

Washington University in St. Louis

Washington University Open Scholarship

Arts & Sciences Electronic Theses and
Dissertations

Arts & Sciences

Summer 8-2021

Simultaneous Bilinguals' Comprehension of Accented Speech

Sita Carraturo

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds



Part of the [Psycholinguistics and Neurolinguistics Commons](#)

Recommended Citation

Carraturo, Sita, "Simultaneous Bilinguals' Comprehension of Accented Speech" (2021). *Arts & Sciences Electronic Theses and Dissertations*. 2554.

https://openscholarship.wustl.edu/art_sci_etds/2554

This Thesis is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS
Department of Psychological & Brain Sciences

Simultaneous Bilinguals' Comprehension of Accented Speech
by
Sita Carraturo

A thesis presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Master of Arts

August 2021
St. Louis, Missouri

© 2021, Sita Carraturo

Table of Contents

| | |
|--|-----|
| List of Figures..... | iii |
| Acknowledgments | iv |
| Abstract of the Thesis | v |
| Chapter 1: Introduction & Background..... | 1 |
| Chapter 2: Experiment 1 | 7 |
| 2.1 Participants | 7 |
| 2.2 Stimuli..... | 8 |
| 2.3 Results | 8 |
| Chapter 3: Experiment 2..... | 10 |
| 3.1 Participants | 10 |
| 3.2 Stimuli | 11 |
| 3.3 Procedure | 12 |
| 3.4 Results | 12 |
| Chapter 4: Experiment 3..... | 15 |
| 4.1 Participants | 16 |
| 4.2 Stimuli | 16 |
| 4.3 Results | 17 |
| Chapter 5: Discussion & Future Directions..... | 18 |
| 5.1 Future Directions | 20 |
| References | 22 |

List of Figures

| | |
|---------------|----|
| Figure 1..... | 27 |
| Figure 2..... | 32 |
| Figure 3..... | 36 |

Acknowledgments

I would like to thank my committee members: Kristin Van Engen, Mitchell Sommers, and Joe Barcroft. I would also like to thank my colleagues Drew McLaughlin and Violet Brown.

Sita Carraturo

Washington University in St. Louis

August 2021

ABSTRACT OF THE THESIS
Simultaneous Bilinguals' Comprehension of Accented Speech

by

Sita Carraturo

Master of Arts in Psychological & Brain Sciences

Washington University in St. Louis, 2021

Professor Kristin Van Engen, Chair

Professor Mitchell Sommers

Professor Joe Barcroft

L2-accented speech recognition has typically been studied with monolingual listeners or late L2-learners, but simultaneous bilinguals may have a different experience: their two phonologies offer flexibility in phonological-lexical mapping (Samuel and Larraza, 2015), which may be advantageous. On the other hand, the two languages cause greater lexical competition (Marian & Spivey, 2003), which may impede successful L2-accented speech recognition. The competition between a bilinguals' two languages is the oft-cited explanation, for example, as to why bilinguals underperform monolinguals in native-accented speech-in-noise tasks (Rogers et al., 2006).

To investigate the effect of bilingualism on L2-accented speech recognition, the current studies compare monolingual and simultaneous bilingual listeners in three separate experiments. In the first study, both groups repeated sentences produced by speakers of Mandarin-accented English whose English proficiencies varied. In the second study, the stimuli were presented in varying levels and types of noise, and a native-accented speaker was included. In each of these

first two studies, the sentences were semantically anomalous (i.e., nonsensical). In the third study, the stimuli were meaningful sentences, presented in a single noise condition, and spoken by either a native speaker or an L2-accented speaker.

Mixed effects models revealed differences in L2-accented speech recognition measures driven by listeners' language backgrounds only in Experiments 2 and 3; in Experiment 1, performance between groups was statistically identical. Results in Experiments 2 and 3 also replicated the prior finding that bilinguals perform worse for native-accented speech in noise.

We propose that neither a flexible phonological-lexical mapping system nor increased lexical competition can alone sufficiently explain the deficit (relative to monolinguals) that simultaneous bilinguals exhibit when faced with L2-accented speech in real-world listening conditions. We discuss the possible implications of processing capacity and cognitive load, and suggest that these two factors are more likely to contribute to experimental outcomes. Future studies with pupillometry to explore these hypotheses are also discussed.

Chapter 1: Introduction & Background

A speech signal's acoustic quality can be affected by speaker-extrinsic factors, such as noise in the environment, and speaker-intrinsic factors (i.e., characteristics about the speaker that affect speech output), such as speech pathologies, speaking rates, or accents. These speaker-extrinsic and speaker-intrinsic factors manifest frequently during communication, and as such, an abundance of research has been devoted to understanding how each affects speech comprehension.

Regarding speech-extrinsic factors, the signal-to-noise ratio (SNR) is a crucial predictor of the impact of environmental noise, such that the louder the background noise is relative to the target signal, the poorer the comprehension of the target signal will be (e.g., Adank et al., 2009; Shi, 2010; Tabri et al., 2011). This is attributed to energetic masking, wherein the noise obscures parts of the speech signal that are necessary for successful comprehension. Similarly, the presence of speech in the background (referred to as *babble*) also interferes with comprehension (e.g., Van Engen & Bradlow, 2007). In the case of babble, it is not just energetic masking at play, but also informational masking: listeners are susceptible to having their attention diverted away from the target stimulus and reallocated to comprehending the background babble (Kim, 2020; Brouwer, Van Engen, Calandruccio, & Bradlow, 2012; Cooke, Garcia Lecumberri, & Barker, 2008).

As for speaker-intrinsic factors, the focus of the present study is speaker accent. While *degradation* is a term that aptly describes the effects of background noise on a speech signal, *deviation* and *variation* are more appropriate terms for describing the comprehension challenges

presented by the accent of a person speaking a language other than their native language. That is, an accented speech signal may be completely unobscured acoustically, but different from the variety spoken by the listener. Given that everyone has an accent, we would like to clarify the terminology used in this paper. “L2-accent” will be used to refer to speech that is (1) accented as a result of the speaker’s relatively late acquisition of a second language and (2) an accent that is not produced by the listener and therefore is likely to be unfamiliar to the listener. “Native accent”, thus, will be used to refer to (1) speech that is produced by a native speaker of the language, and (2) an accent that is shared by the speaker and the listener (i.e., it is native to both parties).

