J Psycholinguist Res (2008) 37:309–329
DOI 10.1007/s10936-008-9069-z

ORIGINAL ARTICLE

What Do Second Language Listeners Know About Spoken Words? Effects of Experience and Attention in Spoken Word Processing

Pavel Trofimovich

Published online: 11 March 2008

Abstract With a goal of investigating psycholinguistic bases of spoken word processing in a second language (L2), this study examined L2 learners' sensitivity to phonological information in spoken L2 words as a function of their L2 experience and attentional demands of a learning task. Fifty-two Chinese learners of English who differed in amount of L2 experience (longer vs. shorter residence in L2 environment) were tested in an auditory word priming experiment on well-known L2 words under two processing orientation conditions (semantic, control). Results revealed that, with more L2 experience, learners become more sensitive to phonological detail in spoken L2 words but that attention to word meaning might eliminate this sensitivity, even for learners with more L2 experience.

Keywords Spoken word processing · Auditory word priming · Second language · Attention · Language experience

Introduction

Whereas research in second language (L2) lexicon can no longer be characterized as "a neglected area of language learning" (Meara 1980), some of its aspects have been investigated less than others (see Bogaards and Laufer 2004; Kroll and Sunderman 2003, for reviews). The role of phonology in L2 lexical processing and learning represents one such aspect. Until recently, the focus of research in L2 lexical learning, at least from a psycholinguistic (processing) perspective, has predominantly been semantic in nature, exemplified by the study of word learning as the process of integrating words into lexical and conceptual networks (e.g., Kirsner et al. 1980; Van Hell and De Groot 1998). This interest in the learning of L2 words as semantic entities is perhaps not surprising, given that meaningful, systematic

P. Trofimovich (✉)
Department of Education/TESL Centre, Centre for the Study of Learning and Performance, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, QC, Canada H3G 1M8
e-mail: pavel.trofimovich@concordia.ca

relationships between prior and new experiences underlie any form of human learning (see, e.g., Bartlett 1932), including the learning of words.

There is a growing body of evidence, however, suggesting that how well L2 words are learned is determined, at least in part, by L2 learners' ability to notice, encode, store, and subsequently retrieve phonological (auditory, acoustic) information about spoken words (Barcroft and Sommers 2005; Speciale et al. 2004; Trofimovich 2005; Trofimovich and Gatbonton 2006). The present study adds to this body of research by investigating whether L2 learners, those with more and less experience with an L2, are sensitive to phonological information in spoken L2 words under two different circumstances: when their attention is directed to the meaning of spoken words as opposed to when it is not. The overall goal of the present study was, therefore, to determine how L2 experience and attention to input interact to result in processing benefits for spoken L2 words.

Semantic Approaches to L2 Word Learning

Craik and Lockhart's (1972) levels-of-processing (LOP) framework is perhaps the best known psycholinguistic approach to the study of word learning from a semantic (meaning-based) perspective (see also Craik and Tulving 1975). The framework is based on the assumption that incoming verbal information can be perceived, encoded, stored, and subsequently used at different levels of processing (e.g., phonemic, graphemic, semantic, imagery, or associative) and that some levels are necessarily "deeper" and more elaborative than others. According to the LOP framework, deeper and more elaborative processing levels (e.g., semantic, associative), as opposed to "shallower" ones (e.g., phonological, graphemic), should lead to long-lasting memory representations for words because a deep, semantic analysis of incoming information ensures its successful integration with existing knowledge structures. In support of this prediction, a number of studies have shown that word recall (a measure of memory for words) in a native language (L1) is best when participants are either explicitly instructed to learn words or implicitly oriented to their semantic properties, for example, by rating word pleasantness (e.g., Bower and Winzenz 1970; Walsh and Jenkins 1973; but see Morris et al. 1977, for criticisms).

Applied to L2 learning, the LOP framework (directly or indirectly) motivated investigations into the applicability and efficiency of deep, elaborative processing in L2 word learning. This research has shown that such semantic elaboration techniques as using imagery to teach words (Kasper 1983; Pressley et al. 1981, 1982), embedding words into meaningful or associative contexts (Beck et al. 1987; Brown and Perry 1991; Vochatzer and Blick 1989), or manipulating words and their synonyms in sentences (Crow and Quigley 1985) all lead to superior memory for L2 words, especially when these techniques are contrasted with rote repetition or using context or pictures alone to introduce words. While informative of the (largely positive) effects of semantic processing on L2 word learning (but see Barcroft 2002, Barcroft 2003; Sagarra and Alba 2006), these studies have little to say, however, about the role of auditory (or perceptual) processing in L2 word learning or about possible mechanisms underlying spoken word learning. Given that a spoken word, in addition to carrying semantic information, is also encoded in a particular phonological "shape" (determined by the sound structure of a language), a comprehensive account of L2 word learning must include spoken word processing (see Beckman and Edwards 2000, for a similar claim in L1 development).

