Research Article

Falling on Sensitive Ears

Constraints on Bilingual Lexical Activation

Min Ju and Paul A. Luce

University at Buffalo, The State University of New York

ABSTRACT—*Spoken word recognition is characterized by multiple activation of sound patterns that are consistent with the acoustic-phonetic input. Recently, an extreme form of multiple activation was observed: Bilingual listeners activated words from both languages that were consistent with the input. We explored the degree to which bilingual multiple activation may be constrained by fine-grained acoustic-phonetic information. In a head-mounted eyetracking experiment, we presented Spanish-English bilinguals with spoken Spanish words having word-initial stop consonants with either English- or Spanish-appropriate voice onset times. Participants fixated interlingual distractors (nontarget pictures whose English names shared a phonological similarity with the Spanish targets) more frequently than control distractors when the target words contained English-appropriate voice onset times. These results demonstrate that fine-grained acoustic-phonetic information and a precise match between input and representation are critical for parallel activation of two languages.*

A critical question in research on bilingual word recognition concerns whether lexical representations in the two languages are processed independently or interactively. For example, Macnamara and Kushnir (1971) argued for independent processing, proposing a "language switch" mechanism that activates the relevant lexicon while switching off the irrelevant one. Several long-term priming experiments have supported the independence hypothesis by showing that phonetically remote translation equivalents (e.g., English *cat*/Spanish *gato*) fail to produce long-term priming effects (Gerard & Scarborough, 1989; Scarborough, Gerard, & Cortese, 1984; Watkins & Peynircioglu, 1983). In addition, functional magnetic imaging of Broca's area in bilinguals has shown that two separate nonoverlapping regions are activated during subvocal production of the two different languages (Kim, Relkin, Lee, & Hirsch, 1997). These studies suggest that at least some aspects of the lexical representations of the two languages are independent.

However, there is evidence demonstrating that lexical representations in bilinguals' two languages interact during recognition (for a review, see Spivey & Marian, 1999; Marian & Spivey, 1999, 2003). Spivey and Marian (1999) showed that even when Russian-English bilinguals are in a putatively monolingual mode, the two languages are activated simultaneously. In their experiment, highly proficient Russian-English bilinguals were asked to move objects according to instructions spoken in either language. The objects included a target (e.g., *marker*), an interlingual distractor whose name in the other language was phonologically similar to the target word (e.g., *marku*, "stamp"), and two control distractor objects whose names were not phonologically similar to the target word. Bilinguals fixated the interlingual distractors more often than the other distractors, suggesting that native- and second-language form-based representations were activated in parallel and competed for recognition. Spivey and Marian's research demonstrates an extreme form of multiple activation in spoken word recognition (e.g., Luce & Pisoni, 1998), such that all lexical representations—even those in different lexicons—may be activated by the spoken input with which they are consistent.

Cross-linguistically, sounds that are similar enough to be nominally categorized as belonging to the same phonetic category may exhibit fine-grained differences to which sophisticated listeners may be sensitive. For example, [t] in French is acoustically different from its English counterpart (e.g., Umeda & Coker, 1975), differing in place of articulation and voice onset time (VOT). VOT is a brief delay between the release burst and glottal pulsing and is a primary acoustic cue that distinguishes voiced sounds (e.g., *b*at) from voiceless sounds (e.g., *p*at). Spanish and English voiceless stop consonants show acoustic differences resulting from subtle differences in articulation. Specifically, Spanish voiceless stops have shorter VOTs than English voiceless stops (Abramson & Lisker, 1973; Flege & Eefting, 1986). Spanish and English word-initial voiceless stops also differ in *aspiration*, that is, in whether a puff of air is emitted during the interval between the offset of the release burst and the onset of the voicing (Ladefoged, 1993). In English, voiceless stops are aspirated in word-initial position (e.g., *p*it). In contrast, Spanish word-initial voiceless stops are unaspirated. Thus, Spanish word-initial voiceless stops differ from these in English in both VOT and aspiration.[1] Fine-grained phonetic variation between similar sounds belonging to the same sound category is referred to as allophonic variation. The variation focused on in this study is cross-linguistic allophonic variation.

