhibiting linguistic items is not domain-specific, and that inhibition in speech perception relies on domain-general mechanisms, at least partially. As recent evidence suggests that monolinguals and bilinguals rely on different inhibitory mechanisms during both linguistic and non-linguistic conflict tasks, it will be interesting to examine whether the type of inhibition that relates to perception advantage is the same type as the one that is used for inhibition at other linguistic levels, as well as whether it is used by bilinguals too. On the same note, this study found a relation between phonological representation and performance on the Simon and Stroop tasks, which measure top-down inhibition. Some argue that automatic bottom-up inhibition plays a role in language processing as well (e.g., Colzato et al., 2008). Therefore, it will be interesting to examine whether such inhibition also plays a role in the perception of pre-voicing or other phonetic features, and in what way the role of the two types of inhibition differs.

This study, then, shows that speakers' perception and, possibly, production might be influenced by cognitive skills in general, and level of inhibitory skill in particular. It thus contributes to our understanding of variation in production and perception, and opens a new window into research on how individual differences in cognitive skills that play a role in speech production and perception can influence linguistic representations.

Appendix A

Target words	Neighbors	Control words
aigu	écu	groggy
audit	otite	taudis
badauds	bateau	bandeau
cabot	capot	sabot
codé	coté	brodé
débit	dépit	fourbi
dégât	déca (=decaf)	yoga
endetté	entêté	modéré
égard	écart	bagarre
ego	écho	argot
égoutter	écouter	dégoûter
embaucher	empocher	saboter
frauder	frotter	broder
logo	locaux	mégot
odeur	auteur/hauteur	grandeur
pagaie	paquet	muguet
pédale	pétales	bandage
redire	retire	raidir
regain	requin	béguin
solidaire	solitaire	validé

Appendix B. Statistical tables

See Tables B1–B6.

Table B1
Table of results for the analysis of the role of the Stroop task.
Formula: Facilitation~(1+Order|Participants)+(1+Stroop score|Items)+Order*Type*Stroop score.

Factor	β	SE	t-Value
Intercept	11.36	12.83	0.885
Type	−6.15	17.90	−0.34
Stroop score	−0.05	0.07	−0.63
Order	−42.01	17.40	−2.41[*]
Type × Stroop score	0.23	0.10	2.22[*]
Type × Order	5.08	23.04	0.22
Stroop score × Order	−0.03	0.09	−0.3
Type × Stroop score × Order	−0.27	0.12	−2.15[*]

Table B2
Table of results for the analysis of the role of the Stroop task for the cases in which the long version was presented first.
Formula: Facilitation~(1|Participants)+(1+Stroop score|Items)+Type*Stroop score.

Factor	β	SE	t-Value
Intercept	11.70	13.83	0.85
Type	−6.72	19.26	−0.35
Stroop score	−0.05	0.08	−0.67
Type × Stroop score	0.23	0.11	2.07[*]

Table B3
Table of results for the analysis of the role of the Simon task.
Formula: Facilitation~(1|Participants)+(1+Simon score|Items)+Order*Type*Simon score.

Factor	β	SE	t-Value
Intercept	6.71	13.09	0.51
Type	−3.98	18.23	−0.22
Simon score	−0. 53	0.47	−1.12
Order	−28.77	17.18	−1.675
Type × Simon score	1.76	0.65	2.71[*]
Type × Order	1.39	24.14	−0.06
Simon score × Order	0.32	0.57	0.55
Type × Simon score × Order	−2.27	0.80	−2.84[*]

Table B4
Table of results for the analysis of the role of the Simon task for the cases in which the long version was presented first.
Formula: Facilitation~(1|Participants)+(1|Items)+Type*Simon score.

Factor	β	SE	t-Value
Intercept	10.41	13.80	0.755
Type	−4.67	19.25	−0.24
Simon score	−0.56	0.42	−1.33
Type × Simon score	1.76	0.57	3.07[*]

Table B5
Table of results for the analysis of the relation between perception and production.
Formula: Facilitation~(1|Participants)+(1|Items)+Produced pre-voicing*Type*Order.

Factor	β	SE	t-Value
Intercept	−13.37	38.25	−0.35
Produced pre-voicing	1.08	0.75	1.45
Type	134.36	54.96	2.445[*]
Order	79.78	52.92	1.51
Produced pre-voicing × Type	−2.95	1.06	−2.80[*]
Produced pre-voicing × Order	−2.48	1.03	−2.41[*]
Type × Order	−218.75	77.18	−2.83[*]
Produced pre-voicing × Type × Order	4.66	1.48	3.16[*]

Table B6

Table of results for the analysis of the relationship between perception and production for the cases in which the long version was presented first.

Formula: Facilitation~(1│Participants)+(1│Items)+Type☰Produced Pre-voicing.

Factor	β	SE	t-Value
Intercept	−12.00	37.14	−0.32
Produced pre-voicing	1.08	0.72	1.50
Type	137.59	52.99	2.60[*]
Produced pre-voicing × Type	−3.04	1.02	−2.98[*]

References

Anderson, M C, Bjork, R A, & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *20*(5), 1063–1087.

