

# Examining talker effects in the perception of native- and foreign-accented speech

Conor T. M<sup>c</sup>Lennan · Julio González

© Psychonomic Society, Inc. 2012

**Abstract** Understanding the circumstances under which talker (and other types of) variability affects language perception represents an important area of research in the field of spoken word recognition. Previous work has demonstrated that talker effects are more likely when processing is relatively slow (M<sup>c</sup>Lennan & Luce, 2005). Given that listeners may take longer to process foreign-accented speech than native-accented speech (Munro & Derwing, *Language and Speech*, 38, 289–306 1995), talker effects should be more likely when listeners are presented with words spoken in a foreign accent than when they are presented with those same words spoken in a native accent. The results of two experiments, conducted in two different countries and in two different languages, are consistent with this prediction.

**Keywords** Spoken word recognition · Speech perception · Priming

Variability can slow recognition of written (Burgund & Marsolek, 1997) and spoken (Bradlow, Nygaard, & Pisoni, 1999) words, supporting theoretical positions with specific representations (Goldinger, 1998). However, variability does not always affect word recognition (M<sup>c</sup>Lennan &

Luce, 2005), supporting theoretical positions with abstract representations (TRACE; McClelland & Elman, 1986).

Consistent with the phonetic-relevance hypothesis (Sommers & Barcroft, 2006), some types of variability are more likely to affect spoken word recognition (Bradlow et al., 1999). Determining which types of surface and allophonic (Luce, M<sup>c</sup>Lennan, & Charles-Luce, 2003; M<sup>c</sup>Lennan, Luce, & Charles-Luce, 2003, 2005) variability are more likely to affect spoken word recognition is an important area of research.

Since there is evidence for both abstract and specific representations, Luce and M<sup>c</sup>Lennan (2005) (see also Luce & Lyons, 1998) suggested that variability might be more likely to affect spoken word recognition at various points during processing. M<sup>c</sup>Lennan and Luce (2005) subsequently provided evidence that abstract representations are more likely to affect early processing and specific representations are more likely to affect later processing. These authors used a long-term repetition-priming paradigm in which they presented listeners with two blocks of spoken words (primes and targets). The target words were either repeated or new. Repeated words were either spoken by the same (match) or by a different (mismatch) talker than the prime words. Crucially, the magnitude of specificity (MOS) (i.e., the advantage for repeated words spoken by the same talker relative to repeated words spoken by a different talker) was more robust during later processing.

MOS was significant in the slower (delayed shadowing; hard lexical decision), but not in the faster (speeded shadowing; easy lexical decision) tasks. The only difference between the two shadowing tasks was that participants in delayed shadowing were instructed to delay their response until a response cue appeared (150 ms after stimulus offset). The only difference between the two lexical decision tasks was that the nonwords were

C. T. M<sup>c</sup>Lennan (✉)  
Department of Psychology, Language Research Laboratory,  
Cleveland State University,  
2300 Chester Avenue,  
Cleveland, OH 44115, USA  
e-mail: c.mclennan@csuohio.edu

J. González  
Departamento Psicología Básica, Clínica y Psicobiología,  
Universitat Jaume I,  
12071 Castellón de la Plana, Spain  
e-mail: gonzalez@psb.uji.es

unwordlike (i.e., low phonotactic probability) in the easy task and wordlike in the hard task. MOS was statistically larger in the hard tasks.

These time course results provided the motivation for the present investigation. Because listeners take longer to recognize words spoken with a foreign accent (Munro & Derwing, 1995), the prediction based on the time-course hypothesis is that talker mismatches should be more likely to affect recognition when words are spoken with a foreign accent.

Previous studies have examined variability using signal degradations that resulted in effortful processing and reduced accuracy (e.g., low-pass filtering, Church & Schacter, 1994; or white noise, Goldinger, 1996). The studies by Goldinger (1996) and Luce and Lyons (1998) were among the first to report reaction times (RTs); the previous studies had focused on accuracy. One aim of the present study was to examine a milder and naturally occurring form of degradation in which accuracy is expected to be high and the main dependent variable is RT.