Address correspondence to Min Ju, Language Perception Laboratory, 245 Park Hall, Department of Psychology, University at Buffalo, The State University of New York, Buffalo, NY 14260; e-mail: mju@buffalo.edu.

---

[1] Although aspiration may differentiate voiceless stops in the two languages, VOT is the primary cue distinguishing Spanish and English voiceless stops (e.g., Abramson & Lisker, 1973; Flege & Eefting, 1986).

Studies of cross-linguistic phone perception suggest that bilinguals are sensitive to subtle acoustic-phonetic differences (Flege, 1984, 1991; Flege & Hammond, 1982). Elman, Diehl, and Buchwald (1977) examined the perception of natural /ba/ and /pa/ syllables varying in VOT by Spanish monolinguals, English monolinguals, and proficient and less proficient English-Spanish bilinguals. Short-VOT tokens were identified as /pa/ by monolingual Spanish speakers, whereas the same tokens were identified as /ba/ by monolingual English speakers. Compared with their less proficient counterparts, highly proficient English-Spanish bilinguals identified significantly more of the syllables as /pa/. This research demonstrates that highly proficient second-language users are sensitive to fine-grained acoustic-phonetic differences among similar sounds in their first and second languages (see also Bohn & Flege, 1993; Hazan & Boulakia, 1993; Khattab, 2000). However, a crucial question remains: What is the role of acoustic-phonetic variation during on-line word recognition in bilingual listeners?

If bilinguals are sensitive to subtle acoustic-phonetic differences between their two languages, allophonic variation may constrain multiple cross-lexicon activation. That is, language-specific acoustic-phonetic cues may help bilingual listeners to selectively access form-based representations from the lexicon with which the acoustic-phonetic information is consistent, in keeping with at least a weak form of the independent account of bilingual word recognition. In contrast, if bilinguals code spoken input into phonemic representations that discard this acoustic-phonetic detail, cross-linguistic allophonic variation may fail to constrain access, so that it would not be language selective, a finding that would support the interactive account of bilingual word recognition. In short, are bilingual listeners sufficiently sensitive to the fine-grained allophonic differences across languages to selectively access words in a given language?

In order to test this hypothesis, we created two sets of stimuli having word-initial voiceless stops in spoken Spanish. One set consisted of unaltered versions of Spanish words articulated in a normal Spanish manner by a proficient Spanish-English bilingual. The other consisted of edited tokens that were identical to the unaltered versions in every way except in the initial segments (the interval from the onset of the initial release burst to the onset of glottal pulsing that contains information about VOT and aspiration). To create this altered set, we edited out the initial segments of the tokens in the first set and replaced them with the initial segments taken from a recording of the same target words articulated in an English-like manner by the same Spanish-English bilingual.

Using head-mounted eye tracking, we compared fixation probability for line drawings of the target (e.g., *playa*, "beach"), an interlingual distractor (an irrelevant English word that has a phonological similarity with the Spanish target word; e.g., *pliers*), and two control distractors (e.g., *ojo*, *mono*). We expected that if a bilingual's two languages are activated in parallel, the fixation probability would be greater for the interlingual distractor than for the two control distractors both when targets contained Spanish-appropriate VOTs and when they contained English-appropriate VOTs; this pattern of results would suggest that bilinguals are not sensitive to subtle allophonic variations during spoken word processing. However, we expected that if the two languages can be selectively accessed on the basis of cross-linguistic allophonic information, the fixation probability would be greater for the interlingual distractor than for the control distractors when the target contained English-appropriate VOTs, but not when the target contained Spanish-appropriate VOTs.

