Monolingual and Bilingual Word Recognition and Word Learning in Background Noise

Giovanna Morini [iD]
Communication Sciences and Disorders Program, University of Delaware, USA

Rochelle S. Newman
Department of Hearing and Speech Sciences, University of Maryland, USA

Abstract
The question of whether bilingualism leads to advantages or disadvantages in linguistic abilities has been debated for many years. It is unclear whether growing up with one versus two languages is related to variations in the ability to process speech in the presence of background noise. We present findings from a word recognition and a word learning task with monolingual and bilingual adults. Bilinguals appear to be less accurate than monolinguals at identifying familiar words in the presence of white noise. However, the bilingual "disadvantage" identified during word recognition is not present when listeners were asked to acquire novel word-object relations that were trained either in noise or in quiet. This work suggests that linguistic experience and the demands associated with the type of task both play a role in the ability for listeners to process speech in noise.

Introduction

In the last 10 to 20 years, there has been a significant increase in the number of studies examining how bilingualism affects human faculties including language and cognition. Most studies examining this topic have focused on non-linguistic measures. For example, much excitement and debate has resulted from the reported finding of a "bilingual cognitive advantage," which claims that bilingualism confers benefits on attention and inhibitory control measures (Bialystok, Craik, Green, & Gollan, 2009; Hilchey & Klein, 2011). Nevertheless, there are various linguistic abilities that might be influenced by bilingualism, and which have received considerably less attention. The

Corresponding author:
Giovanna Morini, Communication Sciences and Disorders Program, University of Delaware, Newark, DE 19713, USA.
Email: gmorini@udel.edu.

present work focuses on one such skill, which is processing speech in the presence of competing background noise.

There are several reasons why bilinguals might develop different processing strategies than monolinguals. Bilinguals are required to "manage" two linguistic systems; this means monitoring (and switching between) the two languages on a regular basis. Furthermore, bilinguals must split lexical information across languages (e.g., words are heard in each language less frequently, and there are entries for the same concept in each lexicon—i.e., translation equivalents), potentially leading to weaker stored word knowledge and more lexical competition (Ecke, 2004; Gollan, Montoya, Fennema-Notestine, & Morris, 2005). As a result, individuals in bilingual environments might develop different strategies compared to monolinguals that help them achieve proficiency in both languages, and which may lead to differences in performance on linguistic measures.

Studies exploring the effect that bilingualism has on language-related tasks have primarily focused on lexical retrieval. This work suggests that when presented with naming tasks, bilinguals produced fewer correct responses compared to monolinguals (Roberts, Garcia, Desrochers, & Hernandez, 2002; Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007). Furthermore, when asked to complete timed/speeded versions of the naming tasks, bilinguals performed more slowly (Gollan et al., 2005) and made more errors (Bialystok, Craik, & Luk, 2008) compared to monolinguals. This pattern of results was true even when bilinguals were asked to produce words in their dominant language (Ivanova & Costa, 2008). Less is known about bilingual performance in other types of linguistic tasks, and specifically in tasks that involve processing speech in different listening environments.

Much of the time, listeners hear speech in environments that are rich in auditory and visual cues (i.e., multiple sounds and objects). To understand different messages, listeners must simultaneously process competing sounds and rely on cognitive abilities and previously acquired linguistic knowledge to interpret the information that is being conveyed. The process of separating two competing sound streams into the specific components that make up each signal and grouping together the elements that make up one stream is often referred to as *stream segregation*. To accomplish this process, listeners must rely on attention to focus on the target speech and simultaneously inhibit the competing signal (Mattys, Davis, Bradlow, & Scott, 2012). In addition to segregating the competing auditory streams, to successfully identify words, the listeners must map a degraded acoustic signal (one that is "incomplete" or partially masked by the noise) onto an existing representation that is presumably complete. Thus, understanding speech in the presence of noise necessitates several subtasks beyond simple lexical retrieval. The present work examines the effect that linguistic experience has on listeners' ability to perform listening-in-noise tasks.

Word recognition and word learning are two fundamental linguistic tasks that must often be accomplished in less-than-optimal listening conditions. Listeners regularly find themselves in settings where they are spoken to in the context of background noise, including high levels of ambient noise found in workplace situations such as open-plan offices or trading areas (review in Beaman, 2005). Given that stream segregation involves not only selecting one auditory source of information over another, but also accessing lexical representations (which may be different across listeners depending on language experience), studying bilinguals' performance during stream segregation tasks is of particular interest.

Prior work examining this topic has mainly been carried out with monolinguals and adult second language (L2) learners, who acquired their L2 later in life, and the focus has been on word recognition, rather than also examining word learning. These studies suggest that both monolingual and L2 speakers show comparable speech recognition abilities in quiet listening conditions, but the latter group performs significantly worse than monolinguals when there is noise in the background (Florentine, 1985a, 1985b; Florentine, Buus, Scharf, & Canévet, 1984). However, L2

learners are typically not balanced in their two languages and are hence likely to have weaker lexical representations (at least in one of their languages) compared to monolinguals or native bilinguals. This is a factor that can negatively impact speech processing in difficult listening environments.

Only a handful of studies have examined balanced bilingual adults' ability to process speech in the presence of competing noise. Mayo, Florentine, and Buus (1997) found that bilinguals who acquired English before the age of six showed better sentence comprehension in babble noise than those who acquired English post-puberty, yet none of the bilinguals performed as well as monolinguals on the task. This work, however, had a number of limitations, including the lack of an assessment of language proficiency in the bilingual participants (making it unclear as to whether the bilinguals were balanced, or only early learners), and very small sample sizes (e.g., only $n = 3$ in the group of early bilinguals). In another study, Meador, Flege, and Mackay (2000) found that native speakers of Italian who had moved to Canada at different ages performed differently on an English word-identification-in-noise task depending on their age of arrival to the English-speaking country. Participants with the earliest age of arrival (an average age of 7 years) performed better compared to the L2 speakers with a later age of arrival, but both bilingual groups performed significantly worse compared to monolinguals when the competing signal (pink noise in this case) was present. But would the same pattern of performance be observed with bilinguals who acquired their two languages even earlier in childhood (i.e., even earlier than 7 years)? A study by Rogers, Lister, Febo, Besing, and Abrams (2006) examined this question with Spanish–English bilinguals who had been exposed to Spanish since birth and to English before the age of 6 (M age of acquisition $= 2.8$ years). Bilingual participants had poorer word recognition scores compared to monolinguals when speech in English was presented in noise (both reverberation and speech-spectrum noise). However, this study once again included a relatively small number of participants ($n = 12$). Finally, a more recent study compared monolinguals', bilinguals', and trilinguals' ability to understand familiar words in the presence of babble noise, and found that both bilinguals and trilinguals had poorer performance on the task compared to monolinguals (Tabri, Chacra, & Pring, 2011). Once again, however, the sample sizes were small—with nine to 13 participants per group.

To our knowledge, only a single study has explored speech processing in noise in a large sample of balanced bilinguals, but this study was conducted with 14-year-olds and not adults (Krizman, Bradlow, Lam, & Kraus, 2017). The study compared monolingual and bilingual adolescents' ability to identify sentences-in-noise, words-in-noise, and tones-in-noise. Monolinguals performed better than bilinguals during the sentences-in-noise task, while bilinguals outperformed monolinguals during the tones-in-noise measure (where the degraded auditory target was non-linguistic) suggesting that the amount of linguistic information available during the task played an important role. This finding also implies that bilingualism might have multiple effects—on the one hand improving stream segregation abilities, while at the same time reducing accuracy during tasks that involve lexical identification. It is unclear, however, whether adolescent and adult bilinguals would show the same pattern of performance, or whether the bilingual disadvantage would hold with other types of linguistic tasks that do not involve identification of familiar words (e.g., novel word learning). It is also unclear whether the type of competing signal (speech babble vs. random noise) plays an important role on bilinguals' ability to process speech in noise.