Munro & Derwing (1995a) found that L2-accented speech does require more processing time than native-accented speech, but they also found that the accentedness of the speech is partially independent of its intelligibility (Munro & Derwing 1995b). That is, speech that is considered to be highly accented can also be highly—if not entirely—intelligible. Nonetheless, in a pupillometry study, McLaughlin & Van Engen (2020) demonstrated that even when L2-accented speech is as intelligible as native speech, listeners exerted more cognitive effort in understanding the L2-accented speech.

There is also a body of research that has studied the combined effects of noise and accent. Wilson & Spaulding (2010) and Rogers et al. (2004) both looked at how different SNRs could interact with L2-accented English produced by speakers of varying proficiency levels. Both studies found that speaker intelligibility was highest for native proficiency speakers, lower for high proficiency speakers, and lowest for low proficiency speakers. Crucially, both studies also found that a moderate amount of masking noise (+10 dB SNR) caused a significant decrease in the intelligibility of high proficiency speakers but did not result in a change for native

proficiency speakers. In addition to these intelligibility scores, Wilson & Spaulding (2010) looked at processing time. They found that listeners' reaction times were shortest for native proficiency speakers and longest for low proficiency speakers. In a study that compared the effects of six different maskers (one energetic, five informational) on Spanish-accented English, Gordon-Salant et al. (2013) found that listeners' ability to cope with maskers was dependent on the accentedness of the talker ("native", "mildly accented", or "moderately accented"): masking was most difficult for the moderately accented talker.

These findings indicate that, while they interact with one another, background noise and accent each present a distinct type of comprehension challenge. In fact, Ferguson et al. (2010) found that older adults (who often perform worse than their younger counterparts on difficult listening tasks) were not affected by foreign accent any more than younger adults were. As a possible explanation, Mattys et al. (2012) points out that unlike background noise, accent is patterned: the phonological variation is (almost by definition) consistent and predictable. This predictability is likely what allows listeners to adapt to L2-accented speech. Clarke & Garrett (2004) demonstrated that listeners can adapt to a foreign accent within a single minute of exposure to it, and Brown et al. (2020) used pupillometry to demonstrate that this is the case even for fully intelligible L2-accented speech.

In the current study, we are interested in understanding the effects of background noise and speaker accent in the context the listener's language background. Specifically, we are interested in the effects of bilingualism on accented speech comprehension. Though bilingual vs. monolingual performance on speech recognition can be indistinguishable in quiet listening conditions, multiple studies have found that bilinguals perform significantly worse than monolinguals when tasked with comprehending native-accented speech in noise (e.g., Bradlow

& Bent, 2002; Krizman, Bradlow, Lam & Kraus, 2016; Mayo, Florentine, & Buus, 1997; Rogers, Dalby & Nishi, 2004; Shi, 2009, 2010; Van Engen & Bradlow, 2007). For example, Krizman et al. (2016) studied a group of high school students (31 English monolinguals and 25 Spanish-English bilinguals) and found that monolinguals performed better than bilinguals in both word-in-noise and sentence-in-noise tasks, though only the latter yielded statistically significant results. This same group of bilinguals, however, performed significantly better than their monolingual counterparts in a test of tone-in-noise (non-linguistic). Though Krizman et al. did not find significant differences in performance between groups for the word-in-noise task, Tabri et al. (2011) and Rogers et al. (2006) did: monolinguals outperformed the bilinguals in single-word-in-noise tasks.

Morini & Newman (2020) suggest that bilinguals' performance in speech-in-noise tasks is rooted in how the bilingual brain stores lexical information. For one, bilinguals' linguistic experiences are distributed between their two languages so that individual lexical items tend to be weaker for the bilingual than for the monolingual (for whom every encounter with a given item will have activated the same lexical representation throughout their life, strengthening it). Schmidtke (2016), for example, found that the difference between monolingual and bilingual performance was greatest for low-frequency lexical items (i.e., words that bilinguals would have had the least exposure to). Moreover, it has been shown that bilinguals' languages are simultaneously activated, such that even during a task in which the language is explicitly specified, there is dual-activation (Marian & Spivey, 2003). In the face of this dual-activation, bilinguals need to allocate some of their cognitive resources to inhibiting the non-target language. It is likely, then, that the presence of noise disproportionately affects bilinguals because their linguistic systems are already managing the combined effects of weaker lexical

activation and the need to inhibit one of their languages (Morini & Newman, 2020; see also: cognitive capacity, e.g., Kahneman, 1973).

The literature on how bilinguals deal with L2-accented speech is relatively limited. Weber et al. (2014) tested two participant groups (monolinguals and bilinguals) and two types of accents (genuine and artificial). In this study, the investigators found that English monolinguals adapted to Italian-accented English, as did Dutch-English bilinguals (whose second language was English). The Dutch-English bilinguals were able to accommodate artificially accented versions of both their L1 (Dutch) and their L2 (English). Meanwhile, the monolinguals failed to accommodate the artificially accented version of English. The authors argue that this finding indicates that phonetic-to-lexical mapping is a more flexible process among bilinguals such that novel pronunciations of words are readily mapped to the intended stored lexical representation, while monolinguals appear to struggle with this. This conclusion has additional support from the findings of Samuel & Larraza (2015): in this study, bilingual participants were taught nonwords that matched the phonotactics of their L1. Crucially, the participants only ever heard one pronunciation of each nonword until the test trials. In the test trials, accented versions of the nonwords were presented and the participants accepted them, identifying them as the intended word despite never having heard these pronunciations before. These authors suggest that this flexibility is adaptive for bilinguals.