Two recent studies have provided additional motivation. First, Vitevitch and Donoso (2011) found more change deafness (i.e., inability to detect a talker change) in an easy than in a hard lexical decision task. Second, Mattys and Liss (2008) found greater talker effects with dysarthric speech than with healthy speech. Both studies support the time-course hypothesis and the notion that slower processing results in greater sensitivity to talker changes.

Both Vitevitch and Donoso (2011) and M<sup>c</sup>Lennan and Luce (2005) manipulated processing speed in the lab. To our knowledge, Mattys and Liss (2008) were the first to examine the time course of talker effects without slowing from lab manipulations or artificially degraded stimuli. According to the authors, “we use the term *naturally occurring degraded speech* to refer to unedited speech stimuli produced by individuals who, for whatever reason, produce speech that is degraded relative to the speech produced by healthy, native speakers” (p. 1236). Consequently, one motivation for the present study was to examine talker effects in another form of naturally occurring degraded speech. Foreign-accented speech is of particular interest because it falls within this definition of naturally occurring degraded speech and, unlike dysarthric speech, can be produced by healthy speakers.<sup>1</sup>

<sup>1</sup> There are other theoretically important distinctions between accented and dysarthric speech, including the nature of the segmental distortion (accented segments may be more phonologically canonical than segments distorted by dysarthria), as well as intelligibility differences (dysarthric speech is likely to be more difficult to perceive) and the frequencies with which listeners are likely to be exposed to dysarthric (less often) and accented (more often) speech.

## Experiment 1: English with foreign-accented speech

### Method

**Participants** A group of 72 participants from the Cleveland State University community were paid or received credit for a course requirement. The participants were right-handed (according to the Edinburgh Handedness Inventory; Oldfield, 1971) native speakers of American English with no reported history of speech or hearing disorders.

**Materials** The stimuli consisted of the words and nonwords used in M<sup>c</sup>Lennan and Luce's (2005) Experiment 2, re-recorded in English by one male and one female native Spanish speaker, both of whom had learned English as adults and spoke with a foreign accent.

The stimuli were recorded in a sound-attenuated room, low-pass filtered at 10 kHz, and edited into individual files. The mean durations for the experimental words produced by the male (583 ms) and the female (574 ms) did not differ,  $t(22) < 1.0$ ,  $p = .79$ .

**Design** The design followed that of Experiment 2 of M<sup>c</sup>Lennan and Luce (2005). Two blocks of stimuli were presented. Half of the stimuli in each block were spoken by each talker, and primes matched, mismatched, or were unrelated to the targets. The talker was the same in the match condition (e.g., *book*<sub>male</sub>, *book*<sub>male</sub>) and different in the mismatch condition (e.g., *book*<sub>male</sub>, *book*<sub>female</sub>). Words in the unrelated condition were unprimed.

Both blocks consisted of 24 trials (half nonwords). The primes consisted of eight experimental words, eight nonwords, and eight control stimuli (four of which were nonwords). The targets consisted of 12 experimental words and 12 nonwords. Eight targets matched the target stimuli, eight mismatched, and eight were controls. All of the nonwords and unrelated stimuli were fillers; the focus of the manipulations and analyses was limited to the experimental words. A 3 (Prime) × 2 (Talker) completely within-participants design was used. Across participants, each word appeared in every condition, but no participant heard more than one version of a word within a block.

**Procedure** The participants performed a lexical decision task in which they decided as quickly and accurately as possible whether the stimulus was a real English word or a nonword by pressing one of two buttons (*word* on the right, and *nonword* on the left) on a SuperLab response box. Between blocks, the participants worked on a filler task for approximately 5 min. The stimuli in both blocks were presented binaurally over Sony headphones. An iMac running SuperLab software (Cedrus Corporation, 2006)

controlled stimulus presentation and recorded RTs, which were measured from stimulus onset to buttonpress onset. If the maximum RT (5 s) was exceeded, the computer recorded an incorrect response and presented the next trial. The stimulus presentation within each block was random.