Taken together, this prior work with bilinguals has mainly examined word recognition in adults who acquired a second language later in life, and has been carried out with relatively small sample sizes. The present studies aim to address some of these limitations and expand our understanding of how linguistic experience may play a role in stream segregation abilities.

2   Experiment I

As a first step, we examined the relationship between language experience and the ability to rely on selective attention and previously acquired knowledge to achieve word recognition in noise. The task utilized a modified version of the word identification paradigm introduced by Rogers et al. (2006) to test monolingual and bilingual young adults' ability to identify familiar words both in quiet and in the presence of competing background noise. During their visit to the lab, participants also completed a general demographic questionnaire. Lastly, the language proficiency of the bilingual participants was assessed using a language history questionnaire. This questionnaire was modeled after the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007) and included some additional questions regarding language exposure, experience, and usage. Prior work by Marian and colleagues has indicated that the use of self-report tools (such as the LEAP-Q) is a valid, reliable, and efficient way of assessing language proficiency in adult bilinguals.

2.1 Method

Participants: A total of 64 participants (32 monolinguals and 32 bilinguals) between the ages of 18 and 24 ($M = 20.2$ years, $SD = 1.55$) recruited from two universities (one in the United States and one in Canada) completed the study. Participants were all right-handed and had no history of attention or speech problems. Additionally, all participants reported having normal hearing. A direct measure of IQ was not obtained, but participants were all undergraduates attending the same universities, and thus were considered to be a more homogenous group than if a broader population had been included. The monolingual group included 10 males and 22 females who were born and raised in the United States or in Canada, who grew up in a monolingual English-speaking household, and who were fluent in English. While 97% of the monolingual participants reported having studied a foreign language for at least one year ($M = 3.9$, $SD = 2.45$), none of them reported being fluent in it.

The bilingual group included nine males and 23 females who had acquired English as well as one other language before the age of 5 (M age of acquisition $= 2.7$, $SD = 1.5$), and who still used both languages on a regular basis (i.e., each language was used at least 30% of the time). Eighteen of the bilingual participants were born and raised in the United States, four were born and raised in the Toronto area in Canada, and 10 were born in a foreign country (five in India, one in China, one in the Philippines, two in Iran, and one in Russia) but had grown up speaking both languages. A detailed distribution of the non-English language (i.e., the second language) is provided in Appendix A. Additionally, 65% of the bilingual participants reported having studied a third language after the age of 5 for at least one year ($M = 3.2$, $SD = 2.4$), but none of them reported fluency in the third language. Based on responses from the language history questionnaire, on a scale from 1 (little to no knowledge) to 7 (like a native speaker), bilingual participants' proficiency level was on average 6.6 ($SD = 0.78$) in English, and 6.0 ($SD = 0.95$) in the other language. Specifically, six of the participants reported equal proficiency in both languages, 18 participants reported higher proficiency in English, and the remaining eight participants indicated higher proficiency in their other language.

Stimuli: The auditory stimuli consisted of a target speech stream and a competing noise signal. As in the study by Rogers et al. (2006), the target speech stimuli were words from the Central Institute for the Deaf (CID) W-22. This is a test that is commonly used to assess word recognition, where items are organized into phonetically balanced lists (Hirsch, Davis, Silverman, Reynolds, Eldert, & Benson, 1952). All words were monosyllabic and in consonant-vowel, vowel-consonant,

or consonant-vowel-consonant format. Some examples were "*knee*," "*wet*," "*young*" (for the complete lists of words see Appendix B). All words were presented at the end of a carrier phrase "repeat the word _____" and recorded in a sound-attenuating booth by a female native speaker of English. Recordings were made at a sampling rate of 44.1 kHz and were digitalized using a 16-bit analog-to-digital converter. Excess silence at the start and end of sentences was removed (with cuts occurring at zero crossings). Additionally, the individual sentences were edited to have the same root mean square (RMS) amplitude, and were then stored on computer disk. The competing signal was white noise presented with a steady-state amplitude envelope delivered in combination with the speech at approximately 70–75 dB sound pressure level (SPL) via headphones. The same sample of white noise was used across trials. This type of noise was chosen because it is one that has been previously used during speech perception tasks (e.g., Carhart, Tillman, & Greetis, 1969; Bradlow & Bent, 2002; Johnstone & Litovsky, 2006), and because it does not contain language-specific features that would share similarity with speech in a particular language (an important factor since the non-English language of our bilinguals varied across participants).

Trials in the noise condition were presented in one of three possible signal-to-noise ratios (SNRs; 0 dB, -4 dB, and -8 dB). This meant that the target speech and the competing noise were heard at the same intensity level, or the background noise was 4 dB or 8 dB more intense than the target speech. The background signal always began 500 ms prior to the target speech and continued until the end of the trial.

Apparatus and procedure: Testing sessions were conducted in a quiet room with participants seated in front of a computer monitor. Audio recordings of the test sessions (for later transcription) were digitally generated using the built-in microphone in the computer and Audacity audio recording software.

Participants were asked to repeat aloud the familiar word heard at the end of the carrier phrase (e.g., repeat the word *young*). A list of 50 target words was presented at each of the three SNRs for the noisy trials, as well as in the quiet condition (for a total of four word lists and 200 words). The list-to-SNR level was counterbalanced across participants so that each list was heard an equal number of times in each of the four conditions across participants. Words were randomly selected for presentation from the different lists by the computer using PsyScope X experiment control system (Cohen, MacWhinney, Flatt, & Provost, 1993). This resulted in trials from different conditions being intermixed. At the beginning of each trial, a black fixation cross flashed in the center of the screen for 500 ms. Immediately after, participants heard the auditory stimulus for that trial presented through the headphones. A 100 ms beep was played through external speakers (captured by the microphone that recorded the test session but not heard through the participants' headphones) at the offset of the target word to allow for later reaction time (RT) coding. Participants were instructed to repeat target words as quickly and as accurately as possible. Participants also saw a message that said "say the word, THEN press [space]" appear on the screen as a reminder. Once participants had verbally produced the target word and pressed the space bar on the keyboard, the next trial began.

