RESEARCH NOTE

Native-language benefit for understanding speech-in-noise: The contribution of semantics*

NARLY GOLESTANI
Institute of Cognitive Neuroscience
University College London
& University Medical School, Geneva
STUART ROSEN
Department of Phonetics and Linguistics
University College London
SOPHIE K. SCOTT
Institute of Cognitive Neuroscience
University College London

Bilinguals are better able to perceive speech-in-noise in their native compared to their non-native language. This benefit is thought to be due to greater use of higher-level, linguistic context in the native language. Previous studies showing this have used sentences and do not allow us to determine which level of language contributes to this context benefit. Here, we used a new paradigm that isolates the SEMANTIC level of speech, in both languages of bilinguals. Results revealed that in the native language, a semantically related target word facilitates the perception of a previously presented degraded prime word relative to when a semantically unrelated target follows the prime, suggesting a specific contribution of semantics to the native language context benefit. We also found the reverse in the non-native language, where there was a disadvantage of semantic context on word recognition, suggesting that such top–down, contextual information results in semantic interference in one's second language.

Introduction

In bilingual populations, residual differences can be seen between native and non-native language processing, even if both languages are spoken well. Thus, it is more difficult for non-native listeners to understand speech in a noisy background than it is for native listeners (Nabelek and Donahue, 1984; Takata and Nabelek, 1990; van Wijngaarden, Steeneken and Houtgast, 2002). This advantage of the native language for speech recognition under poor listening conditions exists even if the second language is spoken very proficiently (Florentine, 1985a, b; Takata & Nabelek, 1990; Mayo, Florentine and Buus, 1997). Furthermore, this native language advantage arises from better use of CONTEXTUAL INFORMATION in the first (native) compared to the second (non-native) language (Florentine, 1985a; Mayo et al., 1997). This was demonstrated using the so-called Speech Perception in Noise (SPIN) sentences (Kalikow, Stevens and Elliott, 1977; Bilger, Nuetzel, Rabinowitz and Rzeczkowski, 1984), in which participants hear sentences in which the final word is of high or of low predictability. The sentences are embedded in different levels of noise, which allows the

investigation of how degrading the 'bottom–up' speech input interacts with 'top–down' knowledge of higher-level linguistic information. These studies showed that although listeners are adversely affected by noise in their native and non-native languages, they are better able to use top–down resources such as contextual information in their native language to predict the identity of a word, and compensate for the loss of the bottom–up information. The use of sentences, however, does not allow one to determine which level of language (e.g. semantics, syntax, prosody) the native language benefit arises from. Specifically, although the SPIN material involves manipulating the SEMANTIC predictability of the final word in sentence-level stimuli (e.g. Bradlow and Alexander, 2007), it is likely that higher-level information other than semantics also contained in the sentences contributes to the ability to predict the identity of the final word – that is, that semantic but also syntactic and prosodic information contained in the sentences implicitly or explicitly help to identify the last word of the sentences.

Here, we propose to extend existing research showing that native listeners benefit from linguistic context under adverse listening conditions by using a new paradigm which will allow to specifically isolate the SEMANTIC 'higher-level' component of language, and to determine whether this level (at least) contributes to the native language advantage when listening to degraded speech. We used an AUDITORY version of the retroactive word

Address for correspondence:
Narly Golestani, Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK
n.golestani@ucl.ac.uk

priming paradigm, which involves the use of word pairs which are either semantically related or unrelated. This paradigm has previously been used in the VISUAL modality to show the importance of semantic context in word processing. Bernstein and colleagues (Bernstein, Bissonnatte, Vyas and Barclay, 1989) showed that the identification of visually masked primes is better when they are followed by semantically related targets than if they are presented alone, constituting an example of 'RETROACTIVE PRIMING'. Conversely, performance is worse when primes are followed by semantically unrelated targets than when they are presented alone (Bernstein et al., 1989). These findings are very interesting because they demonstrate that the ability to identify a degraded word (i.e. to make the most of sub-optimal bottom–up input) depends on higher-level (i.e. semantic) context induced by the identity of a word presented later in time. They demonstrate an interaction between how top–down, higher-level context can interact with and influence bottom–up, low-level processing of speech stimuli.