A Case for a Study of Spoken Word Processing

A first line of evidence for the important role of spoken word processing comes from studies demonstrating beneficial effects of auditory word presentation on word recall (Engle et al. 1980; Engle and Roberts 1982; Gathercole and Conway 1988; Rae 1979). These studies, which compared L1 word recall in the visual and auditory learning conditions, reveal that participants typically recall more words after they have been encountered in the auditory than in the visual modality. Brand and Jolles (1985), who studied the learning of Dutch words by native English speakers, provided similar evidence for the superiority of auditory word learning in the L2. In particular, when tested in a delayed recall task, the learners in their study showed faster response times and made fewer errors after learning spoken than written L2 words (see also Channell 1988). Taken together, these findings suggest that word retrieval from memory may be less effortful for words learned in the auditory than in the visual modality.

A powerful indication of learners' tendency to process L2 words, at least in the initial stages of learning, on the basis of their spoken form comes from studies of word association (Meara 1982, 1984; see also Van Hell and De Groot 1998). Using word association tests, Meara found that L2 learners of French produced word associations that were largely based on phonological rather than semantic similarity (e.g., French *important* [important] produced in response to French *cruche* [pitcher, jug], most likely as a result of confusion with phonologically similar English *crucial*). This pattern of word associations was due to the learners' difficulty in correctly perceiving spoken L2 words, as misperceptions accounted for a majority of inexplicable word associations. These findings suggest that the representation and use of L2 words may, at least in part, be determined by their spoken form.

Additional evidence for the important role of spoken word processing comes from studies citing effects of overt articulation (the production aspect of spoken word processing) on word learning. These studies have revealed that learners typically recall more L2 words after saying them aloud than after reading them silently (Kelly 1992; Rodgers 1969; Seibert 1927), even when the learning task emphasizes meaningful (semantic) properties of words (Pressley et al. 1980). Learners also appear to use overt articulation spontaneously as a word learning strategy in a variety of learning tasks (Hall et al. 1981; Henning 1974) and to prefer overt articulation over such strategies as semantic elaboration or analysis of morphological and orthographic features of words (Lawson and Hogben 1996). As a learning strategy, overt articulation may enhance word learning because it engages auditory rehearsal processes in short-term memory (Baddeley et al. 1975) and provides opportunities for auditory self-perception (Baker and Trofimovich 2006; Cooper et al. 1975), thus ensuring that phonological representations of words are maintained in short-term memory and are later committed to long-term storage (see Baddeley et al. 1998, for review).

Auditory Word Priming

If L2 learners' experience with spoken words plays a significant role in word learning, then it is important, first, to measure the processing benefits of such experience and, second, to describe contexts in which these processing benefits are most likely to arise. The present study used auditory word priming as a methodology to accomplish these goals. In a typical auditory word priming experiment, listeners are exposed to a set of spoken words in a first task and are tested in a second task on another set of words containing both those words that were previously heard and words that are new to this task. In this second task, listeners often demonstrate auditory word priming, a phenomenon of processing facilitation, whereby they benefit from repeated (previously heard) spoken words. For example, in the second task, listeners are more likely to identify (e.g., Jackson and Morton 1984), repeat (e.g., Onishi et al. 2002), and recall from memory (e.g., Pilotti et al. 2000) the words they have heard in recent experience than the words they have not.

This processing advantage has two characteristics relevant to the present study. First, this processing advantage originates at the level of auditory (as opposed to semantic) processing, that is, processing requiring a language user to rely on form-related, phonological characteristics of spoken words (Schacter and Church 1992; Church and Schacter 1994). In other words, the fact that listeners demonstrate auditory word priming is indicative of their sensitivity to phonological information in spoken words. Second, this processing advantage is sensitive to the degree of perceptual overlap between spoken words, such that the amount of processing advantage is largest when two renditions of a spoken word also share one (or more) context-specific (indexical) details: a speaker's voice (Schacter and Church 1992), the intonation of an utterance (e.g., happy vs. angry), its fundamental frequency or pitch (Church and Schacter 1994), or phonetic context (Fisher et al. 2001). For example, repeated words spoken in a familiar voice are recognized more accurately than the same words spoken in an unfamiliar voice (Goldinger 1996; Sheffert 1998). This means that listeners not only demonstrate a processing benefit from "abstract" word-identity matches (repeated words) but also gain an additional processing benefit when "specific" details of spoken words (e.g., a speaker's voice) match across episodes of experience (repeated words and voices). Put differently, auditory word priming is also indicative of listeners' sensitivity to different types of information (abstract and specific) in spoken words.