Coding and measures: Audio recordings from the test sessions were transcribed and analyzed for accuracy and RT by trained coders. Recordings from 12% of the participants were coded by a second coder in order to ensure reliability. *Accuracy* for each group was calculated based on the number of words that participants repeated correctly. Responses were grouped based on listening condition (i.e., Quiet, 0 dB SNR, -4 dB SNR, and -8 dB SNR) in order to calculate the proportion of trials in which the correct target word was produced for each condition. RT was defined as the length of time from the offset of the auditory cue (i.e., the beginning of the beep, which was also the end of the target word) to the onset of the participant's verbal response during trials where a correct response was produced. This meant that only a portion of the trials were included in the
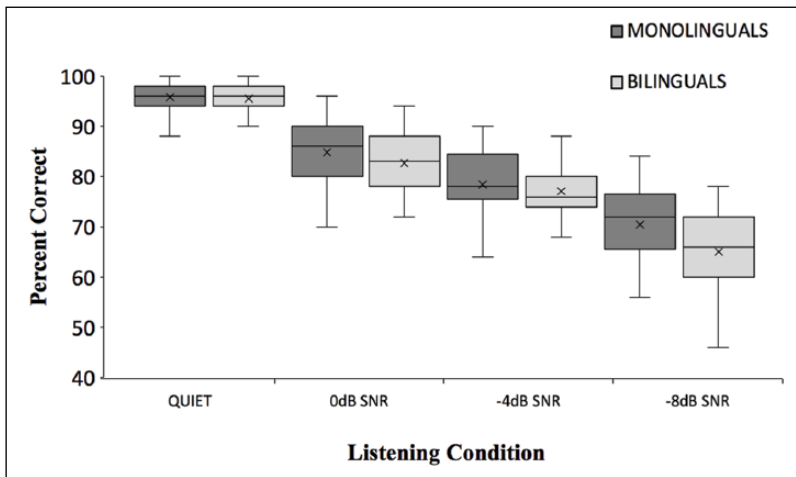
Figure 1. Percentage of correct responses produced by monolinguals and bilinguals in Experiment 1. The x inside the boxes represents the mean.

RT analyses—that is, only those trials with a correct answer; while this leads to a reduced number of data points, it is preferable to analyzing RT measures based on incorrect responses. Verbal responses were hand-coded for RT (rather than analyzed using a voice key) so as to avoid potentially wrong RTs that could result from participants producing disfluencies or other noises prior to their response.

2.2 Results

As shown in Figure 1, initial inspection of the data revealed that there was a gradual effect of noise level on accuracy, and that listeners were still able to process speech even in the most difficult background noise condition, as indicated by relatively high accuracy (note that "chance" performance is essentially zero in this open-set task). This was the case for both language groups. To account for potential variability issues among subjects and/or test items, accuracy was modeled using the lme4 package in an R environment with RStudio. The most complex random effects structure supported by the data included random intercepts for participant and item. The fixed effects are presented in Table 1. The quiet condition served as the reference group for condition. A reliable main effect emerged for condition such that participants performed more poorly in all noise conditions than in the quiet condition ($t = $ -9.31, -14.08, and -20.648 for 0dB, -4dB, and -8dB respectively, $p < .001$ for 0dB, -4dB, and -8dB). No reliable effect of group emerged, but an interaction between group and condition emerged such that bilingual participants experienced a greater reduction in performance than monolingual participants in the -8dB condition as compared to other conditions ($t = $ -3.50, $p < .001$). These analyses suggest that bilinguals are not worse overall at the task, but are more prone to mistakes than monolinguals as the noise level becomes more difficult. It is also worth noting that this difference was present despite the fact that the vast majority of bilingual participants reported that they were as (or more) proficient in English as in their other language; thus, the current pattern does not appear to be the result of testing in a less proficient, or non-dominant, language.

   Next, the RT data were modeled using the same parameters used for accuracy. Only accurate responses were included in the analysis, and all RTs were standardized prior to being entered in the

Table 1. Fixed effects for accuracy data in Experiment 1.

	Estimate	Std error	df	t value	p value
(Intercept)***	0.959	0.016	367.000	59.111	<0.001
Group	-0.008	0.015	207.000	-0.500	0.618
Condition 0dB***	-0.109	0.012	12530.000	-9.314	<0.001
Condition -4dB***	-0.164	0.012	12530.000	-14.078	<0.001
Condition -8dB***	-0.241	0.012	12530.000	-20.648	<0.001
Group x Condition 0dB	-0.013	0.017	12530.000	-0.757	0.449
Group x Condition -4dB	-0.025	0.017	12530.000	-1.514	0.130
Group x Condition -8dB***	-0.058	0.017	12530.000	-3.492	<0.001

Reference groups: Monolingual, Quiet.

Table 2. Fixed effects for RT data in Experiment 1.

	Estimate	Std Error	df	t value	p value
(Intercept)***	0.318	0.086	561.900	3.697	<0.001
Group	0.210	0.122	585.700	1.727	0.085
Condition 0dB	0.016	0.124	10280.000	0.130	0.896
Condition -4dB	0.036	0.126	10320.000	0.288	0.773
Condition -8dB	0.066	0.130	10360.000	0.512	0.608
Group x Condition 0dB	-0.221	0.176	10240.000	-1.257	0.209
Group x Condition -4dB	-0.240	0.179	10270.000	-1.341	0.180
Group x Condition -8dB	-0.242	0.186	10290.000	-1.299	0.194

Reference groups: Monolingual, Quiet.

model. The most complex random effects structure supported by the data included random intercepts for participant and item. The fixed effects are presented in Table 2. No reliable main effects or interactions emerged.

2.3 Discussion

The goal of this experiment was to examine the effect that linguistic experience (i.e., growing up monolingual vs. bilingual) might have on listeners' ability to understand speech in the presence of background noise. Bilinguals did not have an overall difficulty processing speech during the task. However, they did experience a greater performance drop in their ability to recognize familiar words in the -8dB noise condition compared to the quiet condition than the monolingual group did. These results align with previous findings obtained in studies with smaller sample sizes (e.g., Mayo et al., 1997; Meador et al., 2000). Furthermore, they suggest that the bilingual disadvantage during word identification in noise is consistent across different types of background noises including babble (Mayo et al., 1997), pink noise (Meador et al., 2000), reverberation and speech-spectrum noise (Rogers et al., 2006), and now white noise. Additional research is needed, however, to understand the role that the SNR might play.

One possibility is that the results are driven by the demands associated with the specific task, and the way lexical information is stored and accessed in the bilingual mind. Bilingual listeners need to inhibit a competing linguistic system, which may result in having fewer resources available

to process speech; this would not pose a difficulty when listening is easy, but would result in poorer performance as the listening condition became more noisy and stream segregation became more resource-demanding. An alternative explanation is that bilinguals' performance in noise was related to the fact that this task relies heavily on accessing previously acquired knowledge to fill in masked information. Prior work suggests that bilinguals have weaker stored word knowledge and more lexical competitors (Ecke, 2004; Gollan et al., 2005); this splitting of lexical information across two languages potentially leads to the "access pathways" being less automatic. In our task, participants were not only asked to identify words with a signal quality that was reduced by the overlapping white noise, but they were asked to do so as quickly as possible. Furthermore, the presentation of the different SNRs was intermixed throughout the task, which also added difficulty compared to if the different noise levels had been blocked.

When there is high task demand (such as the one caused by segregating competing auditory signals, alternating between SNRs, or providing a response quickly), and in particular when the demand is related to accessing previously acquired lexical information, the weaker stored knowledge in bilinguals may lead to lower response accuracy compared to monolinguals; this is a hypothesis that has been previously introduced in the bilingual literature (e.g., Rogers et al., 2006; Pichora-Fuller, Scheider, & Daneman, 1995). On the other hand, when the task demand is lower (as in the quiet condition, where the signal quality is higher) bilinguals can dedicate more attentional resources to accessing the stored information and compensate for their weaker representations.

There are other linguistic tasks (e.g., word learning) that presumably rely more heavily on processing information that is being perceived from the environment and less on previously stored lexical knowledge. This difference in task demands may lead bilinguals not to experience the same processing costs identified during word-recognition-in-noise tasks. Even though word learning is a phenomenon that is primarily associated with young children who are expanding their vocabulary or with individuals who are acquiring a foreign language, adults (both monolingual and bilingual) are learning new words all the time (Leach & Samuel, 2007). Hence, as a next step we examined monolinguals' and early bilinguals' ability to learn new words in noise.