In the present study, we have adapted the retroactive priming paradigm to the auditory modality to evaluate the differential effect of semantic context on the intelligibility of words embedded in different levels of noise in bilinguals. By using a paradigm which involves the use of word pairs only, we effectively isolate the possible contribution of semantic context in driving the native language benefit during the perception of speech-in-noise.

Native French speakers who were non-proficient, 'late' learners of English, were tested using this paradigm in both their native (French) and in their non-native (English) languages. We predicted that overall performance would improve with higher signal-to-noise ratios (SNR) – that is, performance would be better with lower compared to higher noise levels, and in the first compared to the second language. Further, we predicted that we would find a benefit of context (i.e. better performance during related compared to unrelated trials) in the native compared to the non-native language.

Method

Participants

Nine native French speakers (4 men), who started to learn English in school 'late', after the age of 11, and who spoke English moderately fluently, participated in the study. Participants had a homogeneous language background; all had learned a second language (one of English, German, or Spanish) in school from the ages of 11–18 and a third language (one of English, German, or Spanish) from the ages of 13–18. None spoke a second or third language proficiently, and none had been regularly exposed to a language other than French before the age of 11. All participants gave informed consent to participate in the study, which was approved by the regional ethical committee.

Materials

French semantically related and unrelated word pairs were selected from the database by Ferrand and Alario (Ferrand and Alario, 1998), and English word pairs were selected from the University of South Florida Free Association Norms (Nelson, McEvoy and Schreiber, 1998).

In each language, two sets of stimuli were generated (i.e. two sets of 520 word pairs) such that in list 1, a specific prime was followed by a related target, and that in list 2, it was followed by an unrelated target (see Table 1, lists 1 and 2 for examples of how a stimulus item was constructed, and see Appendix for full list of materials).[1] The lexical frequency and number of syllables of related and unrelated 'targets' were matched across lists (within each language) because more frequent (or common) words are more likely to be recognized than are less common ones (Bradlow and Pisoni, 1999), and because pilot testing revealed that longer words are more likely to be recognized than shorter ones, perhaps because in longer utterances there is usually more phonetic information that survives the noise. The number of syllables of primes were also matched across languages. Word frequency information was taken from the 'Lexique 3' database for the French words (New, Brysbaert, Veronis and Pallier, 2007; www.lexique.org), and from the English Lexicon Project for the English words (Balota et al., 2007). See Table 2 for summary of stimulus information, including the mean number of syllables and the mean word log frequency for the different stimulus types in English and in French. Half of the participants were presented with one list (e.g. list 1 in Table 1) and the other half with the other list (e.g. list 2 in Table 1). This was done to ensure that results are not due to stimulus-specific effects but rather to the manipulations of interest. In addition, two versions of each stimulus list were generated such that in one version, a particular prime was embedded in a certain level of noise or in no noise, whereas in the other version, it was embedded in a different level of noise or in no-noise (see Table 1, lists 3 and 4). In turn, half of the participants tested in each subgroup above were tested with one version (e.g. list 3) and half with the other (e.g. list 4). This was done to control for differential phoneme recognition effects in noise as certain speech sounds survive noise better than others (Miller and Nicely, 1955; Boothroyd, Mulhearn, Gong and Ostroff, 1996). Thus, we wanted to control for the fact that some prime words might be more easily identified at a particular noise level than others due to their different constituent speech sounds. For example, the word "artist" might be more likely to be heard over a certain level of noise than the word "apple". By embedding "artist" in a higher SNR level and "apple"

---

[1] The appendix mentioned in this article is available on the Journal's website as Supplementary Materials accompanying the present article (see journals.cambridge.org/bil, vol. 12 (3)).