METHOD

Participants

Twenty-two students at the University at Buffalo were either paid $10 or given partial credit toward an introductory psychology course for their participation. All were Spanish-English bilinguals with normal or corrected-to-normal vision and reported no speech or hearing disorders. In order to assess the English proficiency of these participants, we administered a language-background questionnaire prior to the experiment. All participants were asked to rate their level of comprehension for spoken English on a 7-point scale (1 = *poor*; 7 = *very good*). They were also asked about their length of exposure to spoken English. Because cross-language VOT research suggests that only highly proficient bilinguals show evidence of establishing new phonemic category boundaries for nonnative sounds (Elman et al., 1977; Flege & Eefting, 1987), we chose only highly proficient bilinguals for this study. That is, only the bilinguals who gave their comprehension of spoken English a rating of 7 and had been exposed to spoken English at least 5 years were included in the analysis. Mean length of exposure to spoken English among these participants was 16 years. These bilinguals either immigrated to the United States at an early age (before the age of 10) or were born in the United States but spoke Spanish at home.

Materials

The target stimuli consisted of 12 Spanish words having word-initial voiceless stops[2] (see Table 1). In addition, there were 12 interlingual distractors. The interlingual distractors (e.g., *pliers*) were English words that were phonologically similar to the Spanish targets (e.g., *playa*, "beach"). On average, the targets and the interlingual distractors overlapped by 2.5 phonemes (average number of phonemes in Spanish words = 4.5; average number of phonemes in English words = 3.8). Twenty-four additional Spanish words were control distractors. The Spanish distractors (e.g., *ojo*, "eye") were not phonologically similar to the targets or the interlingual distractors.

The recorded stimuli were spoken by a native speaker of Spanish (from Chile) who is a fluent second-language speaker of English (with 30 years of exposure). Of the 12 critical trials (i.e., those bearing targets), half contained altered targets with English-appropriate VOT, and the other half contained unaltered targets with Spanish-appropriate VOT. In order to create altered targets, we made two sets of recordings: one with normal Spanish words and the other with the same Spanish words spoken in an English-like manner. The interval from the onset of the initial release burst to the onset of glottal pulsing was edited out of each intact normal Spanish token. The same interval was removed from the corresponding English-like token and then spliced onto the edited Spanish token. In this way, we created two sets of stimuli (altered vs. unaltered) that varied only in the initial segment (approximately 88 ms in duration). The root mean square (RMS) amplitude of the release burst was 56 dB for unaltered targets and 58 dB for altered targets. The RMS amplitude of aspiration was 55 dB for unaltered targets and 60 dB for altered targets. The average VOT

---

[2] Because of the extreme difficulty in finding words meeting all the constraints we placed on stimuli, we were left with a limited number of candidate words. In order to have a sufficient number of items for analysis, we decided to include an item whose interlingual distractor contained a word-medial stop (*cuna-raccoon*). A post hoc analysis revealed that the data pattern for this item was consistent with the pattern for the other words.

TABLE 1

Stimuli

Target		Interlingual distractor		Control distractor 1		Control distractor 2	
Spoken stimulus	Picture and English name	Picture and English name	Spanish name	Picture and English name	Spanish name	Picture and English name	Spanish name
cama	bed	comb	peine	tree	arbol	pencil	lapiz
clavo	nail	clover	trebol	lemon	limon	zebra	cebra
cola	tail	corn	maiz	fence	cerca	monkey	mono
cottora[a]	parrot	coat	abrigo	axe	hacha	nose	nariz
cuna	crib	raccoon	mapache	chair	silla	finger	dedo
pato	duck	pot	olla	skirt	falda	boot	bota
pie	foot	pea	arveja	envelope	sobre	frog	rana
pila	battery	pillow	almohada	eye	ojo	ruler	regla
pipa	pipe	people	gente	fly	mosca	lion	leon
pista	trail	pig	cerdo	cloud	nube	moon	luna
playa	beach	plier	alicate	arrow	flecha	mushroom	hongo
timbre	bell	timber	madera	table	mesa	watch	reloj

[a] Omitted from final analysis.

values were 31 ms for unaltered targets and 88 ms for altered targets. Thus, the amplitudes of release burst, aspiration, and VOT in the altered stimuli were close to typical English values (Stevens, 1998). Because altered targets had longer VOTs than unaltered targets, duration was longer for altered targets (mean duration = 543 ms for altered targets and 486 ms for unaltered targets).