When studying bilinguals, it is generally best practice to specify the bilingual population of interest, as the term “bilingual” encompasses people with diverse language backgrounds that can lead to different experimental outcomes (Larraza, Samuel, & Oñederra, 2016; Luk & Bialystok, 2013; Shi, 2010, 2012). For example, Larraza et al. (2016) studied how Spanish-Basque and French-Basque bilinguals performed on phoneme discrimination tasks in order to

understand the effects of L1 and age of acquisition (the age at which the second language was acquired) on accented-speech perception. Their findings showed that age of acquisition contributes significantly to the outcomes measured, such that the simultaneous bilinguals (those who acquired both their languages in infancy) performed the best. Additional studies of foreign accent and bilinguals have focused on late L2 learners, who themselves speak with accents, and found that their accents directly affect their performance on accented-speech comprehension (Bent & Bradlow, 2003).

The current study is interested in simultaneous bilinguals. Though bilinguals and monolinguals are fundamentally different, simultaneous bilinguals have a crucial factor in common with their monolingual counterparts: both they and the monolinguals are native speakers of American English (the target language of the speech stimuli used in this study). In studying simultaneous bilinguals, we can control for the age of English acquisition across both groups (i.e., minimize the effects of English proficiency). Given that L2-accented speech is a largely phonological phenomenon, the distinction of interest between the two groups is that the simultaneous bilinguals will have had exposure to and experience with multiple phonologies across their lifetime. Thus, the comparison of these two groups' performance in the face of L2-accented speech comprehension allows us to ask whether long-term exposure to multiple phonologies (simultaneous bilingualism) affects the ability to comprehend phonologically variant speech (aka: L2-accented speech). As such, the following three experiments present L2-accented speech to monolingual and simultaneous bilingual listeners.

Chapter 2: Experiment 1

In this experiment, we started with this fundamental question: do simultaneous bilinguals and monolinguals perform differently when comprehending L2-accented speech? Based on prior literature, we had competing hypotheses: (1) due to a more flexible phonetic-lexical mapping system, bilinguals would outperform monolinguals, or (2) due to the fact that bilinguals encounter greater lexical competition, monolinguals would outperform bilinguals.

2.1 Participants

Participants in all three experiments were recruited and compensated using the online subject pool Prolific (<https://www.prolific.co/>). Compensation was set at a rate of \$10/hour. Each of the three experiments was built and deployed using Gorilla Experiment Builder (Anwyl-Irvine, Massoné, Flitton, Kirkham, & Evershed, 2019). Experimental protocols were approved by the Washington University in Saint Louis Institutional Review Board. All participants were between the ages of 18 and 35 (inclusive), had self-reported normal hearing, and were born and living in the USA. Simultaneous bilingualism was defined as having been speaking/understanding English and at least one other language since before the age of three. Bilingual participants were excluded if their non-English language was Chinese (i.e., Chinese, Mandarin, Cantonese, Teochew. etc), as the L2-accent used in our study is Mandarin.

In Experiment 1, data was collected from 145 participants. After exclusions based on demographics, Chinese language proficiency, and incomplete data, 40 bilinguals and 49 monolinguals remained. To match the groups, nine monolinguals were randomly selected for removal, leaving $N = 40$ per group.

2.2 Stimuli

Participants heard English sentences spoken by two non-native speakers of English whose first language was Mandarin Chinese. Based on data from a pilot study, one speaker was approximately 80% intelligible (high proficiency) while the other speaker was approximately 40% intelligible (low proficiency).

The sentences used in this experiment were originally created for the Syntactically Normal Sentences Test (Nye & Gaitenby, 1974), but were recorded locally by the speakers described above. These sentences are semantically anomalous (i.e., meaningless) but are syntactically legal (e.g. *The swift candle talked the sky.*); the syntactic structure does not vary across the set of sentences. The choice of semantically anomalous sentences was motivated by the goal of focusing on the effect of the speaker's phonology: the lack of semantic content reduces the impact of top-down processing, but the syntactic structure helps maintain suprasegmental and prosodic features of natural speech.

The participants were asked to type the sentences after they heard them. Intelligibility scores were based on the correct identification of each of the four content words in a given sentence. Each participant heard 98 sentences (49 from each speaker) resulting in 392 intelligibility scores per participant. Scoring was done automatically, with each keyword coded as correct or incorrect.

2.3 Results

Data were analyzed via generalized linear mixed-effects models, using the *lme4* package in R; methods of analysis and results reporting were implemented per the recommendations of

Brown (2021). Each content word of each sentence was scored independently, meaning accuracy scores are binary. Participants and items (word per given sentence) were included as random intercepts.

Model comparisons (likelihood ratio tests) were used to assess the significance of Group (bilingualism vs. monolingualism). The effect of Group ($\chi^2(1) = 0.0026, p = 0.903$) did not improve model fit, and model estimates indicated that performance was not significantly different between bilinguals and monolinguals ($\beta = 0.0168$; Figure 1).