Table 1. *Examples of sample items in four stimulus lists. In the first part of the table, (1) and (2) are examples of how stimuli were constructed (i.e. by semantically matching foils with primes) and (3) and (4) are examples of how stimuli were NOT constructed (i.e. by phonetically matching foils with primes).*

| List 1 | | List 2 | |
Related target		Unrelated target	
(1a)	(1b)	(2a)	(2b)
prime – target	**prime – foil**	**prime – target**	**prime – foil**
parrot (SNR –7) *– bird*	parrot – eagle	*parrot* (SNR –7) *– cake*	parrot – eagle
(3a)	(3b)	(4a)	(4b)
prime – target	**prime – foil**	**prime – target**	**prime – foil**
parrot (SNR –7) *– bird*	parrot – carrot	*parrot* (SNR –7) *– cake*	parrot – carrot
List 3		List 4	
Related target		Unrelated target	
prime – target	**prime – foil**	**prime – target**	**prime – foil**
parrot (SNR –5) *– bird*	parrot – eagle	*parrot* (SNR –5) *– cake*	parrot – eagle

SNR = speech-to-noise ratio

Note 1. Participants were tested with list 1, list 2, list 3, or list 4, and an equal number of participants was tested with each list. Words in italics represent auditorily presented words (stimulus words), and words not in italics represent visually presented words (test words; participants have to select the word that represents the first auditory, often degraded "prime" word). Lists 1 versus 2 (or lists 3 versus 4) match for 'semantic relatedness' (of the target to the prime), and lists 1 versus 3 (or lists 2 versus 4) match for the level of noise used to degrade the prime.

Note 2. Foils were semantically rather than phonetically matched with the prime words for the following reason. Had we matched the foils phonetically (i.e. by making them rhyme with the primes; see (3) and (4) in Table 1), then ONLY on related trials (see (3) in Table 1) had participants not heard the prime (in this case, "parrot") due to noise; they could then later nonetheless select the word "parrot" from the visually presented words ((3b) in Table 1) since parrot is semantically related to audible "target" word that they have just heard, in this case, "bird" (see (3a)). Importantly, had we matched foils phonetically with the primes, such a strategy would only work for related trials since on the unrelated ones (see (4) in Table 1), neither the prime nor the foil (see (4b)) would be semantically related to the audible target (4a), in this case, "cake". Such a differential response strategy during related compared to unrelated conditions, especially beneficial on trials where the prime is not heard, would have biased our results such that participants would have more often correctly selected the prime during related compared to unrelated trials, biasing their performance in favor of the predicted relatedness effect, but due to semantic relatedness of the visual rather than auditory prime to the target.

Thus, instead of phonetically matching foils with primes, we matched them semantically (see (1) and (2) in Table 1). Here, it can be seen that given that (a) the foil (eagle) is semantically matched with the prime (parrot) both in the related (1b) and unrelated (2b) conditions, and that (b) BOTH are either semantically RELATED to the target in the related condition (in this case "bird", see (1)), or both are semantically UNRELATED to the target in the unrelated condition (in this case "cake", see (2)), participants cannot use the strategy (of using semantic relatedness of the prime to the audible auditory target) described above.

in lower SNR level in one version, and vice versa in the other version, we controlled for such phonetically-driven biases in performance.

The visual foils used in the recognition phase of the task were semantically rather than phonetically matched with the prime word (i.e. the degraded one, to be recognized); see Note 2 to Table 1 for further explanations. Visual foils were also matched with primes with respect to number of syllables; this was done in order to ensure that participants do not use this information to recognize the prime from the foil.

The English and French words were digitally recorded by a multilingual female speaker in an anechoic chamber using a sampling rate of 44.1 kHz with 16 bit quantization. The microphone was positioned 30 cm from the speaker's mouth, at 15 degrees to the mid-sagittal line. The final set of stimuli was created off-line by editing the words at zero-crossings before and after each word. Recordings were normalized with respect to root-mean-squared amplitude and had an average duration of 1.1 seconds.

Several behavioral pilot studies were conducted involving embedding the primes in different SNR levels.