Andruski, Blumstein, & Burton (1994). The effect of subphonetic differences on lexical access. *Cognition*, *52*, 163–187.

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

Baese-Berk, M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes*, *24*, 527–554.

Bialystok, E., Craik, F. I.M, Grady, C., Chau, W., Ishii, R., Gunji, A., et al. (2005). Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage*, *24*, 40–49.

Blumenfeld, H. K., & Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition*, *118*(2), 245–257.

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.

Colombo, L. (1986). Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance*, *12*(2), 226–234.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S. T., La Heij, W., et al. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 302–312.

Craft, J. L., & Simon, J. R. (1970). Processing symbolic information from a visual display: Interference from an irrelevant directional cue. *Journal of Experimental Psychology*, *83*, 415–420.

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

Federmeier, K. D., & Kutas., M. (2005). Aging in context: Age-related changes in context use during language comprehension. *Psychophysiology*, *42*, 133–142.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *10*, 680–698.

Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodriguez-Pujadas, A., Belloch, V., et al. (2010). Bridging language and attention: Brain basis of the impact of bilingualism on cognitive control. *NeuroImage*, *53*, 1272–1278.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning. A distributed model of speech perception. *Language and Cognitive Processes*, *12*(5/6), 613–656.

Goldrick, M., Vaughn, C., & Murphy, A. (2013). The effects of lexical neighbors on stop consonant articulation. *Journal of the Acoustical Society of America*, *134*, EL172–EL177.

Gollan, T. H., Sandoval, T., & Salmon, D. P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science.*, *22*, 1155–1165.

Green, D. W. (1998). Mental control of the bilingual lexicosemantic system. *Bilingualism: Language and Cognition*, *1*, 67–81.

Kong, E., & Edwards, J. (2011) Individual differences in speech perception: Evidence from visual analogue scaling and eye-tracking. In *Proceedings of the XVII international congress of phonetic sciences* (pp. 1126–1129), 17–21 August 2011, Hong Kong: ICPhS.

Law, S. P., Fung, R., & Kung, C. (2013). An ERP study of good production vis-a-vis poor perception of tones in Cantonese: Implications for top-down speech processing. *PLOS ONE*, *8*(1), e54396.

Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics*, *41*, 320–331.

Linck, J. A., Schwieter, J. W., & Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition*, *15*, 651–662.

Lisker, L. (1977) Rapid vs. Rabid: A catalogue of acoustic features that may cue the distinction. *Paper presented at the 94th meeting of the Acoustical Society of America*.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384–422.

Logan, G. D., Zbrodoff, N. J., & Williamson, J. (1984). Strategies in the color-word Stroop task. *Bulletin of the Psychonomic Society*, *22*, 135–138.

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

Luk, Anderson, Craik, Grady, & Bialystok (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and Cognition*, *74*, 347–357.

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

McMurray, B., Samelson, V., Lee, S., & Tomblin, J. B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology*, *60*(1), 1–39.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*, B33–B42.

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 1146–1160.

Meyer, D. E., & Gordon, P. C. (1985). Speech production: Motor programming of phonetic features. *Journal ofMemory & Language*, *24*, 3–26.

Mirman, D., Yee, E., Blumstein, S. E., & Magnuson, J. S. (2011). Theories of spoken word recognition deficits in aphasia: Evidence from eye-tracking and computational modeling. *Brain and Language*, *117*(2), 53–68.

New, B., Pallier, C., Ferrand, L., & Matos, R (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique*, *101*, 447–462. ⟨http://www.lexique.org⟩.

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

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological review*, *115*(2), 357.

Peterson, G., & Barney, H. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, *24*, 175–184.

Repp, B. (1981). Two strategies in fircative discrimination. *Perception & Psychophysics*, *30*, 217–227.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*, 249–284.

Sadat, J., Martin, C. D., Costa, A., & Alario, F. (2014). Reconciling phonological neighborhood effects in speech production through single trial analysis. *Cognitive Psychology*, *68*, 33–58.

Scarborough, R. (2013). Neighborhood-conditioned patterns in phonetic detail: Relating coarticulation and hyperarticulation. *Journal of Phonetics*, *41*(6), 491–508.

Shockley, K., Sabadini, L., & Fowler, C. A. (2004). Imitation in shadowing words. *Perception and Psychophys*, *66*, 422–429.

Stemberger, J. P. (1985). *The lexicon in a model of language production*. New York: Garland.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*(3), 281–284.

Stewart, M. E., & Ota, M. (2008). Lexical effects on speech perception in individuals with "autistic" traits. *Cognition*, *109*(1), 157–162.

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. *Journals of Gerontology: Series B*, *65*, 551–560.

Vitevich, N. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *28*(7), 735–747.

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

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

Wright, R. (2003). Factors of lexical competition in vowel articulation. In J. Local, R. Ogden, & R. Temple (Eds.), *Phonetic interpretation: Papers in laboratory phonology VI*. Cambridge: Cambridge University Press (pp. 75–87).

Yu, A. C. L. (2010). Perceptual compensation is correlated with individuals' "autistic" traits: Implications for models of sound change (8). *PLoS ONE*, *5*, e11950.