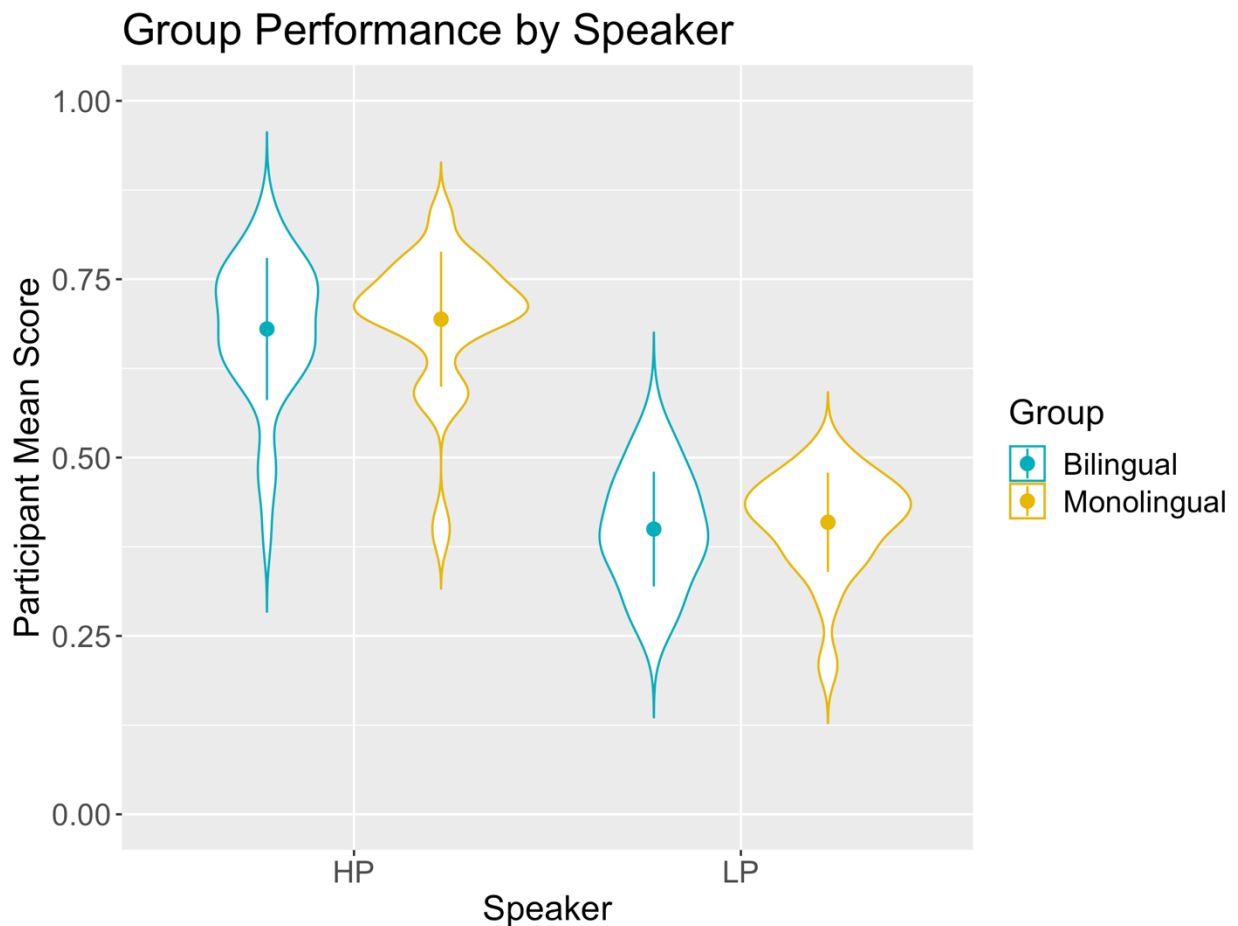


Figure 1. Group performance in Experiment 1. Violin plots show the distribution of participants' averages, per group and per speaker (High Proficiency and Low Proficiency), and points show the group means with standard deviation.

Chapter 3: Experiment 2

Based on the null results of Experiment 1, it may be the case that L2-accented speech does not affect the accuracy of simultaneous bilinguals' processing any more than it affects the accuracy of monolinguals' processing. Nonetheless, the stimuli in the previous study were heard in silence. To make the listening conditions more challenging, we masked half the stimuli with energetic noise and the other half with informational noise, each at two different noise levels.

Adding noise to the stimuli not only makes them more challenging, but arguably also makes the listening condition more realistic given that speech communication rarely ever occurs in pristine listening conditions. Furthermore, to the author's knowledge, a study of L2-accented speech in noise with bilingual listeners has not yet been published. Related studies include Rogers et al. (2004) and Gordon-Salant et al. (2013), but the former was conducted with monolingual listeners while language background information for the latter was not reported (aside from it being stated that the participants were native English speakers).

3.1 Participants

Two participant groups were recruited with the same characteristics as in Experiment 1. Data was collected from 130 participants prior to any exclusions. Of the 130 subjects who participated, nine were excluded for not meeting eligibility criteria, 22 were excluded due to some proficiency in a Chinese language, leaving 53 simultaneous bilinguals and 47 monolinguals. To keep groups even, an additional six bilinguals were randomly selected and removed prior to analysis, leaving $N = 47$ per group.

3.2 Stimuli

Participants heard English sentences spoken by six non-native speakers of English whose first language was Mandarin Chinese, and three native speakers of English. All speakers were female. Based on data from a pilot study, three of the non-native speakers were approximately 80% intelligible in quiet (high proficiency), while the other three were approximately 40% intelligible (low proficiency). The three native speakers were assumed to be approximately 100% intelligible.

The sentences used in this experiment were from the same set as those used in Experiment 1. A total of 264 target sentences were presented in this study (the original sentence set included 211 sentences, but an additional 100 were created in the lab following the syntactic and syllabic parameters of the original set).

Two types of maskers were employed in this study: speech-shaped noise (SSN) and a two-talker babble. The speakers of the two-talker babble were male speakers; differing the genders between the target and babble speakers was a deliberate decision made to provide the listener an intuitive cue for target onset (e.g., Zekveld, Rudner, Kramer, Lyzenga, & Rönnerberg, 2014). The two-talker babble was made with recordings and procedures described in Van Engen, Phelps, Smiljanic, & Chandrasekaran (2014). To create this babble track, the authors concatenated a set of sentences (originally from Bradlow & Alexander, 2007) spoken by each speaker yielding two tracks (one per speaker) of continuous speech. Those two tracks were then mixed together to form a two-talker babble track. This track was duplicated and one was normalized at 68 dB, the other at 62 dB. These two tracks were then randomly divided into 88 shorter files, which were used to mask the target sentences.

The two-talker babble track was also the audio file that was used to create the SSN masker. This was done in Praat (Boersma & Weenink, 2021) per the instructions of Listen Lab (2020). Again, the resulting track was duplicated and one was normalized at 68 and the other at 62 dB. These two tracks were then randomly divided into 88 shorter files that would mask the target sentences.

The target sentences had been normalized at 65 dB prior to masking. Half of the sentences were masked with SSN, and the other half with two-talker babble. Within each of the masking conditions, half of the target sentences were masked using the 68 dB files and the other half were masked using the 62 dB files, resulting in two SNRs: +3 dB SNR and -3 dB SNR; this yielded a total of four masking conditions.

3.3 Procedure

The three speaker proficiency groups (Low Proficiency L2-Accented, High Proficiency L2-Accented, and Native-Accented) and four masking conditions were blocked such that participants heard each of the four conditions of a given talker group before hearing a new talker group. Presentation of these blocks were counterbalanced across participants, and stimuli presentation within conditions was randomized.