Table 2. *Stimulus information: word log frequency and number of syllables of stimuli in (a) English and (b) French.*

	Prime	Target list 1	Target list 2	Foil
(a) English				
Word log frequency	3.92	4.25	4.19	4.16
Number of syllables	1.65	1.43	1.44	1.57
(b) French				
Word log frequency	3.11	3.41	3.53	3.29
Number of syllables	1.71	1.54	1.47	1.67

Note. Word (log) frequency information for English stimuli is based on per 120 million observations (Balota et al., 2007). Thus, the (log) frequency information obtained for the French stimuli from Lexique 3, which was originally based on per one million observations (New et al., 2007), has been adjusted as to reflect per 120 million observations.

It was found that using SNR levels of −7 (highest level of noise), −6, −5, and −4 dB (lowest noise level) was optimal in terms of a) yielding performance that was not at ceiling nor at floor across participants, and b) demonstrating the predicted effect of semantic context (i.e. of relatedness) in the native language of participants. In other words, we were able to see an effect of relatedness (better performance during the related compared to the unrelated trials) on at least one of the SNR levels for each participant. We also included a no-noise condition. We used speech-shaped noise, which approximates the average long term spectrum of the speech of an adult male, and which has a similar effect to the masking produced by a number of other speakers speaking at the same time ("multi-speaker babble").

Procedure

The following procedure was implemented for each participant during a total of four testing sessions in both French and in English. There was a short rest between sessions, and language was alternated between the sessions, while counterbalanced for starting language across participants. Eprime was used to present visual and auditory stimuli, and to collect responses.

For each trial, participants heard a pair of words, half of which were semantically related and half of which were semantically unrelated. The first, 'prime' word was degraded by being presented in different levels of noise (SNR levels: −7, −6, −5, −4 dB, and no-noise), whereas the second, 'target' word was always clearly audible. Immediately after the offset of the target word participants saw two visually presented words, one being the prime and the other being a semantically related foil, and were required to decide which of the two visually presented

words corresponded to the prime by making a button press response (left button for word on left side of screen, and right button for word on right side of screen). They had 1.5 seconds to respond. Each trial lasted four seconds.

The following conditions were included: language (French = native language and English = non-native language), semantic context (related and unrelated), and signal-to-noise ratio (no-noise, −7, −6, −5, −4), resulting in a total of 20 conditions. In each language, there were 26 related and 26 unrelated word pairs at each of the five SNR levels (i.e. there were 26 stimuli per condition), resulting in a total of 130 related and 130 unrelated word pairs per language. Each participant therefore performed a total of 520 trials. The language and SNR conditions were blocked into miniblocks of five trials each, and 'relatedness' was mixed within miniblocks in order to ensure that participants would not adopt different response strategies across relatedness conditions. Testing lasted about 35 minutes per participant.

Results

Figures 1a and 1b show the mean percent correct scores for the English and the French conditions separately.

A three-way (language by relatedness by SNR) repeated-measures analysis of variance was performed on accuracy measures, excluding the no-noise condition which was at or near ceiling in any case. There were main effects of language ($F(1,8) = 9.9$, $p < .05$), showing that performance was better in the native (French) compared to the non-native (English) language, and of noise level or SNR ($F(4,32) = 15.6$, $p < .001$), showing that performance was better with higher compared to lower SNR levels. More specifically, the 'noise level' effect was explained by a strong linear trend ($F(1,8) = 30.7$, $p < .001$), demonstrating that higher SNRs were associated with proportionately better performance. The main effect of relatedness was not significant ($F(1,8) = 0.9$, $p > .05$), but there was a significant language by relatedness interaction ($F(1,8) = 26.2$, $p < .001$). Figure 2 shows the interaction pooled over noise levels (excluding the no-noise condition). From this figure and also from Figure 1a, it can also be seen that unexpectedly, it appears as though the reverse holds in English, where performance appears to be better on the unrelated compared to the related trials. Post-hoc tests confirmed this. In French, performance is better on the related compared to unrelated trials ($F(1,8) = 21.0$, $p < .01$), and in English, performance is better on unrelated compared to related trials ($F(1,8) = 7.6$, $p < .05$).