The altered targets were screened for intelligibility by two native speakers of Spanish. On a scale from 1 (*not intelligible at all*) to 7 (*very intelligible*), all but one word (*playa*, "beach") received a rating of 7. This word was replaced with a new recording and passed rescreening with a rating of 7.

The picture stimuli consisted of 48 black-and-white line drawings. Thirty-one line drawings were selected from the Snodgrass and Vanderwart (1980) picture set, and the other 17 line drawings were hand-drawn. Participants were familiarized with the line drawings through a training session prior to the experiment.

Procedure

Prior to the experimental trials, participants were familiarized with and tested on the line drawings to ensure that they were identified as intended. For this session, participants were seated in front of a 17-in. flat-screen monitor connected to a Macintosh computer. In the familiarization phase, the pictures were presented on the monitor with their Spanish names, one at a time. After looking at a picture and its Spanish name, participants pressed the space bar to make the next picture appear. In the testing phase after familiarization, the pictures were again presented individually, this time without their Spanish names. Participants were required to type in the Spanish name of each picture. A beep sounded if the response was incorrect, and the same picture was shown again with its correct Spanish name before the next trial was presented. At the end of the session, participants were told how many errors they had made. If they made errors, they returned for another session, at which the pictures they named incorrectly were presented again, and they were asked to type the names of the pictures. This process continued until all of the pictures were identified correctly.

After the training session, participants were moved to a testing station equipped with an Eyelink head-mounted eyetracker and two computers. One of the computers controlled stimulus presentation, and the other collected eye movement data. A camera mounted on a lightweight headband worn by the participants monitored eye movements. The center of the pupil and first Purkinje corneal reflection were tracked using an infrared image, to determine the position of the eye relative to the head. Accuracy was better than one degree of arc; head and body movement were virtually unrestricted. Participants were seated about 57 cm from the computer. The experimenter monitored their eye movement throughout the experiment on a separate monitor.

The high-quality digital spoken stimuli were presented at a 44.1-kHz sampling rate at a comfortable listening level over Sony earphones. Ambient noise in the testing room was minimal.

After calibration, four practice trials were presented. Each trial proceeded as follows: A crosshair appeared in the center of the monitor, and participants were instructed to fixate on the crosshair. The experimenter then pressed the space bar to begin a trial. A beep was played to signal the beginning of a trial. A set of four pictures appeared in the corners of a 3 × 3 grid with the crosshair in the center. Each cell in the grid spanned about 5° of visual angle, which was within the resolution of the tracker (better than 1°). A schematic representation of the grid with four pictures is presented in Figure 1. Pictures appeared 500 ms after the beep. The spoken target word was then presented 500 ms after the onset of the pictures. Participants were instructed to remain fixated on the crosshair until the target word was presented. If they moved their eyes between the onset of the pictures and the onset of the target, the trial was aborted. Aborted trials were reshuffled into the remaining trials of the experiment. Locations of the pictures were randomized.

There were 36 trials in total: 12 critical trials (6 with altered targets and 6 with unaltered targets, so VOT was a within-participants factor) and 24 filler trials. Each critical trial included a picture corresponding to a target, a picture corresponding to an interlingual distractor (phonologically similar to the target), and pictures of two other distractors (that did not phonologically resemble the target or the interlingual distractor). The task was to use a mouse to click on the picture that corresponded to the target word as quickly as possible.[3] After the

---

[3] Because Spanish nouns are preceded by gender-marked articles, a carrier phrase (e.g., "click on the _____") could not be used.