In the context of the present study, abstract information available in spoken words might be conceptualized as listeners' knowledge, for example, of syllable structure typical of a language, phonotactic regularities permissible in it, or its phonetic features, phonemes, or syllables. Such information might be accessed and used at a relatively abstract phonological level of processing (Pallier et al. 2001). In turn, specific information refers to acoustic–phonetic details encoded in spoken words, those that are particular to an individual speaker or a given utterance (Goldinger 1996).

Do L2 learners benefit from their experience with spoken words and, if so, what information are they sensitive to? Two recent investigations of auditory word priming have yielded preliminary answers to these questions. In one study, Trofimovich (2005) reported that low-intermediate L2 learners of Spanish overall appeared to benefit from their experience with spoken words, demonstrating an auditory word priming effect. This finding suggested that the learners were able to encode the phonological information in spoken words and to bring it to aid in subsequent word processing. However, the learners demonstrated processing benefits only for words spoken in the same voice (i.e., the same word uttered by two different people evinced no benefits). This finding, in turn, suggested that the learners were not sensitive to abstract information in spoken words and were thus processing them in a speaker-specific manner. In the other study, Trofimovich and Gatbonton (2006) showed that, at least for some low-intermediate L2 learners of Spanish, the processing benefits appeared to be minimal when the learners' attention was drawn, at the time of first exposure to words, to their meaning (cf. Barcroft 2003). This finding suggested that meaning-based processing may divert learners' attention from spoken word form, preventing words from being encoded fully and eliminating processing benefits (Masson and MacLeod 2000).

The Current Study

Taken together, these findings reveal possible constraints on L2 learners' ability to benefit from their experience with spoken words. What is not clear, however, is whether (and how) learners eventually overcome these constraints, for example, as they gain more L2 experience. That is, do learners with more L2 experience come to benefit from repeated exposure to words spoken by different people, thus generalizing across non-identical instances of spoken words and processing them in a speaker-independent manner? Similarly, do learners with more L2 experience learn to benefit from repeated exposure to spoken words despite their attention being drawn to the semantic properties of words? It is only reasonable to suppose that learners with more L2 experience should be able to do so. With increasing amounts of L2 experience, learners would have multiple opportunities to both perceive and produce their L2. A possible outcome of such experience is likely twofold: certain abstract knowledge of the sound and syllable structure of the L2 (Pallier et al. 2001) coupled with the ability to disregard variability in those acoustic–phonetic details of speech (e.g., speaker's voice, speech rate) that are specific to an individual speaker or a particular situation (Goldinger 1996).

At least some learners eventually do learn to perceive and produce L2 speech accurately. First and foremost, these learners learn to attend to phonologically relevant details in L2 speech, for example, those that signal the phonetic distinction between *right* and *light* (Iverson et al. 2003). Likewise, they learn to disregard variability in context-specific details of L2 speech, for example, those differentiating a male and a female speaker saying *right* or *light* (Bradlow et al. 1997). Last but not least, they also learn to allocate attention efficiently among various aspects of the L2, for example, between *right* signaling location (as in *to the right of*) and *light* signaling color hue (Segalowitz and Frenkiel-Fishman 2005).

The present study attempted to answer both questions posed above by testing native Chinese learners of English with relatively more and relatively less L2 experience, defined here as the learners' length of residence (LOR) in Canada, in tests of auditory word priming. These tests included spoken L2 words that either shared or did not share the specific details of a speaker's voice. If learners are sensitive to both abstract and specific information in spoken words, then they should demonstrate processing benefits for repeated words that are spoken in the same and in different voices. The prediction here was that the extent of processing benefits obtained should depend on LOR, such that these benefits should be larger for learners with a longer LOR than for learners with a shorter LOR. These tests were also conducted under two conditions: when learners' attention was directed to the meaning of spoken words (semantic condition) as opposed to when it was not (control condition). If a semantic processing of spoken words diminishes the amount of processing benefits, then learners should demonstrate smaller benefits when they are instructed to attend to the meaning of words than when they are not. The prediction here was that learners with a longer LOR will be less susceptible to this effect of semantic processing than learners with a shorter LOR. Overall, these tests sought to clarify processing bases of learners' experience with spoken L2 words.