Discussion

Using a novel paradigm that isolates the semantic level of speech (i.e. word pairs rather than sentences), we
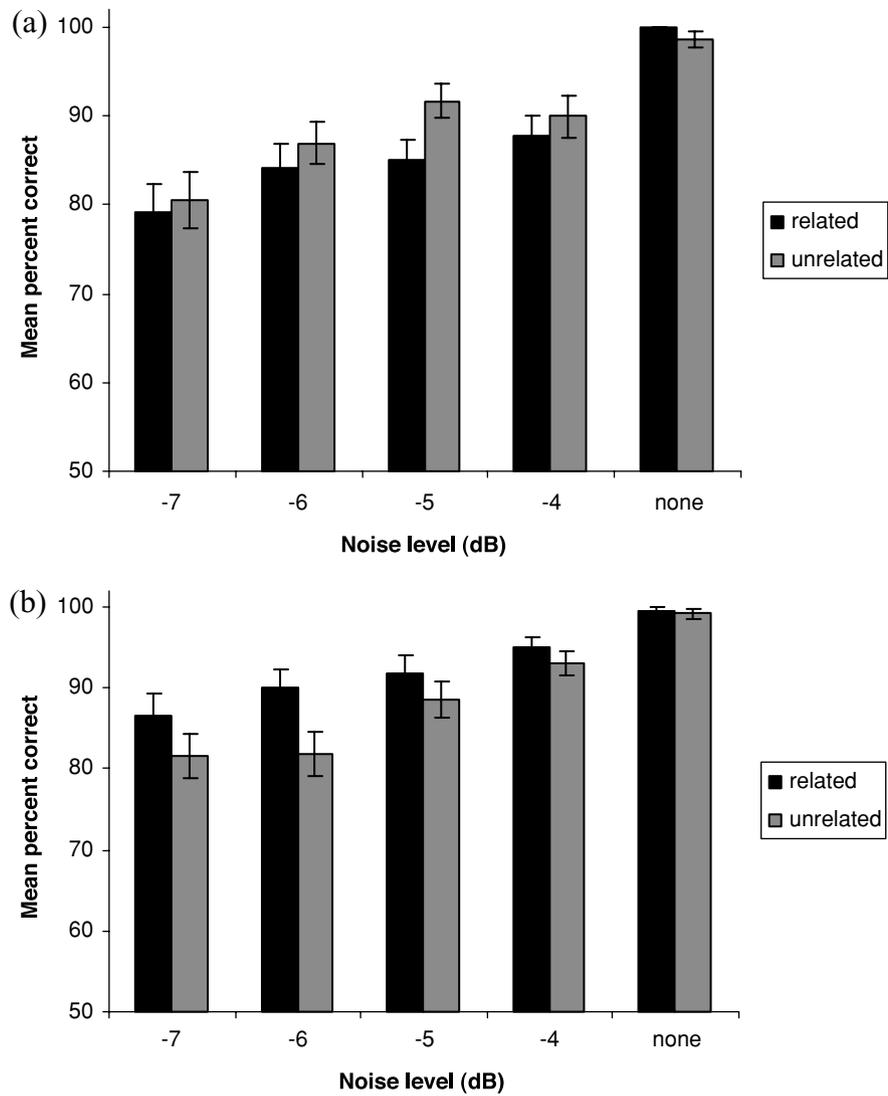
Figure 1. Mean percent correct scores in English (a) and French (b). Error bars indicate standard error of the mean (SEM).
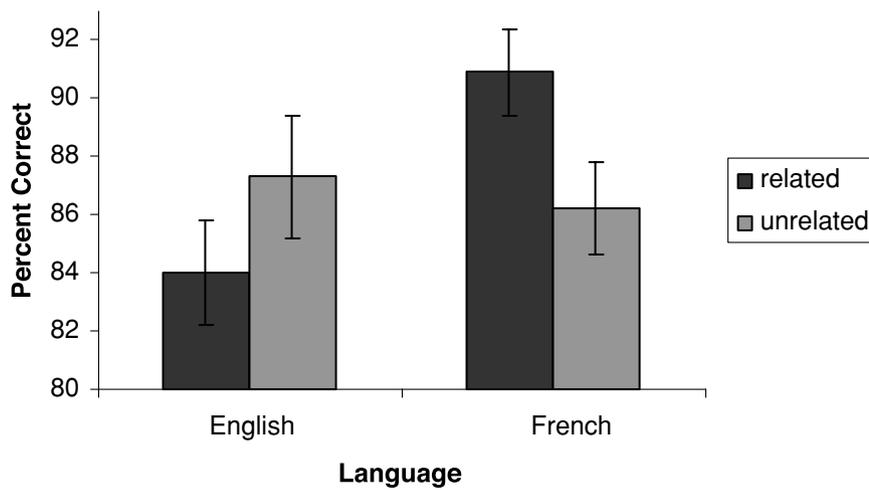