## Results

Following M<sup>c</sup>Lennan and Luce (2005), RTs less than 500 or greater than 2,500 ms were excluded (two RTs). Three of the participants were also excluded.<sup>2</sup> The overall accuracy to the experimental words in the target block was 96 %.

A Prime × Target repeated measures analysis of variance (ANOVA) was performed on the mean RTs to correct responses.<sup>3</sup> The main effect of prime was significant,  $F_1(2, 126) = 6.90, p = .001, MSE = 18,670.34, \eta_p^2 = .10$ ;  $F_2(2, 22) = 6.36, p = .007, MSE = 3,399.00, \eta_p^2 = .37$ . Because the focus was on evaluating priming and talker effects, the comparisons of primary interest were between the match and control conditions (magnitude of priming, or MOP) and between the match and mismatch conditions (magnitude of specificity, or MOS).

As the measure of MOP, we calculated match RTs minus control RTs; match RTs minus mismatch RTs served as the measure of MOS. There are other potential ways to calculate MOP, including (match+mismatch) / 2 – control, or mismatch – control. However, we chose to assess MOP on the basis of match minus control in order to be consistent with M<sup>c</sup>Lennan and Luce (2005) (as well as with other similar studies). Also, inspection of the means in Tables 1 and 2 reveals that such alternative calculations of MOP would have led to the same overall conclusions, albeit to somewhat weaker MOPs.

As is shown in Table 1, comparisons consisting of paired one-tailed  $t$  tests revealed significant MOP and MOS:

<sup>2</sup> One participant was removed because no filler task was given to this participant, due to experimenter error, and thus the target block began immediately following the prime block. A second participant was removed because of an overall mean RT to correct responses during the target block that was greater than two standard deviations above the grand mean. A third participant, who reported difficulty hearing the stimuli, was also removed.

<sup>3</sup> Two dummy variables representing the allocation of participants to experimental lists were included in the ANOVA solely to reduce the estimate of random variation (Pollatsek & Well, 1995); effects involving the dummy variables are not reported. Note that traditional item analyses were not appropriate for the present experiments. Raaijmakers (2003; see also Raaijmakers, Schrijnemakers, & Gremmen, 1999) has argued that conducting separate item analyses in designs that use counterbalanced lists is unfounded. Nevertheless, we will report item analyses, more because of convention than because of their appropriateness. The reader should bear in mind these caveats in interpreting the significance levels of all item tests reported for the present experiments.

**Table 1** Reaction times, standard errors (in parentheses), and magnitudes of specificity (MOS) and priming (MOP) for Experiment 1

Experiment 1	M	MM	C	MOS	MOP
English (with foreign- accented speech)	870 (13)	898 (13)	931 (16)	-28*	-61*

M, MM, and C refer to the match, mismatch, and control conditions, respectively. \*Significant effects in the MOS (match – mismatch) and MOP (match – control) planned comparisons

$t_1(68) = 3.08, p < .001, \text{Cohen's } d = 0.37$ ;  $t_2(11) = 3.01, p = .01, d = 0.99$ , and  $t_1(68) = 1.84, p = .035, d = 0.22$ ;  $t_2(11) = 1.34, p = .10, d = 0.40$ , respectively.<sup>4</sup> The difference between the mismatch and control conditions was also significant,  $t_1(68) = 1.80, p = .038, d = 0.22$ ;  $t_2(11) = 2.26, p = .022, d = 0.83$ .

## Discussion

The results of Experiment 1 are consistent with the time-course hypothesis. Recall that M<sup>c</sup>Lennan and Luce (2005) did not obtain talker effects in the same easy lexical decision task (Exp. 2A).