Method

Participants

The participants were 65 native Chinese speakers (23 female, 42 male). A total of 52 speakers (17 female, 35 male) were included in the final data analyses. The remainder were excluded because of an early (childhood) exposure to English in Canada (6) and large portions of data lost due to equipment failure (7). The participants, who ranged in age between 19 and 42 years (mean age = 29.8), were born in China. All but two cited Mandarin as their mother tongue (one reported Cantonese, one identified both Cantonese and Mandarin). The participants had arrived in Canada at an average age of 27.4 (17.6–37.0) and had resided there for a mean of 2.4 years (0.4–6.4 years). With no exception, all participants had received primary and secondary education in China. All had started learning English in China at a mean age of 11.3 (5–16 years) as part of formal classroom instruction. Prior to arriving in Canada, the participants had completed on average 3 semesters of English instruction in middle school (0–3), 3 semesters in high school (2–3), and 3 semesters in university (1–10), for an average of 5.7 semesters all together. Upon their arrival in Montreal and prior to taking part in this study, the participants had completed on average two additional English composition courses at Concordia University (1–4). Using a 10-point scale (1 = *extremely poor*, 10 = *extremely fluent*), the participants estimated their overall proficiency in Chinese and English at a mean of 8.9 (7–9) and 5.4 (1–8), respectively.

At the time of testing, all participants were students at Concordia University in Montreal (an English-language-medium university), which suggests that they used English often; all had been exposed only to the Canadian variety of English. Because length of residence in Canada (LOR) is an adequate measure of L2 experience only if learners speak the L2 often (Flege and Liu 2001), the participants were asked to estimate their daily use of English and Chinese in writing, reading, listening, and speaking on a 0–100% scale. The analysis of these self-ratings indicated that the participants, on average, used English relatively often on a daily basis ($M = 59\%$, $SD = 18.3$) and that their daily use of English was more frequent than their daily use of Chinese ($M = 35\%$, $SD = 16.0$), $t(51) = 5.55$, $p < .0001$, $r = .61$.

The participants were assigned to two equal groups ($n = 26$), with non-overlapping distributions, based on their LOR in Canada: one with a mean LOR of 1.3 years (0.4–2.2 years), the other with a mean LOR of 3.5 years (2.3–6.4 years). Two tests were then administered to ensure that the two groups differed only in terms of their LOR, a variable of primary interest here. The first test was a reading task in which the participants read a simple 440-word story in English and were recorded directly onto a computer using a Plantronics (DSP-300) stereo microphone. Two measures were derived from these recorded and subsequently transcribed readings. The first measure was speech rate, a fluency measure, calculated by dividing the number of uttered words by the total duration of the sample (which included pauses and hesitations), yielding a speech-rate ratio (words per second). Durations were measured to the nearest millisecond from the display of digital speech-analysis software (*CoolEdit 2000*). The obtained speech-rate ratios ranged between 57.9 and 141.2, with a mean of 111.5. The second measure was error rate, an accuracy measure, computed by dividing the number of words containing an error by the total number of words in the sample. Errors were defined as false starts, omitted words or phonemes, and any errors in sentence structure, morphology, or syntax. Common pronunciation mistakes, particularly those typical of Chinese learners of English (e.g., /r/-/l/ substitutions, mispronounced word-final stops), were not considered errors; these mistakes were equally common to both LOR groups. The obtained error rates ranged between .03 and .30, with a mean of .07.

The second test was a speaking task in which the participants spoke extemporaneously for about two minutes in English in response to a simple prompt (*talk about your experiences settling down in Canada*). The recorded speaking samples were subsequently presented to a panel of 10 judges (9 females, 1 male) to assess the degree of accentedness, comprehensibility, and fluency in the participants' speech. The judges, native speakers of English (mean age = 41.1 years; range = 22–58), listened to a short excerpt from each participant's recording (mean duration = 19.9 s, range = 15.5–24.7) and rated each sample using a 9-point Likert scale for degree of foreign accent (1 = *heavily accented*, 9 = *not accented at all*), comprehensibility (1 = *hard to understand*, 9 = *easy to understand*), and fluency (1 = *not fluent at all*, 9 = *very fluent*). The judges listened to speech samples individually and circled the appropriate rating, upon hearing each sample, on a response sheet before them. The obtained scores, computed for each participant by averaging the ratings given by the 10 judges (interrater reliability $\alpha = .92 - .94$), ranged between 2.4 and 7.6, with a mean accenteness rating of 4.4, a mean comprehensibility rating of 5.6, and a mean fluency rating of 5.5.