Figure 2. Language by relatedness interaction. Error bars indicate standard error of the mean (SEM).

have shown that semantic context contributes to the intelligibility advantage of the native over the second language when listening to speech-in-noise. Specifically, we found that in the native but not in the non-native language, a semantically related target facilitates the perception of a previously presented degraded prime relative to when a semantically unrelated target follows the prime (i.e. there is a retroactive priming effect in the native but not in the non-native language). This suggests that the semantic level of language, specifically, contributes to the native language advantage for speech-in-noise. Previous behavioral work has also shown a differential effect of linguistic context when processing speech-in-noise in the native versus the non-native language in bilinguals, but only at the sentence level (Florentine, 1985a, b; Mayo et al., 1997). For example, it has been shown that in monolinguals but not in late bilinguals, higher noise is associated with a greater context benefit (Mayo et al., 1997). Given that previous studies have used sentences, it has not until now been known which level(s) of language specifically contribute to this native language advantage (syntax, semantics, or prosody). We show for the first time that semantic information specifically contributes to this effect, though not necessarily exclusively. In addition, we have extended previous semantic retroactive priming findings (see Bernstein et al., 1989) to the auditory modality (in the native language only), by degrading, or 'masking' spoken words by presenting them in different levels of noise. In other words, we have shown that the effect of semantic context on word perception is modality independent, and that this top–down modulation of bottom–up processing thresholds occurs regardless of whether the degraded bottom–up speech input is visual or auditory. Last, it is unlikely that the observed effects are occurring at the response stage since we ensured that, at this stage, visually presented foils are always semantically matched to the visually presented prime.

Interestingly, in a related functional magnetic resonance imaging study using the same paradigm with a new group of bilinguals, we found a predicted influence of semantic relatedness in the native and not in the non-native language in 'higher-level' components of the language, attention and executive brain networks during top–down, context-driven processing, and in 'lower-level' parts of the language system during bottom–up, stimulus-driven processing. This latter study thus provides complementary evidence for the neural basis of the native language semantic context benefit during the perception of speech-in-noise (Golestani, Obleser and Scott, 2009).

We also found that although the context by language interaction was driven by the predicted semantic context advantage in the native language (French), there was also, in the non-native language (English), a significant DISADVANTAGE of semantic context on word recognition performance. In other words, the native French speakers

do better on UNRELATED compared to related trials when hearing degraded words in English. This unexpected result is interesting, and could be explained in the following way. We speculate that in the participants' less fluent language (English), hearing degraded words followed by semantically related words results in semantic interference. This interpretation is consistent with previous findings showing that in people learning English as a second language, semantic relatedness of test items interferes with performance on a rhyming test (such tests are typically used to assess reading readiness; Moreira and Hamilton, 2006). Here, individuals are shown four pictures representing words; the first picture represents the target word, and they are asked to choose, from the three other pictures, the one that rhymes with the target. In people who speak English non-natively, performance is poorer than in native speakers mainly because they tend to select the word that is semantically related to the target, rather than the one that rhymes with it. This suggests that relatively greater semantic processing tends to interfere with rhyme task performance in non-native compared to in native speakers. In our study, individuals are not asked to attend to the second, related or unrelated 'target' word. They are not told anything about it at all, but are rather asked to attend to the first, often degraded word, and to try to recognize it among the two visual words presented later. We speculate that as with the rhyming task, in the non-native language, semantic processing, which maybe takes place automatically and/or without conscious awareness, is somehow more prominent than in the native language (maybe due to an automatic tendency to want to translate the words into the native language), and thus fewer attentional and/or processing resources remain for performing the sometimes difficult task at hand, that being recognition of acoustically degraded words.