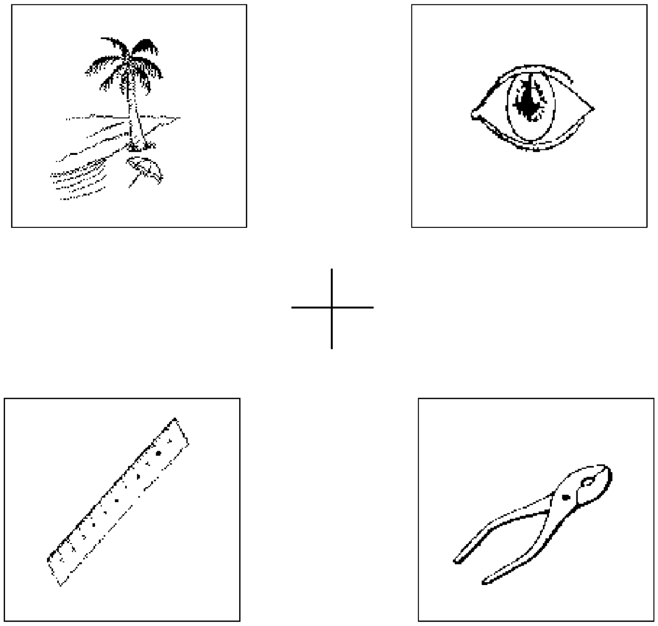

Fig. 1. Example of a stimulus display with *playa* (''beach'') as the target and *pliers* as the interlingual distractor.

participant responded, the pictures disappeared. Calibration was monitored by the experimenter and adjusted between trials when necessary.

Data Coding

The number of saccades made to each picture in each trial was recorded so we could compute fixation probability for the pictures corresponding to the target, the interlingual distractor, and the other two distractors. Accuracy for clicks on correct target pictures was 98%. Because mean latency of the main task was 1,429 ms, fixation data were analyzed for a 1,500-ms window beginning at the onset of the spoken stimuli.[4] Four participants were excluded because they performed many trials with peripheral vision, resulting in few eye movements. The remaining trials in which participants made no eye movements were treated as missing trials and constituted 6% of total trials. Some participants preferred to name the target picture for *cottora* ("parrot") as *loro* (another name for parrot). Given the potential confusion, we excluded this item in the final analysis.

RESULTS

Figure 2 presents mean fixation percentages (number of fixations to each picture as a percentage of the total number of trials on which it appeared) for the interlingual distractor and the two control distractors (averaged together) for the English-appropriate VOT (altered) and Spanish-appropriate VOT (unaltered) target stimuli.[5] Target fixation occurred on 96% of trials. Because our hypothesis concerned only relative fixation to interlingual and control distractors, target fixations were not included in the analysis.

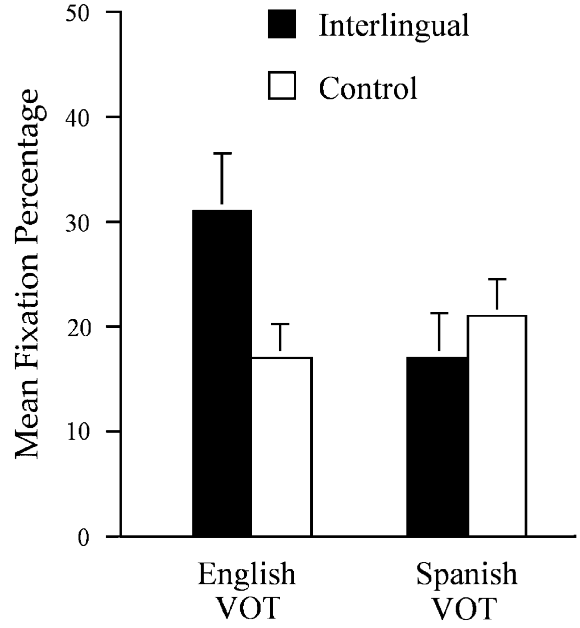

Fig. 2. Mean fixation percentage for interlingual and control distractors for target words with English-appropriate and Spanish-appropriate voice onset times (VOTs).