3   Experiment 2

Word learning and word recognition are essential linguistic tasks. On the surface, the two may appear to be similar, in that they both involve processing the speech signal to extract a message that is being conveyed. However, there is presumably more potential for stored information to play a role in word recognition compared to word learning. Understanding spoken language requires mainly being able to combine *existing* lexical knowledge with the acoustic information that is being received almost simultaneously. For language learning, on the other hand, the individual must primarily focus on generating *new* connections between unfamiliar words that are heard and their corresponding meaning, and storing this newly acquired information.

Furthermore, several lines of work support the notion that knowledge of multiple languages may actually facilitate performance during word-learning tasks in adulthood. Kaushanskaya and Marian (2009) found that bilinguals were better than monolinguals at learning novel words in a foreign language. The authors argued that bilingualism leads to the development of a phonological memory system that is more "efficient," and hence allows listeners to better remember new terms, even when the phonological information is not familiar (Papagno & Vallar, 1995). However, a more recent word-learning study in which monolingual and bilingual adults were matched in their phonological memory capacity still revealed the same group differences in performance (Kaushanskaya, 2012), suggesting that phonological memory was not the reason for the group

differences. Kaushanskaya and Rechtzigel (2012) subsequently examined the possibility that bilinguals and monolinguals might differ in how they process the semantic information associated with novel terms. Data from this study revealed that the bilingual advantage was greater for novel words that corresponded to concrete (compared to abstract) referents. Words that have concrete meanings are not only thought to lead to a richer activation of stored lexical representations (Grondin, Lupker, & McRae, 2009), but in the case of bilinguals these words also share a greater semantic overlap across the two languages compared to words that correspond to abstract referents (Van Hell & De Groot, 1998). Therefore, Kaushanskaya and Rechtzigel (2012) introduced the possibility that the concreteness of the word leads to a stronger activation of the semantic system during encoding, particularly when the meaning of the novel term is activated across the two linguistic systems.

Nevertheless, all of the above findings were gathered using word-learning tasks that relied on paired associates learning methods, where the novel terms were trained through associations with previously acquired translation equivalents (e.g., learn novel word ____, which in English means ____). While this type of approach is relevant to how adults might learn a second language (after having acquired a first language), learning words through translations is not necessarily how bilinguals acquire their two language systems early in life, nor is it necessarily the means by which adults typically learn new words (e.g., technical vocabulary in a new field).

In this second experiment, we explored the relationship between language experience and listeners' ability to process speech in noise in order to learn entirely novel word-object relations. The task relied on a modified version of the word-learning measure introduced by Gupta (2003) and tested monolinguals' and bilinguals' word-learning skills both in quiet and in the presence of background noise. As in Experiment 1, participants completed a general demographic questionnaire and a language history questionnaire.

3.1 Method

Participants: A total of 64 participants (32 monolinguals and 32 bilinguals) between the ages of 18 and 33 (M = 20.4 years, SD = 2.3) recruited from the same two universities in the United States and in Canada completed the study. Participants were all right-handed and had no history of attention or speech problems. All participants reported having normal hearing. As in the first experiment, a direct measure of IQ was not obtained, but participants were all undergraduates attending the same universities, and thus were considered to be a more homogenous group than if a broader population had been included. The monolingual group included 14 males and 18 females who were born and raised in the United States or in Canada, who grew up in a monolingual English-speaking household, and who were fluent in English. While 90% of the monolingual participants reported having studied a foreign language for at least one year (M = 3.5, SD = 2.20), none of them reported being fluent in it.

The bilingual group included nine males and 23 females who had acquired English as well as one other language before the age of 5 (M age of acquisition = 2.3, SD = 1.5), and who still used both languages on a regular basis (i.e., each language was used at least 30% of the time). Fourteen of the bilingual participants were born and raised in the United States, six were born and raised in the Toronto area in Canada, and 12 were born in a foreign country (two in Hong Kong, one in Ghana, one in Saudi Arabia, one in Nigeria, five in India, one in Zimbabwe, and one in Brazil) but had grown up speaking both languages. Of these 12 participants born abroad, six had moved to the US or Canada while they were infants. A detailed distribution of the non-English language (i.e., the second language) is provided in Appendix A. Additionally, 71% of the bilingual participants reported having studied a third language after the age of 5 for at least one year (M = 3.4,

Table 3. Spelling constraints on non-word construction for Experiment 2.

	Word formation
Non-word	$syll_{\#1}$ + $syll_{\#2}$ + $syll_{\#3}$ + $syll_{\#4}$
$syll_{\#1}$	Consonant + Vowel
$syll_{\#2}$	Consonant + Vowel
$syll_{\#3}$	Consonant + Vowel
$syll_{\#4}$	Consonant + Vowel + Coda
	Syllable formation
Consonant	b, ch, d, g, j ,k, p, t
Vowel	a, e, i, o, u
Coda	b/be, c/ce, d/de, f/fe, g/ge, l/le, m/me, n/ne, nt, p/pe, r/re, rt, s, st, t/te, th, ve, x, ze

$SD = 2.6$), but none of them reported fluency in the third language. Based on responses from the same language history questionnaire used in Experiment 1, on a scale from 1 to 7 bilingual participants' proficiency level was on average 6.5 ($SD = 0.67$) in English, and 6.0 ($SD = 1.07$) in the other language, with 13 participants reporting equal proficiency in both languages, 13 reporting higher proficiency in English, and six reporting higher proficiency in their other language. These scores were not significantly different from the ones obtained from the bilingual group in Experiment 1.

Stimuli: The auditory stimuli for the word-learning task consisted of 78 four-syllable non-words (six targets and 72 foils) that were generated based on the constraints in Table 3, which followed English phonotactics (these were the same rules used by Gupta, 2003). All novel words had second syllable stress (e.g., *che-CHE-pa-tile)* and were recorded by the same female speaker from Experiment 1 (for a complete list of words see Appendix C). All non-words were presented in isolation (as opposed to being included in a sentence) and were recorded and edited using exactly the same procedure described in Experiment 1. The competing signal was once again white noise presented with a steady-state amplitude. As in the first experiment, the same sample of white noise was used across trials.[1] Trials in the noise condition were presented using a -2dB SNR. There were two factors that influenced the decision to use this noise level. First, given that it was a learning task that required repeated training, there was a limited number of trials that could be presented in each condition (i.e., from a design point, including multiple SNRs would have made the task too long). Second, given the higher level of difficulty of the task (compared to the word recognition task in Experiment 1), the -8dB and even the -4dB SNRs would likely have made it too hard for participants to accurately learn the novel words. Hence, the easier SNR was chosen. The background signal always began 500 ms prior to the target speech and continued until the end of the trial.

The visual stimuli (the objects to be learned) consisted of six colored pictures of novel "creatures" created using SPORE[(TM)] Creature Creator software. All objects had different features and colors and were presented on a white background (see Appendix D). Objects were selected because they were easily distinguishable from one another and because they would not have any previously existing names.

Apparatus and procedure: Testing sessions were conducted using the exact same setup as in Experiment 1. Participants were asked to wear headphones and to repeat aloud non-words that in some situations corresponded to a particular novel object that appeared on the screen. Audio

Table 4. Structure of word learning task in Experiment 2.