Interestingly, it has previously been shown that during the perception of sentences in noise, non-native listeners are better at recognizing the final word only if it is predictable and if acoustic enhancements are available (i.e. during the perception of clear speech – this refers to sentences recorded in a 'clear' speaking style, as opposed to a plain, conversational speaking style; Bradlow and Alexander, 2007). The fact that non-native listeners can make use of contextual information when presented with 'clear' speech suggests that they simply require greater signal quality in order to do so than do native listeners. In turn, given that the pattern of performance for native and for non-native listeners is the same but only with a different 'baseline' SNR or signal quality level needed to show context benefits, we speculate that native and non-native listeners have a different 'threshold' for being able to make use of higher-level, linguistic context. This threshold difference may be due to different processing resources in the native versus second

languages. Specifically, in non-fluent, late bilinguals, speech processing is much more automatic and practiced in the native language, and so when signal processing of the bottom–up input is made more challenging either by signal degradation or by poorer articulation and other such factors, processing in the non-native language may suffer more because attentional and processing resources reach a bottle-neck both at the higher-level linguistic and at the lower-level, speech input processing levels, whereas they become limited only at the latter in the native language.

In our study, there is likely also a contribution to the perception of speech-in-noise at the phonetic level, especially since we used semantically and not phonetically matched foils. For example, it may be easier to recognize the degraded word "grass" from the foil "green" compared to recognizing the degraded word "yard" from the foil "green" even though both pairs of words are semantically related since the sound "s" in "grass" might survive noise better than do any of the sounds in the word "yard" (Miller and Nicely, 1955). It is noteworthy that the differential semantic context effects were strong enough to be detected over and beyond the variability in performance likely conferred by the fact that some primes contained more 'noise-robust' phonemes than others. In addition to the general effects of phonetic contribution to the perception of speech-in-noise, it is possible that phonetic information modulates performance differently in the native versus the non-native language. In other words, it might be the case that the native French speakers we tested in this study were better able to 'extract' phonetic information from noise in the French compared to in the English condition. This would be predicted by the findings of Cutler, Weber, Smits and Cooper, (2004), who showed that native listeners were better at identifying phonemes (CV and VC syllables) embedded in noise than non-native listeners. In this study, however, the native language advantage for phonetic perception did not increase as a function of the noise level, suggesting that although the native language advantage may be at least in part driven at the phonetic level (since there is a native language advantage using stimuli that effectively isolate the phonetic level of speech such as CV and VC syllables), it is not exclusively phonetically driven. This is consistent with our results, which suggest that semantics specifically contributes to the native language advantage for speech-in-noise intelligibility. Our findings do not, however, exclude the possibility that other levels of speech (e.g. prosody, syntax) also contribute to the perception of speech-in-noise. Further, we speculate that the semantic and phonetic levels of speech may interact in contributing to the native language advantage for the perception of speech-in-noise. In other words, it is possible that higher-level, semantic context modulates a lower-level perceptual threshold for phonetic perception (e.g. the word "grass" followed by "green" may survive noise more than would the word "brass" followed by "yard" due to the former but not the latter pair being semantically related). In other words, semantic context may actively interact with the extent to which acoustic cues can be utilized to assist the perception of speech-in-noise. Conversely, phonetic information likely also modulates the strength of the relatedness effect. An interaction between the semantic and phonetic levels of speech is supported by research showing that auditory semantic priming effects diminish if phonetic segments of the prime word are acoustically distorted (Andruski, Blumstein and Burton, 1994; Utman, Blumstein and Burton, 2000).