A 2 × 2 repeated measures analysis of variance was computed by participants (F_1) and items (F_2). VOT (English-appropriate vs. Spanish-appropriate) and distractor (interlingual distractor vs. two averaged control distractors) were within-participants factors. There was a significant main effect of VOT, $F_1(1, 17) = 5.61$, $p < .03$, and $F_2(1, 10) = 6.80$, $p < .03$, demonstrating that bilinguals fixated non-target objects more frequently when targets had English-appropriate VOTs than when they had Spanish-appropriate VOTs. There was no significant main effect of distractor ($F_1 = 1.54$ and $F_2 < 1$). However, there was a significant interaction of VOT and distractor, $F_1(1, 17) = 5.56$, $p < .04$, and $F_2(1, 10) = 5.32$, $p < .05$.

Planned comparisons revealed a significant difference between the interlingual distractor and the averaged control distractors for targets with English-appropriate VOTs, $F_1(1, 17) = 4.83$, $p < .05$, and $F_2(1, 10) = 5.72$, $p < .04$, indicating that bilinguals fixated the interlingual distractor more frequently than the control distractors only when the target contained English-appropriate VOT. For targets with Spanish-appropriate VOTs, there was no significant difference between the interlingual distractor and the averaged control distractors ($F_1 = 1.29$ and $F_2 < 1$), indicating that bilinguals fixated the interlingual distractor and the two averaged control distractors equally frequently when the target contained Spanish-appropriate VOT.

DISCUSSION

Our results demonstrate that fine-grained acoustic-phonetic information has demonstrable effects on activation of language-appropriate representations during bilingual word recognition. Specifically, when listening to Spanish words with English-appropriate VOTs, Spanish-English bilinguals fixated the interlingual distractors (pictures whose English names were phonologically similar to Spanish target words) more frequently than the control distractors. In contrast, when listening to target words with Spanish-appropriate VOTs, participants fixated the

---

[4] Because it takes approximately 200 ms on average to launch a saccade (Matin, Shao, & Boff, 1993), we conducted another analysis on fixations excluding the initial and final 200 ms. We obtained similar results.

[5] Because 6% of the trials contained no eye movements, all data were normalized by total number of trials completed. Percentages do not sum to 100% because participants looked at more than one picture on each trial.

two types of distractors equally frequently. The finding that the interlingual distractors were activated more frequently than the control distractors when targets had English-appropriate VOTs suggests that cross-lexicon activation depends on bottom-up acoustic-phonetic input and that differences in VOT between Spanish and English may have caused the reduced activation of the interlingual distractors when targets contained Spanish-appropriate VOTs. Thus, although our findings are not entirely consistent with Macnamara and Kushnir's (1971) notion of language switch, they support a model in which lexical activation is primarily driven by the bottom-up acoustic-phonetic input. Our results also suggest that activation of lexical representations by highly proficient bilinguals is exquisitely sensitive to subtle differences in the acoustic-phonetic structure of the input (see also McMurray, Tanenhaus, Aslin, & Spivey, 2003) and that parallel activation of two languages is contingent upon a precise match of fine-grained acoustic-phonetic information (see also Bürki-Cohen, Grosjean, & Miller, 1989).

Further, our finding that Spanish-English bilinguals use fine-grained allophonic variation to activate language-specific lexical representations is consistent with previous research suggesting that highly proficient bilinguals have language-specific phonetic category boundaries (e.g., Elman et al., 1977; Flege & Eefting, 1987). Moreover, our results indicate that cross-linguistic allophonic variation may play a role in bilingual word recognition even when listeners are in a putatively monolingual mode.