Step/trial		Stimulus	Participant's response
1.		Non-word foil	Repeat non-word
2.	Training 1	Non-word target 1 + Picture of creature 1	Repeat non-word
3.	Training 2	Non-word target 1 + Picture of creature 1	Repeat non-word
4.		Non-word foil	Repeat non-word
5.	Early test	Cue: Picture of single creature 1	Name creature
6.		[Repeat trials 1–5 with creature 1, Target 1]	
7.		[Repeat trials 1–6 with creature 2, Target 2]	
8.	Joint test	Cue: Picture of single creature	Name creature
9.	Joint test	Cue: Picture of single creature	Name creature
10.		[Repeat trials 1–9 two more times]	
11.		[Repeat trials 1–10 with creature and Target 3, 4]	
12.		[Repeat trials 1–10 with creature and Target 5, 6]	

1block = 22 trials.

recordings of the entire test sessions were digitally generated using the built-in microphone in the computer. The experimental design is outlined in Table 4, and included multiple blocks and trials. During the first four trials of each block, participants heard a non-word and repeated it aloud. Two of these were foil trials (where "distractor" non-words were presented auditorily without a referent on the screen) and the other two were training trials (where the target label was presented, accompanied by a picture on the screen of the corresponding referent). Thus, participants heard the target word (and saw its referent) twice, and heard two other foil words once each. After these four trials, the participants received a test trial, in which they were cued with the image of the referent from that block and were asked to produce the corresponding label.

In order to avoid immediately repeating the trained word, the training trials either occurred on the first two trials, or on trials 2 and 3; a foil always occurred immediately before the test trial. This type of design was chosen so as to minimize the effects of proactive and retroactive interference. As discussed in Gupta (2003), prior word learning studies with adults have frequently used list-learning paradigms where participants are asked to repeat lists of novel words. One concern with this approach is that proactive interference (i.e., when older items get in the way of trying to recall more recent items) and retroactive interference (i.e., when more recent items get in the way of trying to recall previous items) have been widely reported in list-learning situations (Crowder, 1976; Postman, 1976; Wickens, Born, & Allen, 1963; Gupta 1995). These findings come primarily from work with monolinguals and it is unclear whether or not bilinguals experience the same level of interference. If bilinguals and monolinguals differ in the extent of such interference, then relying on a list-learning task might give an impression of group differences that is strictly tied to the list format and not word learning or stream segregation abilities present in each group. Therefore, in the current design we avoided presenting the word-picture pairs in lists. Additionally, we included distractor/foil items, since asking adult participants to learn a single word-picture pairing would make the task too easy, and would not provide a sensitive measure.

The five-trial sequence was then repeated a second time for the same word. The same process then occurred for a second word-object pair (a total of 20 trials). This was followed by two joint test trials (one for each word) in which participants saw one of the two referents they had learned (in random order) and were asked to name it. This entire sequence was considered one block;

participants repeated this entire block sequence two more times with the same words, and then again with two more sets of words (for a total of nine blocks, in which they learned six new word-object pairs). The order of presentation of the picture cues was randomized by the computer using PsyScope experiment control system (Cohen et al., 1993). Different foils were used within and across blocks. Training trials for half of the word-object pairs were presented in the -2dB SNR noise condition, while the other half were presented in quiet. This resulted in trials from the two conditions being intermixed (i.e., within each block one target was trained in noise and one target was trained in quiet, and the computer randomly selected which condition was presented first). Which target word occurred in noise versus silence, and which label was associated with which object, were counterbalanced across participants so that each target word was heard an equal number of times in noise and in quiet, and each referent was paired with the six different target words.

At the beginning of the testing session, participants were told that they would be completing a word learning task. They were asked to repeat aloud all the words that they heard during training, but to particularly focus on learning the words that corresponded to an image on the screen. Additionally, participants were instructed to provide responses as quickly and as accurately as they could. Following these initial guidelines, participants completed a practice block to ensure that they understood the task. At the start of each training trial participants saw a message that said "repeat the word, THEN press [space]" appear on the screen for 2000 ms. Immediately after, participants heard the non-word and saw the corresponding object appear on the screen. Trials during which a foil word was presented were identical, except for the fact that no object was seen. Participants were given 4500 ms to provide a verbal response and press the space bar, at which point the current trial ended and the following one began. If no response was provided during this time window, the trial timed out and the next one began. Test trials began with the message "say the name of the object, THEN press [space]" flashing on the screen for 2000 ms. This event was followed by a picture of the trained referent appearing on the screen and a 100 ms beep being played through external speakers at the onset of the image (so as to be captured by the computer's microphone but not heard through the participants' headphones). The image of the object remained on the screen until participants provided a verbal response and pressed the space bar. At that point the test trial ended and the following trial began.

After completing the task, the experimenter asked each bilingual participant whether any of the novel words heard during the study was a "real word" in their other language. Two of the participants stated that there was one word that sounded similar to a word in their other language, but none identified any of the novel words as being part of their vocabulary (i.e., identical to a word in their other language).

Coding and measures: Participants' verbal responses during test trials were transcribed and analyzed for accuracy as well as RT, and compared across the two groups. Responses were grouped by listening condition. Accuracy was calculated based on the number of words that participants produced correctly on the joint test trials (at the end of each block). Responses were counted as correct when at least three out of the four syllables in the word were accurately produced.[2] RT was defined as the amount of time from the onset of the visual cue (i.e., the beginning of the beep, which corresponded to the moment when the object appeared on the screen) to the onset of the participant's verbal response during joint test trials where a correct response was produced. It is not uncommon to examine RT data on word learning studies where there are accuracy differences (e.g., Gaskell & Dumay, 2003). Additionally, participants' verbal responses during training trials in the noise condition were transcribed and analyzed for accuracy. This was done in order to identify if words were heard accurately in that condition to begin with. As in the first experiment, verbal responses were hand-coded.
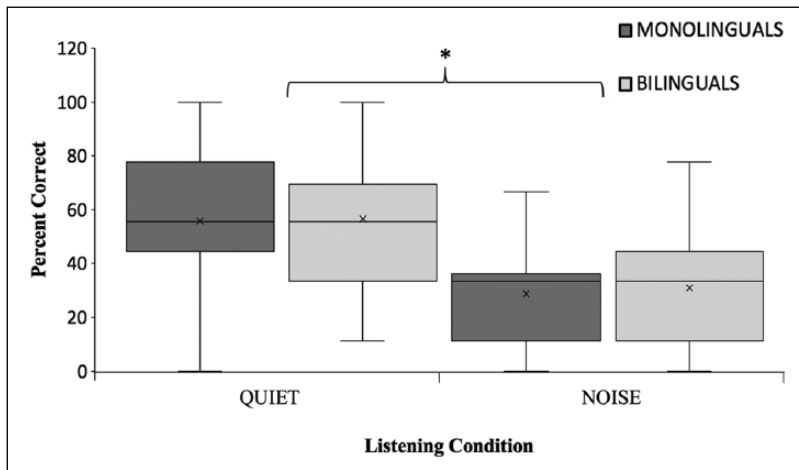
Figure 2. Percentage of correct responses produced by monolinguals and bilinguals in Experiment 2. The x inside the boxes represents the mean.

Table 5. Fixed effects for accuracy data in Experiment 2.

	Estimate	Std error	df	t value	p value
(Intercept)***	0.538	0.079	8.300	6.841	0.001
Group	-0.010	0.056	95.500	-0.179	0.859
Condition***	-0.259	0.035	1080.000	-7.428	<0.001
Group x Condition	0.029	0.049	1080.000	0.592	0.554

Reference groups: Monolingual, Quiet.