In sum, our findings shed light on the differential interaction between sensory (bottom–up) and cognitive (top–down) processing and resources in the native and non-native languages in bilinguals. These results have important implications for understanding language and communication in bilinguals that are generalizable to real life situations, where speech and communication often take place in noisy external and internal environments. This work also has implications for bilingual individuals with severe hearing impairments, or for cochlear implant users who do not hear speech as clearly as do individuals who have normal hearing. As seen from the results of previous work, generalizability to real-life noisy internal and external environments may be especially important in the non-native language in late bilinguals, since it appears that these individuals do not benefit as much from context as do individuals in their native language in non-optimal listening environments. In a second language, an individual may not have access to the same contextual information as one might have in a native language (richer vocabulary, better expectations of what a muffled word might be if heard in the context of a sentence, greater familiarity with words and with word associations). Our results support the idea that in one's native language, semantic context is important in increasing speech intelligibility in noisy environments.

References

Andruski, J. E., Blumstein, S. E. & Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition*, 52 (3), 163–187.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, D. L., Nelson, D. L., Simpson, G. B. & Treiman, R. (2007). The English Lexicon Project. *Behavioral Research Methods*, 39 (3), 445–459.

Bernstein, I. H., Bissonnette, V., Vyas, A. & Barclay, P. (1989). Semantic priming: Subliminal perception or context? *Perception & Psychophysics*, 45 (2), 153–161.

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M. & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research*, 27 (1), 32–48.

Boothroyd, A., Mulhearn, B., Gong, J. & Ostroff, J. (1996). Effects of spectral smearing on phoneme and word recognition. *Journal of the Acoustical Society of America*, 100 (3), 1807–1818.

Bradlow, A. R. & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *Journal of the Acoustical Society of America*, 121 (4), 2339–2349.

Bradlow, A. R. & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *Journal of the Acoustical Society of America*, 106 (4.1), 2074–2085.

Cutler, A., Weber, A., Smits, R. & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America*, 116 (6), 3668–3678.

Ferrand, L. & Alario, X. (1998). Normes d'associations verbales pour 366 noms d'objets concrets. *L' Année Psychologique*, 98, 689–739.

Florentine, M. (1985a). Speech perception in noise by fluent, non-native listeners. Presented at the Acoustical Society of Japan.

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

Golestani, N., Obleser, J. & Scott, S. K. (2009). Bilingual speech perception in the brain: Complementary involvement of higher- and lower-level components of the language network during context- versus stimulus-driven processing. Ms., University College London. [Submitted]

Kalikow, D. N., Stevens, K. N. & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, 61 (5), 1337–1351.

Mayo, L. H., Florentine, M. & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research*, 40 (3), 686–693.

Miller, G. A. & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, 27 (2), 338–352.

Moreira, S. & Hamilton, M. (2006). Goats don't wear coats: An examination of semantic interference in rhyming assessments of reading readiness for English language learners. *Bilingual Research Journal*, 30 (2), 547–557.

Nabelek, A. K. & Donahue, A. M. (1984). Perception of consonants in reverberation by native and non-native listeners. *Journal of the Acoustical Society of America*, 75 (2), 632–634.

Nelson, D. L., McEvoy, C. L. & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://w3.usf.edu/FreeAssociation/ (retrieved May 16, 2009).

New, B., Brysbaert, M., Veronis, J. & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, 28 (4), 661–677.

Takata, Y. & Nabelek, A. K. (1990). English consonant recognition in noise and in reverberation by Japanese and American listeners. *Journal of the Acoustical Society of America*, 88 (2), 663–666.

Utman, J. A., Blumstein, S. E. & Burton, M. W. (2000). Effects of subphonetic and syllable structure variation on word recognition. *Perception & Psychophysics*, 62 (6), 1297–1311.

van Wijngaarden, S. J., Steeneken, H. J. & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native talkers. *Journal of the Acoustical Society of America*, 112 (6), 3004–3013.