Finally, although our results for the altered stimuli are consistent with Spivey and Marian's (1999) findings for Russian-English bilinguals, we failed to replicate their findings using unaltered Spanish targets (see also Weber & Cutler, 2004). If overall phonological similarity drives parallel activation of the two languages, we should have observed greater activation of the interlingual distractors than of the control distractors in both altered- and unaltered-target conditions, given that the two sets of stimuli were identical except for the initial segment. One reason for this discrepancy between our results and those of Spivey and Marian may be that their stimuli contained words starting with a variety of sounds, including nasals (e.g., *marker*) and fricatives (e.g., *fish*), whereas all of our stimuli began with voiceless stops. Lacking a strong acoustic cue (e.g., voicing) that might inhibit cross-lexicon activation, Spivey and Marian's stimuli may have encouraged a greater degree of cross-lexicon activation.

Acknowledgments—We thank James Sawusch and William Schmidt for their help throughout the experiment and Alba Arreys, Mauricio Cavallo, Fanny Miranda, and Nick Schlienz for their assistance in developing the stimuli and collecting data. This research was supported by Research Grant No. R01 DC 0265801 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

REFERENCES

Abramson, A.S., & Lisker, L. (1973). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics*, *1*, 1–8.

Bohn, O.S., & Flege, J.E. (1993). Perceptual switching in Spanish/English bilinguals: Evidence for universal factors in stop voicing judgments. *Journal of Phonetics*, *21*, 267–290.

Bürki-Cohen, J., Grosjean, F., & Miller, J.L. (1989). Base-language effects on word identification in bilingual speech: Evidence from categorical perception experiments. *Language and Speech*, *32*, 355–371.

Elman, J., Diehl, R., & Buchwald, S. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America*, *62*, 971–974.

Flege, J. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, *76*, 692–707.

Flege, J. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, *89*, 395–411.

Flege, J., & Eefting, W. (1986). Linguistic and developmental effects on the production and perception of stop consonants. *Phonetica*, *43*, 155–171.

Flege, J., & Eefting, W. (1987). Cross-language switching in stop consonant production and perception by Dutch speakers of English. *Speech Communications*, *6*, 185–202.

Flege, J., & Hammond, R. (1982). Mimicry of non-distinctive phonetic differences between language varieties. *Studies in Second Language Acquisition*, *5*, 1–18.

Gerard, C., & Scarborough, D. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 305–315.

Hazan, V., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech*, *36*, 17–38.

Khattab, G. (2000). VOT production in English and Arabic bilingual and monolingual children. *Leeds Working Papers in Linguistics*, *8*, 95–122.

Kim, K.H.S., Relkin, N., Lee, K., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, *388*, 171–174.

Ladefoged, P. (1993). *A course in phonetics*. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers.

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

Macnamara, J., & Kushnir, S. (1971). Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior*, *10*, 480–487.

Marian, V., & Spivey, M. (1999). Activation of Russian and English cohorts during bilingual spoken word recognition. In *Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp. 349–354). Mahwah, NJ: Erlbaum.

Marian, V., & Spivey, M. (2003). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, *24*, 173–193.

Matin, E., Shao, K., & Boff, K. (1993). Saccadic overhead: Information processing time with and without saccades. *Perception & Psychophysics*, *53*, 372–380.

McMurray, B., Tanenhaus, M., Aslin, R., & Spivey, M. (2003). Probabilistic constraint satisfaction at the lexical/phonetic interface: Evidence for gradient effects of within-category VOT on lexical access. *Journal of Psycholinguistic Research*, *32*, 77–97.

Scarborough, D., Gerard, L., & Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behavior*, *23*, 84–99.

Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215.

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

Stevens, K.N. (1998). *Acoustic phonetics*. Cambridge, MA: MIT Press.

Umeda, N., & Coker, C. (1975). Subphonemic details in American English. In G. Fant & M. Tatham (Eds.), *Auditory analysis and perception of speech* (pp. 539–564). New York: Academic Press.

Watkins, M., & Peynircioglu, Z. (1983). On the nature of word recall: Evidence for linguistic specificity. *Journal of Verbal Learning and Verbal Behavior*, *22*, 385–394.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*, 1–25.

(RECEIVED 3/3/03; REVISION ACCEPTED 5/7/03)