The 26 participants in each LOR group were then randomly assigned to two experimental conditions—control and semantic, with 13 participants in each. A series of 2 × 2 (LOR × condition) analyses of variance (ANOVAs) were conducted to determine if the participants assigned to each of the four groups differed on a number of demographic and proficiency variables. These analyses revealed no statistically significant main effects or interactions for six of the eight variables examined: (a) speech-rate ratio, (b) error rate, (c) degree of foreign accent, (d) comprehensibility, (e) fluency, and (f) semesters of English instruction completed, Fs(1, 48) < 2.70, ps > .11. However, similar two-way ANOVAs revealed significant main effects of LOR (with no significant effects of condition or significant LOR × condition interactions) for two other variables examined: (a) age at the time of arrival (AOA) in Canada and (b) chronological age, Fs(1, 48) > 16.15, ps < .0001. As reported earlier, the participant differed in their LOR in Canada, such that both groups with a longer LOR (assigned to the control and semantic conditions) had resided in Canada significantly longer than both groups with a shorter LOR. Although all participants were adults, the participants with a longer LOR were also slightly older than the participants with a shorter LOR at the time of their arrival in Canada, ts(24) > 2.78, ps < .01, rs > .49. As a consequence, the participants with a longer LOR were slightly older than the participants with a shorter LOR at the time of testing, ts(24) > 3.91, ps < .001, rs > .62. These analyses thus confirmed that learners' LOR and their AOA were confounded in this study (cf. Hakuta et al. 2003; Stevens 2004; Trofimovich and Baker 2006). The implications of this are discussed for each analysis conducted. The characteristics of the four participant groups appear in Table 1.

Table 1 Background and language proficiency characteristics of participants

Measure	Condition							
	Control				Semantic			
	Short LOR		Long LOR		Short LOR		Long LOR	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Chronological age[a]	26.6	1.3	33.8	1.3	25.8	1.5	33.2	1.2
Age of arrival[a] (AOA)	25.2	1.2	30.3	1.2	24.6	1.5	29.8	1.1
Length of residence[a] (LOR)	1.4	0.2	3.6	0.4	1.2	0.1	3.5	0.3
English self-rating[b]	5.8	1.1	5.7	1.2	5.2	1.3	5.9	1.0
Overall percent English use[c]	48.5	22.0	63.4	15.8	60.5	15.5	63.7	17.1
Speech rate (words/minute)	116.9	2.9	108.9	5.0	109.5	5.9	110.5	4.5
Reading errors (error rate)	5.2	0.8	8.1	2.0	8.0	1.3	6.3	1.1
Degree of foreign accent[d]	4.6	0.3	3.8	0.2	4.4	0.2	4.6	0.3
Comprehensibility[e]	5.9	0.3	5.0	0.3	5.5	0.2	6.0	0.3
Fluency[f]	5.6	0.3	5.3	0.3	5.3	0.3	5.9	0.3
Semesters of English taken	10.7	2.0	10.5	4.2	9.8	1.8	10.5	2.3

Notes: [a] In years. [b] Measured on a 10-point scale (1 = *extremely poor*, 10 = *extremely fluent*). [c] Measured on a 0–100% scale. [d] Measured on a 9-point scale (1 = *heavily accented*, 9 = *not accented at all*). [e] Measured on a 9-point scale (1 = *hard to understand*, 9 = *easy to understand*). [f] Measured on a 9-point scale (1 = *not fluent at all*, 9 = *very fluent*)

Table 2 Characteristics of English word sets

Word characteristics	English word sets	
	Set 1	Set 2
Mean number of syllables	2.3	2.2
Mean word frequency	180	159
Mean word duration	501	489

Notes: Word frequency estimates are based on Kucera and Francis (1967); word duration (in milliseconds) was measured from digitized word recordings