3.4 Results

As in Experiment 1, data were analyzed using generalized linear mixed-effects models per the guidance of Brown (2021). Likelihood ratio tests were conducted between the full model and each of three reduced models. The reduced models were identical to the full model except

that in each one, one of the following factors was eliminated: Group, Speaker Proficiency, and SNR. These model comparisons revealed that effects of Group ($\chi^2(1) = 6.9003, p = 0.008618$), Speaker Proficiency ($\chi^2(2) = 201.22, p = 2.2 \times 10^{-16}$), and SNR ($\chi^2(1) = 26.255, p = 2.992 \times 10^{-7}$) improved model fit, while Masking Condition did not ($\chi^2(1) = 0.2094, p = 0.6473$). Model estimates indicated that monolinguals performed slightly better than bilinguals ($\beta = 0.413$).

A full interaction model with all two-way interactions was compared with each of two reduced interaction models: one without a Group by Speaker Proficiency interaction, and the other with no Group by SNR interaction. These model comparisons revealed that a Group by Speaker Proficiency interaction improved model fit ($\chi^2(2) = 54.033, p = 1.849 \times 10^{-12}$), but a Group by SNR interaction did not ($\chi^2(1) = 0.916, p = 0.3385$). The Group by Speaker Proficiency results are plotted in Figure 2.

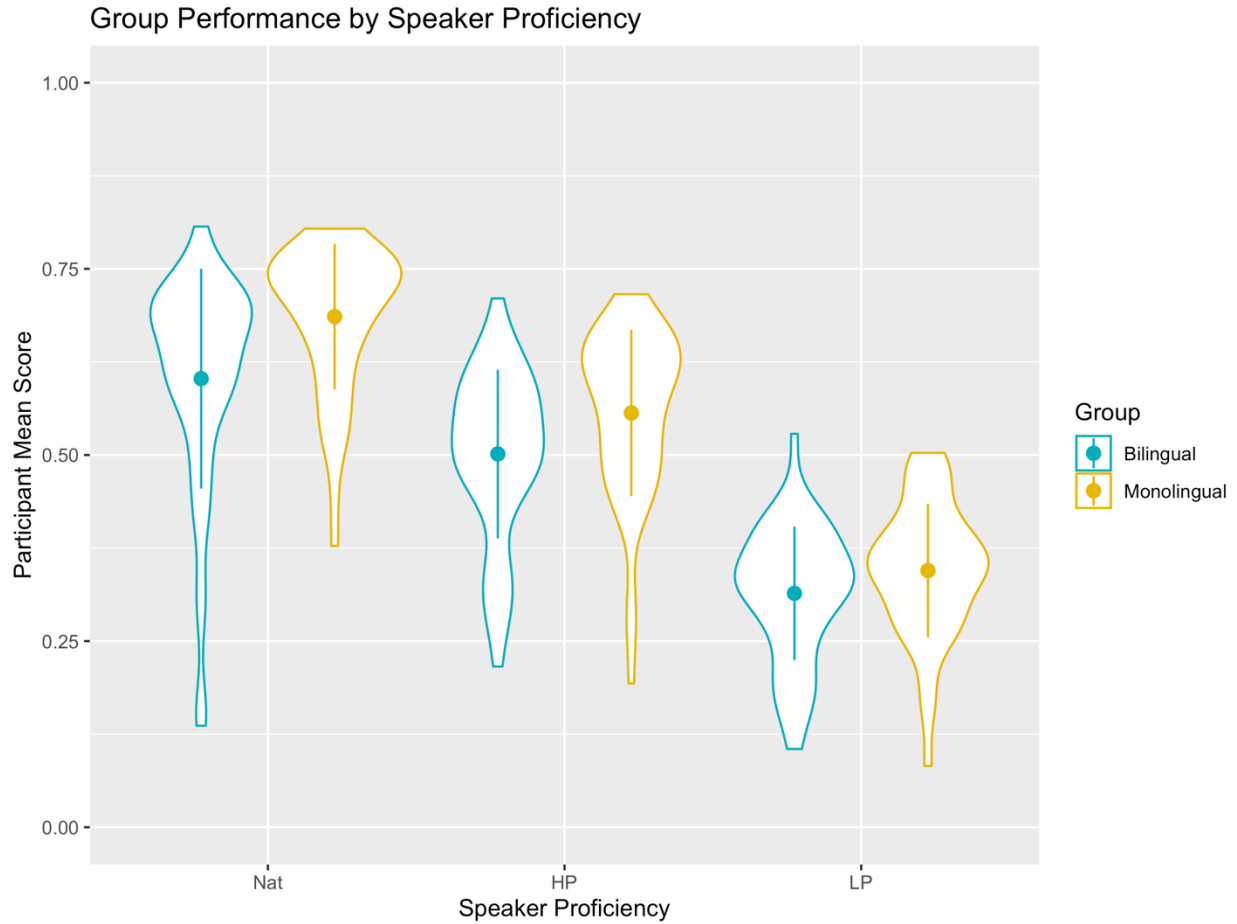


Figure 2. Group performance in Experiment 2. Violin plots show the distribution of participants' averages, per group and per speaker proficiency (Native, High Proficiency, and Low Proficiency), collapsed across masking conditions. Points show the group means with standard deviation.

The significant interaction between Group and Speaker Proficiency reveals that, as speaker proficiency increases from Low Proficiency L2-accented to Native, monolingual performance increases more than bilingual performance does. That is, as the speaker's proficiency goes down, the monolingual advantage diminishes; as seen in previous studies, monolinguals outperform bilinguals on native-accented speech in noise (e.g., Bradlow & Bent, 2002; Krizman, et al., 2016; Mayo, et al., 1997; Rogers, et al., 2004; Shi, 2009, 2010; Van Engen

& Bradlow, 2007), and yet as in our first experiment, the two groups perform similarly with L2-accented speech.