A combined ANOVA revealed that the Prime × Experiment (M<sup>c</sup>Lennan & Luce's, 2005, Exp. 2A, with native-accented speech, vs. the present Exp. 1, with foreign-accented speech) interaction was not significant,  $F < 1.0, MSE = 16,735.97, p = .658, \eta_p^2 = .003$ . Nevertheless, in addition to the statistically significant MOS effect found in the present experiment (-28 ms), but not in M<sup>c</sup>Lennan and Luce's Experiment 2A (-8 ms), an independent one-tailed  $t$  test revealed significantly longer RTs in the present experiment (900 ms) than in Experiment 2A of M<sup>c</sup>Lennan and Luce (773 ms),  $t(135) = 8.04, p < .01, d = 1.37$ , supporting the claim that foreign-accented speech slows processing, allowing specificity effects to emerge. However, an additional (two-tailed)  $t$  test revealed longer stimulus durations in the present experiment (579 ms) than in Experiment 2A of M<sup>c</sup>Lennan and Luce (373 ms),  $t(23) = 11.07, p < .01, d = 3.20$ .

Consequently, in order to investigate the relationship between foreign-accented speech and talker effects further, we conducted Experiment 2. The primary motivation for Experiment 2 was to provide a direct, within-study comparison of talker effects as a function of accent. Half of the

<sup>4</sup> One-tailed tests are reported for tests with an a priori prediction about the direction of an effect (e.g., faster RTs in the match than in the mismatch condition). Also, Cohen's  $d$  was calculated for within-participants data using an online effect size calculator created by the Cognitive Flexibility Laboratory (June 18, 2008; retrieved July 7, 2010, from [www.cognitiveflexibility.org/efficientsize/](http://www.cognitiveflexibility.org/efficientsize/)). The typical effect size interpretations for Cohen's  $d$  are 0.2 = small, 0.5 = medium, and 0.8 = large.

**Table 2** Reaction times, standard errors (in parentheses), and magnitudes of specificity (MOS) and priming (MOP) for [Experiment 2](#)

Experiment 2	M	MM	C	MOS	MOP
Spanish (with native- accented speech)	777 (21)	773 (24)	806 (28)	+4	-29*
Spanish (with foreign- accented speech)	927 (21)	984 (24)	1,031 (28)	-57*	-104*

M, MM, and C refer to the match, mismatch, and control conditions, respectively. \* Significant effects in the MOS (match – mismatch) and MOP (match – control) planned comparisons

participants heard words spoken by a native speaker, and half heard the same words spoken by a non-native speaker with a foreign accent. Furthermore, the durations of the native- and foreign-accented experimental words were equivalent, allowing us to rule out a duration-based explanation.

## Experiment 2: Spanish with native- and foreign-accented speech

### Method

**Participants** A group of 72 participants from the Universitat Jaume I (Spain) community were paid or received credit. The participants were right-handed native Spanish speakers with no reported history of speech or hearing disorders.

**Materials** All of the stimuli, shown in the [Appendix](#), were recorded in Spanish by one male and one female native American English speaker with a foreign accent, and by one male and one female native Spanish speaker with a native accent.<sup>5</sup>

The stimuli were recorded, filtered, and edited as in [Experiment 1](#). The mean word frequency for the experimental words was 981 per five million, according to LEXESP (Sebastián-Gallés, Martí, Carreiras, & Cuetos, 2000). The mean durations for the experimental words produced by the native (580 ms) and non-native (577 ms) speakers did not differ,  $t(46) < 1.0$ ,  $p = .857$ .

**Design** The design is identical to that of [Experiment 1](#), with the exception of adding the between-participants factor

<sup>5</sup> Both of the Spanish speakers and all of the Spanish participants were native speakers of Spanish (i.e., Spanish was their first and primary language). However, it is extremely difficult to find individuals in this area of Spain who are monolingual speakers. Consequently, many participants and both speakers also spoke Catalan, English, or both.

Accent (native or foreign). Half of the participants heard words and nonwords produced by the native Spanish speakers, and half heard the same stimuli produced by the native American English speakers in Spanish with a foreign accent.