3.2 Results

Accuracy performance is summarized in Figure 2. Word learning accuracy was modeled using the same procedure described in Experiment 1. The most complex random effects structure supported by the data included random intercepts for participant and item. The fixed effects are presented in Table 5. The quiet condition served as the reference group for condition. A reliable main effect emerged for condition such that participants performed more poorly in noise than in the quiet condition ($t = -7.43, p < .001$). No additional reliable main effects or interactions emerged.

This suggests that the listening condition in which novel words were trained influenced accuracy during test trials. However, the linguistic experience of the listeners did not affect this measure. In general, both monolinguals and bilinguals were less accurate at learning new words when there was white noise present in the background during training ($M = 30\%$ accuracy, $SD = .19$) than when the non-words were trained in the quiet condition ($M = 56\%$ accuracy, $SD = .23$), and this difference was significant, $t(62) = 9.08, p < .001$, Cohen's $d = 1.14$. But there was no indication that bilingual learners were more affected by noise (26% accuracy decrement) than were monolingual learners (27% accuracy decrement).

Next we examined the RT data. It is important to note that the number of critical observations for the RT measure in this second experiment is rather low, since only correct trials were considered and overall accuracy was low for the task. Additionally, four of the monolingual participants

Table 6. Fixed effects for RT data in Experiment 2.

	Estimate	Std error	df	t value	p value
(Intercept)	0.252	0.129	69.006	1.953	0.055
Group	-0.164	0.183	68.936	-0.893	0.375
Condition	-0.136	0.137	429.814	-0.989	0.323
Group x Condition	0.004	0.193	429.352	0.019	0.985

Reference groups: Monolingual, Quiet.

and four of the bilinguals had to be excluded from the RT analysis since they did not have an RT score for one of the two experimental conditions (i.e., since RT was only calculated for trials where a correct response was provided, if a participant produced a correct response during trials corresponding to one training condition, but not to trials in the other condition, they were excluded from the analysis). Given the more limited data available, analyses for RT in this task should be cautiously interpreted.

RT was modeled once again using the lme4 package in an R environment with RStudio. The most complex random effects structure supported by the data included random intercepts for participant and item. The fixed effects are presented in Table 6. No reliable main effects or interactions emerged, perhaps because of substantial variability combined with a limited amount of data (since incorrect responses were not analyzed).

Additionally, analyses of participants' responses during training trials in the noise condition revealed that monolinguals were slightly better at accurately perceiving words that were trained in noise ($M = 57\%$ accuracy, $SD = .21$) compared to bilinguals ($M = 53\%$ accuracy, $SD = .22$), however this difference was not significant, $t(62) = .79$, $p > .05$. Furthermore, of those words that were produced correctly during training, monolinguals "retained" (i.e., also produced correctly during test trials) 13.6% of them, compared to 13.8% for bilinguals. This difference was once again not significant, $t(62) = .12$, $p > .05$. This suggests that there were no differences across groups in their ability to perceive words in noise during training, nor to learn/retain words that were trained (and correctly heard) in noise.

3.3 Discussion

The goal of this second experiment was to evaluate the notion that differences in performance during stream segregation tasks between monolinguals and bilinguals might be associated with task demands. Findings from the word identification paradigm in Experiment 1 suggested that bilinguals experienced more processing costs compared to monolinguals when there was noise in the background (i.e., in -8dB white noise), and were therefore less accurate at comprehending speech that was presented in this condition. In Experiment 2, monolinguals and bilinguals were asked to complete a language task that relied less on accessing existing lexical knowledge to identify words, and more on (a) generating new word-object relations, and (b) storing this newly acquired information. Performance during this word-learning task revealed that, not surprisingly, it was considerably more difficult for listeners to learn new words in the presence of noise than in quiet. However, this was equally the case both for monolinguals and bilinguals; differences in linguistic experience did not lead to greater (or lesser) accuracy during the task. Bilinguals were just as accurate as monolinguals across conditions. This pattern of results differs from what was observed in the word recognition task, where bilingualism was associated with poorer performance during word identification in noise.

Taken together, these findings support the idea that depending on the language experience of the individual and the demands associated with the linguistic task, participants might be weighting information sources differently. To succeed in Experiment 2, participants were mainly required to focus on the visual and auditory stimuli during training, store this information, and later recall it (and, for those words taught in the presence of noise, to selectively attend to the target signal). Apparently, these tasks are not themselves subject to disadvantages caused by multiple-language experience. Given that all non-words followed phonotactic rules of English and corresponded to referents that would not have had a lexical equivalent in the bilinguals' other language, this task likely did not lead to the same lexical competition that bilinguals experienced in the word recognition task. Moreover, the current task did not require use of stored representations, supporting the idea that weaker stored representations or interference in gaining access to stored representations may be the cause of bilinguals' poorer performance in Experiment 1. Additionally, during Experiment 2, participants presumably had more time for rehearsal as words were trained and learned, whereas in Experiment 1 there were potentially greater demands simultaneously. The differences in the approaches that were necessary to succeed in each of the tasks could explain why bilinguals would experience more processing costs in Experiment 1, but not in Experiment 2.

The word learning task from Experiment 2 was chosen because it presumably relies primarily on establishing new connections between novel words and their meanings, as well as the use of selective attention (a skill that may be enhanced by the experience of bilingualism). With this in mind, it was somewhat surprising that bilinguals were not *more* accurate than monolinguals in the noise condition. It is possible that the difficulty level in this condition was too high for a group difference to be observable. While participants in both groups did manage to learn novel words when there was white noise that was 2dB louder than the target speech (as indicated by the fact that accuracy performance was not at floor for either group in this condition), performance levels suggest that this was a very difficult task, and any cognitive advantages associated with bilingualism might not have been enough to generate a significant group difference at this noise level. Adjusting the degree of difficulty (e.g., using an easier SNR, providing additional training) could potentially lead to more drastic differences in accuracy across groups in this type of listening condition.

4   General discussion

The present work set out to investigate the effect that linguistic experience has on monolinguals' and bilinguals' ability to process speech in noise. To date, most studies on bilingual stream segregation have been conducted with L2 learners (who acquired one of their languages later in life), or with very small groups of adult early bilinguals. Hence, the present findings expand our knowledge of this topic, while addressing limitations of prior work. In addition, prior work has tended to focus primarily on word recognition, whereas the current study explores both word recognition and word learning with very similar groups of participants.

Data from two experiments suggest that bilinguals are less able than monolinguals to process speech in noise, but *only* during word-identification tasks. When completing a word-learning task, bilinguals performed stream segregation and acquired novel words just as well as monolinguals. Previous theories of bilingual language processing suggest that there are different factors between bilinguals and monolinguals that are potentially at play. During tasks that require identifying and producing words in a single language, bilinguals, but not monolinguals, must deactivate/inhibit the non-target language in order to successfully identify words in the linguistic system that is being targeted (Grosjean, 1997; MacKay & Flege, 2004). This extra step that bilinguals must complete leads to additional demands on the attentional resources that are available for speech processing. This element might be particularly problematic for bilinguals during tasks where there is an even

greater processing load, as in the case of word identification in noise (e.g., Mayo et al., 1997; Meador et al., 2000). This would explain why during Experiment 1 bilinguals showed poorer accuracy compared to monolinguals when words were presented in higher levels of noise, but not during the easier listening condition (i.e., the quiet condition) with lower task demands. Furthermore, bilinguals presumably have weaker stored word knowledge as a result of hearing words in a given language less often compared to monolinguals (Ecke, 2004; Gollan et al., 2005), and the weaker representations may lead to lower response accuracy—particularly in tasks that place a large premium on accessing stored information. Yet when the strength of the representations is matched across the two groups (as in Experiment 2), then the difference in performance across groups is no longer present.