That the difference between monolingual and bilingual performance in noise decreases in L2-accented conditions brings us back to the null results of Experiment 1. Specifically, it is unclear what factors are causing these groups' performance to converge when it is well-established that the underlying processes are quite different for the two groups. In our third experiment, we manipulate the semantic content of the stimuli to assess whether the lack of semantic context in Experiments 1 and 2 could be contributing to the results by bringing monolingual performance down.

Despite an abundance of research showing that monolinguals and bilinguals process language differently, no group differences emerged in this task. We believe there are several possible interpretations for this result. It is possible that both the flexible phonetic-lexical mapping system and the greater lexical competition cancel one another out, such that the behavioral outcome appears the same for simultaneous bilinguals as for monolinguals. It could also simply be that the L2-accented speech was not difficult enough to yield group differences. This is the possibility that we explore in the next experiment.

Chapter 4: Experiment 3

There is some evidence that monolinguals make better use of context cues than their bilingual counterparts do (Skoe & Karayanidi, 2019; Shi, 2010). Given that the previous experiments in this series have used semantically anomalous stimuli, we now investigate how this lack of context might have affected group performance in Experiment 1. By comparing the

comprehension of meaningful sentences produced by a native-accented speaker to that by an L2-accented speaker (both sets in identical noise conditions), we pose the following question: Do monolinguals benefit from semantic content more than bilinguals do even in the presence of L2-accented speech in noise?

4.1 Participants

Two participant groups were recruited with the same characteristics as in Experiments 1 and 2. The participant target for this study is 50 per group, but at the time of this writing, data has been collected from 87 participants. The results reported here are based on this number, though data collection continues.

Of the 87 participants, seven were excluded for not meeting eligibility criteria, and an additional three were excluded for some proficiency in a Chinese language, leaving 49 monolinguals and 29 bilinguals. After data was scored, group and individual means were calculated. Any participant who performed three standard deviations below or above the group mean was eliminated; this eliminated one bilingual and one monolingual participant. Finally, to keep groups equal, 20 monolinguals were randomly excluded, resulting in $N = 28$ per group.

4.2 Stimuli

Participants heard English sentences spoken by 1 female native speaker of American English, and 1 female non-native speaker whose first language was Mandarin Chinese.

The sentences used in this experiment had the same sentence structure as those in the previous experiments, but these sentences made sense (e.g., “The gray mouse ate the cheese.”).

All the stimuli were masked with the SSN maskers from Experiment 2, resulting in a +3 dB and -3 dB SNRs.

4.3 Results

Again, generalized mixed effects models were employed. A model comparison between the full model and a reduced models revealed that effect of Group ($\chi^2(1) = 8.105, p = 0.004414$) and SNR ($\chi^2(1) = 62.453, p = 2.729 \times 10^{-15}$) improved model fit. Model estimates indicated that monolinguals performed slightly better than bilinguals ($\beta = 0.4799$).

A model comparison between a full interaction model and reduced interaction models revealed no significant interactions with SNR, but that a Group by Speaker interaction improved model fit ($\chi^2(1) = 25.906, p = 3.584 \times 10^{-7}$). Specifically, bilingual listener performance was more detrimentally affected by the L2 accent than was monolinguals' performance (Figure 3).

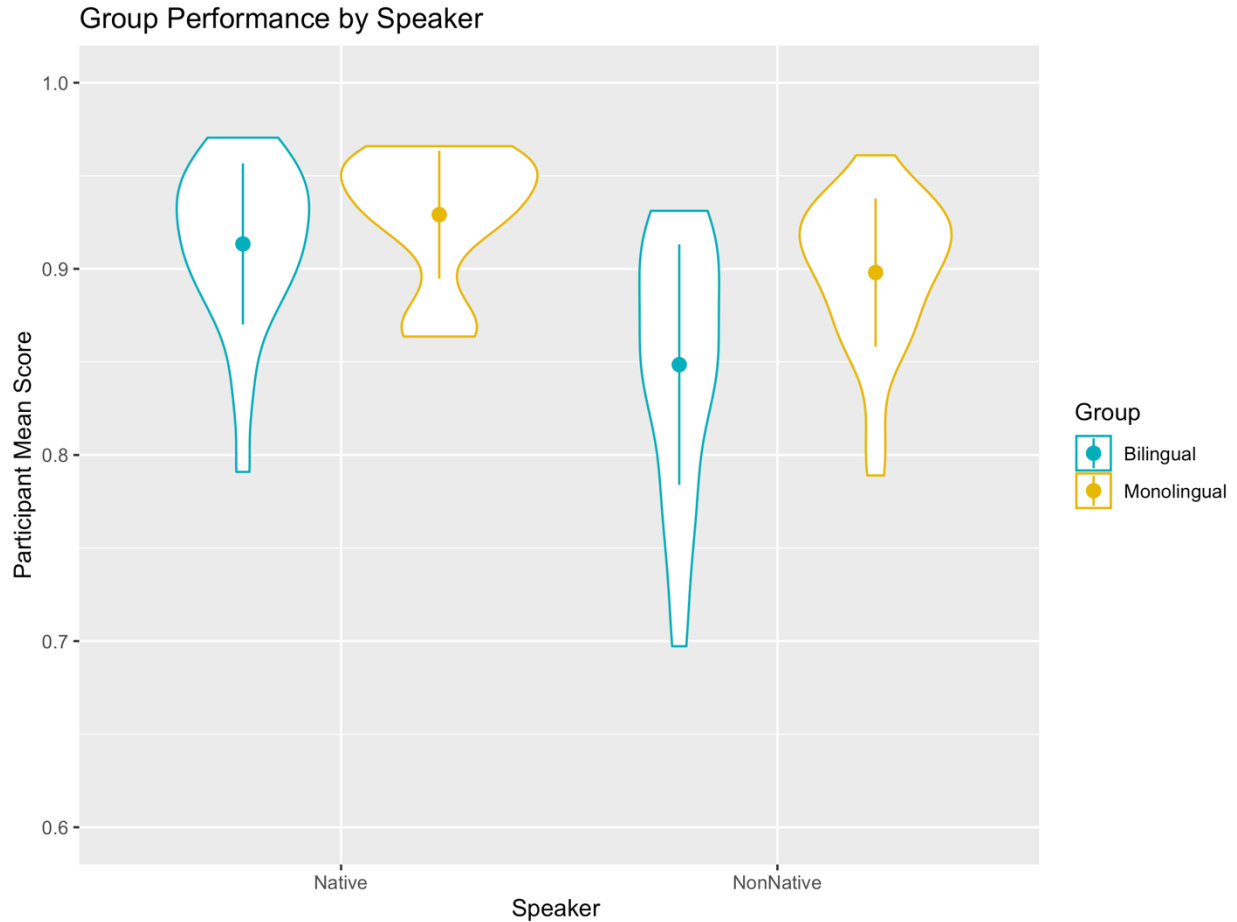


Figure 3. Group performance in Experiment 3. Violin plots show the distribution of participants' averages, per group and per speaker. Points show the group means with standard deviation.