**Procedure** The procedure was identical to that of [Experiment 1](#), except that the stimuli were presented over AKG-K55 headphones and the experiment was controlled by Inquisit 1.33 software on a Pentium PC, which recorded RTs.

### Results

No RTs were less than 500 or greater than 2,500 ms.<sup>6</sup> The overall accuracy to the experimental words in the target block was 91 %.

A Prime  $\times$  Target  $\times$  Accent mixed-factor ANOVA was performed on the mean RTs to correct responses. As expected, native-accented words were responded to more quickly (785 ms) than were foreign-accented words (981 ms),  $F_1(1, 60) = 48.64$ ,  $p < .001$ ,  $MSE = 85,023.36$ ,  $\eta_p^2 = .45$ ;  $F_2(1, 22) = 37.90$ ,  $p < .001$ ,  $MSE = 42,345.44$ ,  $\eta_p^2 = .63$ . Again, the MOP and MOS were of primary interest. The crucial difference between the present experiment and [Experiment 1](#) was our ability to directly evaluate talker effects in the native- and foreign-accented conditions.

The Prime  $\times$  Accent interaction was marginally significant,  $F_1(2, 120) = 2.60$ ,  $MSE = 21,995.14$ ,  $p = .079$ ,  $\eta_p^2 = .04$ ;  $F_2(2, 44) = 1.04$ ,  $MSE = 11,715.51$ ,  $p = .362$ ,  $\eta_p^2 = .05$ . Consequently, MOS and MOP analyses were performed separately for the native- and foreign-accented conditions, as shown in [Table 2](#).

In the native-accent condition, MOP was significant,  $t_1(35) = 1.85$ ,  $p = .04$ ,  $d = 0.31$ ;  $t_2(11) = 2.95$ ,  $p = .01$ ,  $d = 0.86$ , and MOS did not approach significance,  $t_1(35) < 1.0$ ,  $p = .38$ ,  $d = 0.05$ ;  $t_2(11) < 1.0$ ,  $p = .44$ ,  $d = 0.05$ . The difference between the mismatch and control conditions was also significant,  $t_1(35) = 2.17$ ,  $p = .019$ ,  $d = 0.37$ ;  $t_2(11) = 2.30$ ,  $p = .021$ ,  $d = 0.74$ .

In the foreign-accented condition, both MOP and MOS were significant,  $t_1(35) = 3.04$ ,  $p < .001$ ,  $d = 0.55$ ;  $t_2(11) = 2.22$ ,  $p = .02$ ,  $d = 0.64$ , and  $t_1(35) = 2.39$ ,  $p = .01$ ,  $d = 0.41$ ;  $t_2(11) = 1.00$ ,  $p = .17$ ,  $d = 0.25$ , respectively. The difference between the mismatch and control conditions was not

<sup>6</sup> Three participants in the foreign-accented condition had errors on both of the trials in the control condition for the male talker. As a result, for these three participants, there were no RTs to correct responses for that condition. Consequently, each of these three participants' means for their remaining five conditions was used to replace the missing value. Note that this affected less than 1% of the means from this experiment (i.e., three out of 432 condition means, or 72 participants  $\times$  6 condition means for each participant).



significant,  $t_1(35) = 1.17$ ,  $p = .126$ ,  $d = 0.20$ ;  $t_2(11) = 1.39$ ,  $p = .096$ ,  $d = 0.40$ .

A critical final comparison, consisting of an independent one-tailed  $t$  test, was performed in order to directly compare the MOS in the native- and foreign-accented conditions. These results provided further evidence that MOS was greater in the foreign-accented ( $-57$  ms) than in the native-accented ( $+4$  ms) condition,  $t_1(70) = 2.24$ ,  $p = .01$ ,  $d = 0.53$ ;  $t_2(22) < 1.0$ ,  $p = .21$ ,  $d = 0.34$ .

## Discussion

The results of [Experiment 2](#) are consistent with the time-course hypothesis. We are not arguing that talker effects would never be expected in native-accented speech; such evidence already exists (McLennan & Luce, 2005). Rather, our argument is that talker effects are more likely to occur when processing is relatively slow, and consequently, are more likely with foreign-accented speech.