Findings from the present work expand our understanding of bilingual stream segregation abilities. Prior work with this population has focused primarily on word recognition, and has led to claims regarding a bilingual disadvantage when processing speech in noise (Florentine, 1985a, 1985b; Florentine, Buus, Scharf, & Canévet, 1984; Mayo, Florentine, & Buus, 1997; Meador, Flege, & Mackay, 2000; Rogers, Lister, Febo, Besing, & Abrams, 2006). Our work provides some of the first data associated with bilingual word-learning in noise. The current findings not only suggest that bilinguals are able to acquire words in difficult listening conditions, but they are not worse than monolinguals at doing so. If bilinguals were truly worse at the act of stream segregation, this would presumably impact their ability to separate the target speech from the noise during training trials in Experiment 2. Since they seemed to learn equivalently well as monolinguals, it does not appear that they experience poorer stream segregation *per se.* Additionally, qualitative examination of the types of errors individuals made during Experiments 1 and 2 suggested that they were similar in both monolinguals and bilinguals, further supporting the idea of comparable stream segregation abilities across groups. In general, the errors that participants produced during the word recognition task revealed difficulty with consonants (in particular noise-like consonants such as voiceless stops and fricatives), but vowels and syllable structure were preserved (e.g., if the target word was "*ten*," they might produce "*pen*"). Since consonants generally have lower intensity and shorter durations than vowels, this is the pattern we might expect if the noise was masking portions of the target signal (energetic masking). However, it is also what might be expected if there were difficulties in stream segregation, since the background noise would be more likely to interact with an aperiodic segment than a vocalic one. During the word-learning task, individuals altered the phonetic composition of syllables themselves, but preserved the number of syllables and the word stress (e.g., if the target word was "*che-CHE-pa-tile*," they might produce "*che-CHE-ah-peel*"). But in neither case were these errors markedly different across groups. Thus, it appears that while bilinguals may show poorer performance recognizing known words in noise, the difficulty does not lie in the perception of the speech signal per se.

Our data also contribute to the existing literature on bilingual and monolingual word learning more generally. Prior work had found that bilingualism was associated with advantages in the acquisition of novel words (e.g., Kaushanskaya & Marian, 2009; Kaushanskaya, 2012). In the present study, bilingualism did not lead to higher accuracy scores. Here we used a word-learning task in which participants were asked to learn an entirely novel word-object relation. This is quite different from many prior word-learning studies with bilinguals that have focused on translation equivalents (such as the Kaushanskaya studies). We chose this novel-word approach because, on the surface, it seemed to more closely resemble the learning process of early bilinguals, who are acquiring their vocabulary for both languages simultaneously. In contrast, studies that ask participants to connect novel words with previously stored translation equivalents would seem to be a better example of the type of learning encountered by later L2 learners, who have already acquired one linguistic system

before attempting to learn a second. As a result, L2 learners may have developed particular expertise with this form of learning that may not generalize more broadly. It is, therefore, informative to know how bilingual performance compares across different types of word-learning tasks.

This work also informs existing theories of how monolingual adults learn words in noisy settings. Only a limited number of studies have examined word learning in adulthood, and even fewer studies have investigated the effect that noise might have on this ability. Our data suggest that monolingual adults, who usually are relatively good (and better than bilinguals) at comprehending speech in noise, have the same difficulty as bilinguals learning new words in this type of listening condition. Follow-up work should investigate what is driving this effect.

Future work with bilinguals should also examine stream segregation abilities across tasks using a within-subjects approach. In the present set of studies, we did not explicitly test the similarities and possible differences between groups in the two experiments. While participants were all undergraduate students, who were comparable in age, gender, and self-reported language fluency, and who reportedly had normal hearing and no history of speech or attention problems, the two groups might have differed in other factors (e.g., there may have been small differences in auditory thresholds). Having the same participants complete both tasks could help confirm that the difference in performance between word recognition and word learning measures is not due to listener-specific variations in the ability to process speech in noise.

Taken together, data from our two experiments suggest that *both* linguistic experience and task demands influence the success with which listeners process speech in noise. While bilinguals might have greater difficulty compared to monolinguals identifying previously acquired words in difficult listening conditions, they do not show the same disadvantage during the encoding of new perceptual information. This might suggest that bilingualism leads to reliance on a different set of cognitive processes during stream segregation tasks or on a different weighting of available cues; bilinguals seem to approach the tasks of processing competing auditory signals slightly differently, such that advantages/disadvantages are specific to the particular type of task.

Acknowledgements

Funding

Notes

1. It should be noted that using the same sample of white noise could result in listeners learning the characteristics of the particular noise token over the course of the experiment. This could lead to improved performance on the task compared to if more randomly fluctuating noise tokens were provided throughout the experiment. Nevertheless, the same approach was used for both monolinguals and bilinguals and across experiments and is thus unlikely to result in any differences between groups or across tasks.

2.　Several studies examining word learning have implemented accuracy measures that rely on partial word-form knowledge (e.g., Meara & Ingle, 1986; Schmitt, 1998; Storkel et al., 2006; Barcroft, 2008). The reasoning behind this approach is that whole word forms are not necessarily acquired automatically, and instead the process of learning new labels might take place gradually as the different parts of the word are processed and stored "bit-by-bit" (Barcroft & Rott, 2010). Given the length of the words and the difficulty of the listening-in-noise task in the current experiment, we felt it was important to rely on a measure that would be sensitive to the partial learning of information, rather than an all-or-nothing accuracy score. This approach also allowed us to set a "minimum criteria" of what was considered to be a correct production, and examine whether participants actually learned at least an approximation to the word as a whole.

ORCID iD

Giovanna Morini ⓘD https://orcid.org/0000-0002-0034-8213

References

Barcroft, J. (2008). Second language partial word form learning in the written mode. *Estudios de Linguistica Aplicada*, *26*, 53–72.

Barcroft, J., & Rott, S. (2010). Partial word form learning in the written mode in L2 German and Spanish. *Applied Linguistics*, *31*, 623–650.

Beaman, C. P. (2005). Auditory distraction from low-intensity noise: A review of the consequences for learning and workplace environments. *Applied Cognitive Psychology*, *19*(8), 1041–1064.

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, *10*, 89–129.

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

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America*, *112*, 272–284.

Carhart, R., Tillman, T. W., & Greetis, E. S. (1969). Perceptual masking in multiple sound backgrounds. *Journal of the Acoustical Society of America*, *45*, 694–703.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*, 257–271.

Crowder, R. G. (1976). *Principles of learning and memory*. Mahwah, NJ: Lawrence Erlbaum.

Ecke, P. (2004). Words on the tip of the tongue: A study of lexical retrieval failures in Spanish–English bilinguals. *Southwest Journal of Linguistics*, *23*, 33–63.

Florentine, M. (1985a). Non-native listeners' perception of American-English in noise. *Proceedings of Inter-noise '85*, 1021–1024.

Florentine, M. (1985b). Speech perception in noise by fluent, non-native listeners. *Proceedings of the Acoustical Society of Japan*, H-85–16.

Florentine, M., Buus, S., Scharf, B., & Canévet, G. (1984). Speech perception thresholds in noise for native and non-native listeners. *Journal of the Acoustical Society of America*, *75*, Suppl. 1, s84.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*, 105–132.

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

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory and Cognition*, *33*, 1220–1234.

Grondin, R., Lupker, S. J., & McRae, K. (2009). Shared features dominate semantic richness effects for concrete concepts. *Journal of Memory and Language*, *60*, 1–19.