Chapter 5: Discussion & Future Directions

We began with the question: Does a simultaneous bilingual's lifelong experience with multiple phonologies benefit them when listening to phonologically unfamiliar speech? Existing data informed competing hypotheses: bilinguals may experience greater lexical competition, causing them to perform worse than monolinguals, or bilinguals have a more flexible phonetic-to-lexical mapping system that would allow them to outperform their monolingual counterparts.

In Experiment 1, we tested our question straightforwardly: we presented simultaneous bilinguals and monolinguals with semantically anomalous sentences produced by two L2-accented speakers with distinct levels of proficiency. In the absence of any other manipulations, bilingual and monolingual performance in this task was indistinguishable from one another.

In an effort to draw out differences that may elucidate what is contributing to bilingual performance in L2-accented speech comprehension, we followed up with a second experiment in which we introduced background noise. In Experiment 2, we were able to replicate the finding that bilinguals struggle more with background noise than monolinguals do, but we also found that the benefit that monolinguals exhibit over bilinguals diminishes with L2-accented speech. Again, there was no clear interpretation as to the effect that bilingualism has on L2-accented speech comprehension.

Finally, in a third experiment we manipulated the semantic content of the sentences. According to previous findings (e.g., Skoe & Karayanidi, 2019), monolinguals benefit more from semantic content than bilinguals do. With that in mind, we presented monolinguals and bilinguals with meaningful sentences (i.e., sentences with semantic context) produced by one native-accented female speaker and one L2-accented female speaker, both in noise. Bilinguals were outperformed by monolinguals in both conditions, but the disparity was magnified in the L2-accented speech condition. This result is an interesting one: in Experiments 1 and 2, bilingual and monolingual performance *converged* in L2-accented conditions relative to native-accented conditions, while in this experiment the performance between the two groups became more disparate in the L2-accented condition. The result may indicate that the cognitive effort involved in sentence processing is a crucial factor in the difference between monolingual and bilingual listening comprehension performance. That the simultaneous bilinguals were able to leverage

context clues when native-accented speech was acoustically degraded, but unable to do so when the speech was L2-accented *and* acoustically degraded, suggests that these combined acoustic/phonetic challenges overload the processing stream in bilinguals such that the efficacy of any semantic processing is reduced. This is worth further investigation, particularly given that the presence of background noise as well as semantic context make this experiment the most comparable (of the three presented here) to real-world listening conditions.

In sum, the combined results of these three experiments suggest that lifelong experience with multiple phonologies is not sufficient to overcome other obstacles of speech comprehension. From the results of Experiments 2 and 3, it appears that a linguistic system that manages more than one language may be overburdened by each additional factor that introduces adversity to the listening environment.

5.1 Future Directions

Despite some interesting results from Experiments 2 and 3, the results of Experiment 1 remain enigmatic. This is likely due to the nature of the measurements used in these three experiments: intelligibility data alone cannot reveal more subtle differences in mental processes. Pupillometry, on the other hand, allows researchers to index pupil dilation as a psychophysiological measure of cognitive effort (see: Van Engen & McLaughlin, 2018). This measure is able to reveal differences between the effects of listening conditions that intelligibility scores alone cannot capture (e.g., Winn, Edwards, & Litovsky, 2015) As such, planned follow-up experiments include employing measurements of cognitive load (dual task, pupillometry; see: Brown et al., 2020) and studies of perceptual adaptation (e.g., Baese-Berk, McLaughlin, &

McGown, 2020) to investigate differences between the underlying mechanisms of L2-accented speech comprehension in monolinguals and simultaneous bilinguals.

References

- Adank, P., & Janse, E. (2010). Comprehension of a novel accent by young and older listeners. *Psychology and aging, 25*(3), 736.
- Adank, P., Davis, M.D., Hagoort, P. (2012) Neural dissociation in processing noise and accent in spoken language comprehension. *Neuropsychologia, 50*, 77-84.
- Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 520.
- Anwyl-Irvine, A.L., Massoné J., Flitton, A., Kirkham, N.Z., Evershed, J.K. (2019). Gorilla in our midst: an online behavioural experiment builder. *Behavior Research Methods*. Doi: <https://doi.org/10.3758/s13428-019-01237-x>
- Baese-Berk, M. M., McLaughlin, D. J., & McGowan, K. B. (2020). Perception of non-native speech. *Language and Linguistics Compass, 14*(7), e12375.
- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America, 114*(3), 1600-1610.
- Boersma, Paul & Weenink, David (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.50, retrieved 20 June 2021 from <http://www.praat.org/>
- Bradlow, Ann R., and Jennifer A. Alexander. "Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners." *The Journal of the Acoustical Society of America* 121.4 (2007): 2339-2349.
- Brown, V. A. (2021). An Introduction to Linear Mixed-Effects Modeling in R. *Advances in Methods and Practices in Psychological Science, 4*(1), 2515245920960351.
- Brown, V. A., McLaughlin, D. J., Strand, J. F., & Van Engen, K. J. (2020). Rapid adaptation to fully intelligible nonnative-accented speech reduces listening effort. *Quarterly Journal of Experimental Psychology, 73*(9), 1431-1443.
- Clarke, C.M., & Garrett, M.F. (2004). Rapid adaptation for foreign-accented English. *Journal of the Acoustical Society of America 116*(6), 3647-3658.
- Cooke, M., Garcia Lecumberri, M. L., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America, 123*(1), 414-427.
- Ferguson, S.H., Jongman, A., Sereno, J.A., & Keum, K.A (2010). Intelligibility of Foreign-Accented Speech for Older Adults with and without Hearing Loss. *Journal of the American Academy of Audiology, 21*(3), 153-162.