Although both experiments involved foreign-accented speech, the following data suggest that listeners were indeed accessing the intended lexical items. First, accuracy in the lexical decision task was quite high (96 % and 91 % in [Experiments 1](#) and [2](#), respectively). Second, we collected additional data in order to address this issue directly. Ten new native speakers of American English at Cleveland State University were asked to identify each of the experimental words for the English stimuli (produced with a Spanish accent), and 10 new native speakers of Spanish at the Universitat Jaume I were asked to identify each of the experimental words for the Spanish stimuli (produced with an American English accent). The results for the English stimuli were as follows: The mean percentages correct for the stimuli produced by the male and the female talker were 98 % and 94 %, respectively. Furthermore, the mean percentage correct for the experimental words was 96 %. The results for the Spanish stimuli were as follows: The mean percentages correct for the stimuli produced by the male and the female talker were 95 % and 96 %, respectively. Furthermore, the mean percentage correct for the experimental words was 95 %. In short, for both the English and the Spanish stimuli, the foreign-accented words were intelligible across speakers and items. These data provide further evidence that the present results are not simply indicative of a decision under optimal conditions versus decision under uncertainty. Although many studies using degraded stimuli may result in relatively low accuracy indicative of some greater degree of uncertainty, accuracy in the present experiments was quite high, and RT was the main dependent variable.

We performed one final analysis, directly comparing the combined MOS from the two native-accented conditions (Exp. 2A of McLennan & Luce, 2005, and our Exp. 2) and

the two foreign-accented conditions ([Experiments 1](#) and [2](#)). The results of this one-tailed  $t$  test revealed significantly greater MOS in foreign- than in native-accented speech,  $t(207) = 2.05$ ,  $p = .02$ ,  $d = 0.28$ .

## General discussion

The present study demonstrated that talker effects are more likely in foreign-accented speech, consistent with the time-course hypothesis. The evidence is particularly strong, given that we not only found greater MOS in foreign-accented speech in our between-study comparison (Exp. 1), but also in our within-experiment comparison (Exp. 2). However, because Clarke and Garrett (2004) have shown that listeners adjust to foreign accents quickly when presented with longer utterances (complete sentences rather than isolated words), the present pattern of results may be restricted to isolated word recognition. That is, if listeners typically adjust to foreign accents quite rapidly, they may quickly revert to their default pattern of results, in which talker effects are less likely to affect their perception of spoken language. Nevertheless, the role that talker-specific representations play when listeners are presented with longer utterances of foreign-accented speech remains an empirical question that should be addressed in future studies.

The present study advances our understanding of the circumstances under which talker-specific details affect spoken word recognition (McLennan, 2006; 2007) by providing evidence of greater talker effects with foreign-accented speech. To our knowledge, this is the only published study examining the time course of talker effects when listeners' processing was relatively slow, without creating this slowing with lab manipulations or artificially degraded or disordered speech. The present results support the use of the same theoretical framework in accounting for talker effects in listeners' perceptions of clear speech, as well as naturally occurring degraded speech produced by dysarthric speakers and healthy speakers with a foreign accent. Furthermore, the present results provide important new information beyond the results with dysarthric speech.

Some researchers have discussed the role that attention may play in listeners' perception (e.g., Nygaard, 2003) and acquisition (e.g., Francis & Nusbaum, 2002) of abstract and more fine-grained acoustic-phonetic structure. Although we have interpreted our results in terms of the time-course hypothesis, both our results and the time-course hypothesis are compatible with an attention-based account. The degree of task difficulty may affect the way that listeners attend to the signal. When the task is easy, it may be sufficient for listeners to attend to only a few relevant phonemic distinctions in order to perform the task successfully. On the other hand, when the task is difficult, the listener may need to

devote more attentional resources to a finer level of phonetic detail, which in turn results in more robust talker effects.