Grosjean, F. (1997). Processing mixed language: Issues, findings, and models. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 225–253). Hillsdale, NJ: Erlbaum.

Gupta, P. (1995). *Word learning and immediate serial recall: Toward an integrated account*. PhD thesis, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology, 56A*, 1213–1236.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin and Review*, *18*, 625–658.

Hirsch, I., Davis, H., Silverman, S., Reynolds, E., Eldert, E., & Benson, R. (1952). Development of materials for speech audiometry. *Journal of Speech and Hearing Disorders*, *17*, 321–337.

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

Johnstone, P. M., & Litovsky, R. Y. (2006). Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *Journal of the Acoustical Society of America*, *120*, 2177–2189.

Kaushanskaya, M. (2012). Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition*, *15*, 470–489.

Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*, 705–710.

Kaushanskaya, M., & Rechtzigel, K. (2012). Concreteness effects in bilingual and monolingual word learning. *Psychonomic Bulletin and Review*, *19*, 935–941.

Krizman, J., Bradlow, A. R., Lam, S. S. Y., & Kraus, N. (2017). How bilinguals listen in noise: Linguistic and non-linguistic factors. *Bilingualism: Language and Cognition*, *20*, 834–843.

Leach, L., & Samuel, A. G. (2007). Lexical configuration and lexical engagement: When adults learn new words. *Cognitive Psychology*, *55*, 306–353.

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

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*, 940–967.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*, 953–978.

Mayo, L., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research*, *40*, 686–693.

Meador, D., Flege, J. E., & Mackay, I. A. (2000). Factors affecting the recognition of words in a second language. *Bilingualism: Language and Cognition*, *3*, 55–67.

Meara, P., & Ingle, S. (1986). The formal representation of words in an L2 speaker's lexicon. *Second Language Research*, *2*, 160–71.

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology, 48A*, 98–107.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America*, *97*, 593–608.

Postman, L. (1976). Interference theory revisited. In J. Brown (Ed.), *Recall and recognition*. New York, NY: Wiley.

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology*, *16*, 635–645.

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*, 465–485.

Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, *48*, 281–317.

Storkel, H. L., Armbrüster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research*, *49*, 1175–92.

Tabri, D., Chacra, K. M. S. A., & Pring, T. (2011). Speech perception in noise by monolingual, bilingual and trilingual listeners. *International Journal of Language and Communication Disorders*, *46*, 411–422.

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

Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Behavior*, *2*, 440–445.

Appendix A: Distribution of the non-English language for bilingual participants.

Experiment 1		Experiment 2	
Language background	Number of participants	Language background	Number of participants
Bangla	1	Albanian	1
Cantonese	3	Arabic	1
Farsi	2	Bangla	1
French	3	Cantonese	4
Greek	1	Czech	1
Gujarati	1	French	1
Hebrew	1	Hindi	2
Hindi	2	Macedonian	1
Korean	1	Mandarin	1
Mandarin	4	Portuguese	2
Polish	1	Shona	1
Russian	1	Spanish	6
Sinhala	1	Tamil	4
Spanish	4	Twi	2
Tagalog	1	Urdu	1
Tamil	3	Vietnamese	1
Telugu	1	Wolof	1
Urdu	1	Yoruba	1

Appendix B: Word lists from the CID W-22 used in Experiment 1

List 1	List 2	List 3	List 4
an	your	bill	all
yard	bin	add	wood
carve	way	west	at
us	chest	cute	where
day	then	start	chin
toe	ease	ears	they
felt	smart	tan	dolls
stove	gave	nest	so
hunt	pew	say	nuts
ran	ice	if	ought
knees	odd	out	in
not	knee	lie	net
mew	move	three	my
low	now	oil	leave
owl	jaw	king	of

Appendix B. (Continued)

List 1	List 2	List 3	List 4
it	one	pie	hang
she	hit	he	save
high	send	smooth	ear
there	else	farm	tea
earn	tear	this	cook
twins	does	done	tin
could	too	use	bread
what	cap	camp	why
bathe	with	wool	arm
ace	air	are	yet
you	and	aim	darn
as	young	when	art
wet	cars	book	will
chew	tree	tie	dust
see	dumb	do	toy
deaf	that	hand	aid
them	die	end	than
give	show	shove	eyes
true	hurt	have	shoe
isle	own	owes	his
or	key	jar	our
law	oak	no	men
me	new	may	near
none	live	knit	few
jam	off	on	jump
poor	ill	is	pail
him	rooms	raw	go
skin	ham	glove	stiff
east	star	ten	can
thing	eat	dull	through
dad	thin	though	clothes
up	flat	chair	who
bells	well	we	bee
wire	by	ate	yes
ache	ail	year	am

Appendix C: Non-word list used in Experiment 2.

Target words	Foil words	
bibukadir	batibagobe	jugugochop
chechepatile	bejachubuf	jupobegif
dekabagom	bidutobom	kajibobole
jupitoduce	bopapotup	kepechepof
kutopechef	butitogot	ketijopert
tabetogobe	chabugupun	kigatachaf
	chachugokal	kijadachop
	chagakejis	kikuchepeve
	chakobekole	kobadubore
	chetagateg	kudopotire
	chetukopime	kugojochame
	chibutochid	pagijochod
	chichetitex	pebedachupe
	chobedakile	pegakebun
	chokukebare	pikicheburt
	dakujugir	popobiduce
	datipekode	puchakipin
	dijigochut	puchupipent
	dochagagize	pukabagade
	dokidutal	pukajugife
	duchipijar	tapugukot
	dudekajup	tatepagist
	gadipagale	tebukibob
	gedechobege	techibogule
	gibipagunt	titagubet
	gipabijem	tochupopabe
	gipechejel	tojachigave
	gokepetor	tojukided
	gopokebige	tokukuchive
	gugekipert	topejabot
	gujukojert	tuchichedupe
	jaguchajont	tudidojope
	jedokibile	tugukuchere
	jidakopuge	tukabagag
	jojujugif	tupadubuce
	jugatojert	tutibubith
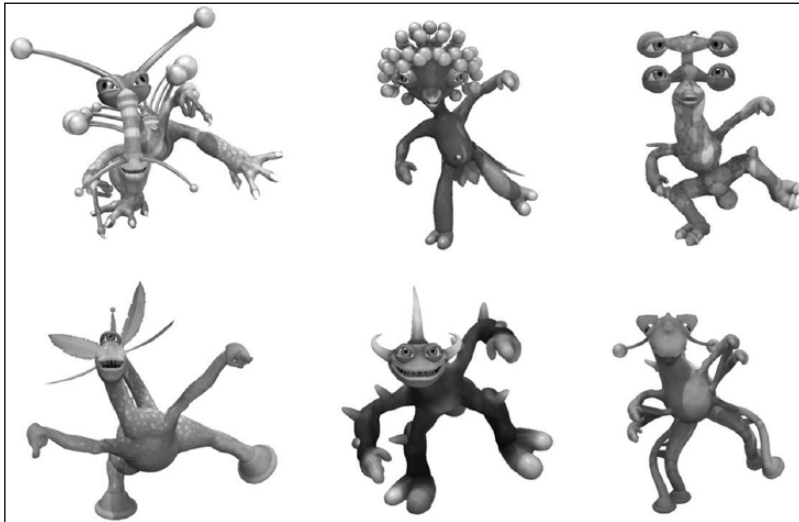
Appendix D: Novel creatures used in Experiment 2.