- Gittleman, S. & Van Engen, K. J. (2018). Effects of noise and talker intelligibility on judgments of accentedness. *Journal of the Acoustical Society of America*. 10.1121/1.5038653
- Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society: JINS*, 13(2), 197-208.
- Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., Cohen, J. I., & Waldroup, C. (2013). Recognition of accented and unaccented speech in different maskers by younger and older listeners. *The Journal of the Acoustical Society of America*, 134(1), 618-627.
- Kahneman, D. (1973). Attention and effort (Vol. 1063, pp. 218-226). Englewood Cliffs, NJ: Prentice-Hall.
- Kim, J. (2020). *The Interplay Between Interference Control and L2 Proficiency in L2 Auditory Sentence Comprehension in the Presence of Verbal and Non-verbal Masking* (Doctoral dissertation, City University of New York).
- Krizman, J., Bradlow, A. R., Lam, S. S. Y., & Kraus, N. (2016). How bilinguals listen in noise: linguistic and non-linguistic factors. *Bilingualism: Language and Cognition*, 1-10.
- Lane, H. (1963). Foreign accent and speech distortion. *The Journal of the Acoustical Society of America*, 35(4), 451-453.
- Larraza, S., Samuel, S.G., & Oñederra, M.L. (2016) Listening to accented speech in a second language: first language and age of acquisition effects. *Journal of Experimental Psychology* 42(11), 1774-1797.
- Listen Lab. (2020, March 24). *Praat 11 - Make speech-shaped noise* [Video]. YouTube. <https://www.youtube.com/watch?v=9tRcPoDtx2s>
- Lucks Mendel, L., & Widner, H. (2016). Speech perception in noise for bilingual listeners with normal hearing. *International journal of audiology*, 55(2), 126-134.
- Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, 25(5), 605-621.
- Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism*, 6(2), 97.
- Mattys, S.L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Process* 27, 953–978.
- McLaughlin, D. J., & Van Engen, K. J. (2020). Task-evoked pupil response for accurately recognized accented speech. *The Journal of the Acoustical Society of America*, 147(2), EL151-EL156.

- McLaughlin, D., Baese-Berk, M., Bent, T., Borrie, S., & Van Engen, K. J. (2018). Coping with adversity: individual differences in the perception of noisy and accented speech. *Attention, Perception, and Psychophysics*. doi: 10.3758/s13414-018-1537-4
- Morini, G., & Newman, R. S. (2020). Monolingual and Bilingual Word Recognition and Word Learning in Background Noise. *Language and speech*, 63(2), 381-403.
- Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*, 45(1), 73-97.
- Munro, M.J., Derwing, T. M. (1995b). Processing Time, Accent, and Comprehensibility in the Perception of Native and Foreign-Accented Speech. *Language and Speech* 38(3), 289-306
- Nye, P. W., and Gaitenby, J. H. (1974). "The intelligibility of synthetic monosyllabic words in short, syntactically normal sentences," Haskins Laboratories Status Report on Speech Research, SR-37/38, pp. 169–190.
- Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics*, 27(3), 465.
- Rogers, C.L., Dalby, J., Nishi, K. (2004). Effects of noise and intelligibility of Chinese-accented English. *Lang and Speech*, 47, 139–154.
- Samuel, A.G. & Larraza, S. (2015). Does listening to non-native speech impair speech perception? *Journal of Memory and Language*, 81, 51-71.
- Schmidtke, J. (2016). The bilingual disadvantage in speech understanding in noise is likely a frequency effect related to reduced language exposure. *Frontiers in Psychology*, 7, 678.
- Shi, L-F. (2012). Contribution of linguistic variables to bilingual listeners' perception of degraded English sentences. *Journal of Speech, Language, and Hearing Research*, 55 219-234.
- Shi, L. F. (2010). Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition.
- Skoe, E., & Karayanidi, K. (2019). Bilingualism and speech understanding in noise: Auditory and linguistic factors. *Journal of the American Academy of Audiology*, 30(2), 115-130.
- Tabri, D., Michelle, K., Chacra, A., & Pring, T. (2011). Speech perception in noise by monolingual, bilingual, and trilingual listeners. *International Journal of Language & Communication Disorders*, 46(4), 411-422.
- Van Engen, K. J., & Bradlow, A. R. (2007). Sentence recognition in native-and foreign-language multi-talker background noise. *The Journal of the Acoustical Society of America*, 121(1), 519-526.

Van Engen, K. J., & McLaughlin, D. J. (2018). Eyes and ears: Using eye tracking and pupillometry to understand challenges to speech recognition. *Hearing research*, 369, 56-66.

Van Engen, K. J., Phelps, J. E., Smiljanic, R., & Chandrasekaran, B. (2014). Enhancing speech intelligibility: Interactions among context, modality, speech style, and masker. *Journal of Speech, Language, and Hearing Research*, 57(5), 1908-1918.

Weber, A., Di Betta, A. M., & McQueen, J. M. (2014). Treack or trit: Adaptation to genuine and arbitrary foreign accents by monolingual and bilingual listeners. *Journal of phonetics*, 46, 34-51.

Wilson, E., & Spaulding, T. (2010). Effects of noise and speech intelligibility on listener comprehension and processing time of Korean-accented English. *Journal of Speech, Language, and Hearing Research*, 53(6), 1543-1555.

Winn, M. B., Edwards, J. R., & Litovsky, R. Y. (2015). The impact of auditory spectral resolution on listening effort revealed by pupil dilation. *Ear and hearing*, 36(4), e153.

Zekveld, A. A., Rudner, M., Kramer, S. E., Lyzenga, J., & Rönnerberg, J. (2014). Cognitive processing load during listening is reduced more by decreasing voice similarity than by increasing spatial separation between target and masker speech. *Frontiers in Neuroscience*, 8, 88.