Also, it may seem as though we are positing that talker-specific representations are qualitatively distinct from abstract representations, and that talker-specific representations do not play any role until later in processing. However, we are not arguing for either of these points. First, although our findings are consistent with qualitatively distinct representations, this is not necessarily the case. It is possible that abstract information and talker-specific details are part of a more distributed representation. Second, although our findings provide empirical evidence that abstract information and talker-specific details affect processing at different points in time, it is not necessarily the case that talker-specific details play no role until later. Rather than assuming that talker-specific representations (if qualitatively distinct) or talker-specific aspects of a distributed representation are not playing any role early (such that it takes longer for this information to become activated), it is possible that all sources of information play a role immediately, but some sources simply take longer for their effects to be detected. In this way, the time-course hypothesis does not necessarily posit that talker-specific information will play no early role, but rather that the effects of talker-specific information will always play a larger role later during processing.<sup>7</sup>

One final point merits discussion. Although the role that surface information, including talker-specific details, plays in the perception of spoken words remains an important issue, researchers have only examined one of the two directions of these effects. Researchers have manipulated surface information, most frequently the talker, and examined the effect that this manipulation has on listeners' ability to recognize the linguistic information (the spoken words). However, effects in the opposite direction have remained relatively unexplored: Researchers could manipulate the linguistic information (e.g., high vs. low frequency) and examine the effect that this manipulation has on listeners' subjective perception of the surface information (e.g., the strength of a non-native speaker's foreign accent). Shah and M<sup>c</sup>Lennan (2008) have begun to investigate the effect that ease of lexical access (primed vs. unprimed) has on listeners' accent ratings. Also, Nygaard and Queen (2008; see also Nygaard & Lunders, 2002) provided evidence that emotional tone of voice can affect listeners' processing of

the linguistic content of spoken words. Studies in which both directions in the relationship between linguistic and surface information are examined should lead to a more complete understanding of how listeners represent and process both types of information.

**Author Note** This research was supported by Research Grant No. R03 DC 007316 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health, and by the Ministry of Science and Innovation of Spain, Research Grant No. PSI2009-10067 (subprogram PSIC), and BRAINGLOT (Grant Consolider-Ingenio 2010). We thank Steven Beyer, Jaime Deyling, Alisa Maibauer, Teresa Markis, Jessica Newell, Timothy Patrick, Maura Krestar, and several undergraduate research assistants for their assistance conducting [Experiment 1](#).

## Appendix

**Table 3** Stimuli used in [Experiment 2](#)

Words	Nonwords
sueño	sueto
tiro	tizo
fuerza	fuerma
techo	techa
guerra	guerre
dedo	deda
sangre	sangri
grito	grimo
tiempo	tiempi
polvo	polvi
padre	padra
nieve	niele
jefe	jefi
hijo	hico
guardia	guardio
bosque	bosca

The last four rows are stimuli for the unrelated fillers trials used in the prime block only; the remaining rows are the experimental trials

## References

- Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics*, *61*, 206–219. doi:10.3758/BF03206883
- Brouwer, S., Mitterer, H., & Huettig, F. (2012). Speech reductions change the dynamics of competition during spoken word recognition. *Language and Cognitive Processes*, *27*, 539–571. doi:10.1080/01690965.2011.555268

<sup>7</sup> There is evidence that blocked and intermixed presentations of casually and carefully produced speech can lead to a different pattern of results (Brouwer, Mitterer, & Huettig, 2012). Consequently, future investigations should examine whether differences in the specificity effects for native- and foreign-accented speech are limited to the blocked design, or would extend to a mixed design as well.

- Burgund, E. D., & Marsolek, C. J. (1997). Letter-case-specific priming in the right cerebral hemisphere with a form-specific perceptual identification task. *Brain and Cognition*, *35*, 239–258.
- Cedrus Corporation. (2006). *SuperLab Pro (4.0)*. San Pedro, CA: Author. Retrieved from [www.superlab.com](http://www.superlab.com).
- Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 521–533. doi:10.1037/0278-7393.20.3.521
- Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, *116*, 3647–3658.
- Francis, A. L., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 349–366. doi:10.1037/0096-1523.28.2.349
- Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *22*, 1166–1183.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251–279. doi:10.1037/0033-295X.105.2.251
- Luce, P. A., & Lyons, E. A. (1998). Specificity of memory representations for spoken words. *Memory & Cognition*, *26*, 708–715. doi:10.3758/BF03211391
- Luce, P. A., & M'Lennan, C. T. (2005). Spoken word recognition: The challenge of variation. In D. B. Pisoni & R. E. Remez (Eds.), *Handbook of speech perception* (pp. 591–609). Malden: Blackwell.
- Luce, P. A., M'Lennan, C. T., & Charles-Luce, J. (2003). Abstractness and specificity in spoken word recognition: Indexical and allophonic variability in long-term repetition priming. In J. Bowers & C. Marsolek (Eds.), *Rethinking implicit memory* (pp. 197–214). Oxford: Oxford University Press.
- Mattys, S. L., & Liss, J. M. (2008). On building models of spoken word recognition: When there is as much to learn from natural “oddities” as artificial normality. *Perception & Psychophysics*, *70*, 1235–1242. doi:10.3758/PP.70.7.1235
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86. doi:10.1016/0010-0285(86)90015-0
- M'Lennan, C. T. (2006). The time course of variability effects in the perception of spoken language: Changes across the lifespan. *Language and Speech*, *49*, 113–125.
- M'Lennan, C. T. (2007). Challenges facing a complementary-systems approach to abstract and episodic speech perception. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 67–70). Saarbrücken: ICPS.
- M'Lennan, C. T., & Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 306–321. doi:10.1037/0278-7393.31.2.306
- M'Lennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 539–553. doi:10.1037/0278-7393.29.4.539
- M'Lennan, C. T., Luce, P. A., & Charles-Luce, J. (2005). Representation of lexical form: Evidence from studies of sublexical ambiguity. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 1308–1314. doi:10.1037/0096-1523.31.6.1308
- Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native- and foreign-accented speech. *Language and Speech*, *38*, 289–306.
- Nygaard, L. C. (2003). Perceptual stability and information variation: A commentary on Remez, Goldinger, Azuma, and Local. *Journal of Phonetics*, *31*, 345–349.
- Nygaard, L. C., & Lunders, E. R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory & Cognition*, *30*, 583–593. doi:10.3758/BF03194959
- Nygaard, L. C., & Queen, J. S. (2008). Communication emotion: Linking affective prosody and word meaning. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1017–1030.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113. doi:10.1016/0028-3932(71)90067-4
- Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 785–794. doi:10.1037/0278-7393.21.3.785
- Raaijmakers, J. G. W. (2003). A further look at the “language-as-fixed-effect fallacy”. *Canadian Journal of Experimental Psychology*, *57*, 141–151. doi:10.1037/h0087421
- Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How do deal with “the language-as-fixed-effect fallacy”: Common misconceptions and alternative solutions. *Journal of Memory and Language*, *41*, 416–426. doi:10.1006/jmla.1999.2650
- Sebastián-Gallés, N., Martí, M. A., Carreiras, M., & Cuetos, F. (2000). *LEXESP: Una base de datos informatizada del español*. Barcelona: Universitat de Barcelona.
- Shah, A. P., & M'Lennan, C. T. (2008). The role of foreign-accentedness in lexical processing. *Proceedings of Meetings of Acoustics*, *2*(060003), 1–13.
- Sommers, M. S., & Barcroft, J. (2006). Stimulus variability and the phonetic relevance hypothesis: Effects of variability in speaking style, fundamental frequency, and speaking rate on spoken word identification. *Journal of the Acoustical Society of America*, *119*, 2406–2416.
- Vitevitch, M. S., & Donoso, A. (2011). Processing of indexical information requires time: Evidence from change deafness. *Quarterly Journal of Experimental Psychology*, *64*, 1484–1493. doi:10.1080/17470218.2011.578749