Materials

The materials consisted of 72 common English words recorded by six adult native speakers of American English (3 males, 3 females) from Illinois (see Appendix A). (For detailed information on stimulus preparation, see Trofimovich 2005.) The recorded words spanned a range of written frequencies (21–1,171), with a mean of 169.6 (Kucera and Francis 1967), and were on average 2.3 syllables long, with a mean duration of 495 ms. The words were not edited in any other way and were thus representative of the acoustic–phonetic variability found in speech. The words were first divided into two sets of 36 words for which six words were randomly chosen from each speaker's productions, with each speaker therefore contributing six words to each set. Both sets were matched for mean number of syllables, mean word frequency, and mean word duration (see Table 2 for set characteristics). Then, the two sets of 36 words were used to construct four study-test list pairs. Each pair contained a 36-word study list and a 72-word test list. The test list included the entire list of study words (36 repeated words) and a list of test words (36 unrepeated words). Of the 36 study words included in a test list, half were spoken by the same speaker (i.e., in the same voice) and half were spoken by a different speaker (i.e., in a different voice), where each word that was spoken by a male speaker in a study list was spoken by a female speaker in a test list, and vice versa. Across the 4 study-test list pairs, each word was spoken equally often in a male and a female voice. Each word appeared equally often in study and test lists.

Procedure

The testing, which lasted approximately 45 min per participant, was conducted individually in a quiet location. The experimental procedure was controlled by a Macintosh G4 iMac computer running *PsyScope* software (Cohen et al. 1993). Speech latency data (reaction times) were recorded using a Sony F-25 microphone connected to a *PsyScope* button box, which acted as a voice key and provided millisecond timing accuracy. The microphone was positioned approximately 25 cm from the participant's mouth. Auditory stimuli were presented via a Koss R/80 stereo headset. All instructions were given in English.

In the first phase of the experiment, the study phase, the participants listened to 46 words (36 study words, 5 primacy and 5 recency fillers, used as list protectors) presented with a 5-s inter-stimulus interval (ISI). As the participants were listening to the spoken words, they performed (depending on the condition) a secondary task. In the semantic condition, the purpose of the task was to orient the participants' attention during processing to the meaning-based properties of the words. The participants were asked to rate word pleasantness (pleasantness of the word meaning) on a 7-point scale (1 = *meaning is unpleasant*, 7 = *meaning is pleasant*). Upon hearing each word, they circled the appropriate rating on an answer sheet.

Used in studies investigating depth of (semantic) processing (e.g., Schacter and Church 1992), instructions to rate word pleasantness increase the likelihood that language users access word meaning during processing. In rating word pleasantness, the participants appeared to use the entire rating scale; the resulting ratings, used for attentional manipulation only, were not analyzed further. In the control condition, no attentional orientation was imposed. However, to make this condition as comparable as possible to the semantic condition, the participants were asked to track the presentation order for each spoken word in a list. Upon hearing each word, they circled the number (from 1 to 46) that corresponded to each word's serial position on a list. Thus, the two conditions imposed different attentional demands on the participants (attention to word meaning vs. no explicit attentional orientation) but included a similar manual (motor) response (circling each word's pleasantness rating vs. checking off each word's serial position on a list).

Following the study phase of the experiment, the participants performed a simple 3–4 min distractor task in which they solved 48 simple arithmetic problems (e.g., $45 - 19 = ?$) printed on a sheet of paper. The purpose of this task was to clear the participants' short-term memory and to allow some time to pass between the study and the test phases of the experiment. In the second phase of the experiment, immediately following the distractor task, the participants performed an immediate repetition task (Onishi et al. 2002; Trofimovich 2005). In this task, the participants listened to 80 words (72 test words, 4 primacy and 4 recency fillers, again used as list protectors) presented with a 5-s ISI and repeated each word as rapidly and as accurately as possible speaking into the microphone. A research assistant was present throughout the testing session, noting all instances of erroneous word repetitions (e.g., *thirty* for *dirty*) and failures of the voice key. The participants were equally frequently assigned to the four study-test list pairs.

Data Analysis

Because the objective of this experiment was to assess the amount of processing facilitation for known, as opposed to novel, L2 words, it was important that the participants know or at least be fairly familiar with the English words. (See Stark and McClelland 2000, for a comparison of word priming for words and non-words in English.) The participants' word familiarity was established in a word-knowledge test administered at the end of the testing session. The participants were asked to translate each English word into Chinese and then rate on a 7-point scale (1 = *I don't know the word at all*, 7 = *I know the word very well*) how familiar they were with the word. The obtained translation accuracy counts and word familiarity ratings were then used as criteria for identifying well known and familiar words. The words selected for further analysis were all translated accurately and were rated, on average, as highly familiar ($M = 6.6$, $SD = .20$).

The primary measure used in this experiment was response latency, defined as the length of time (in milliseconds) between the offset of the stimulus word and the onset of the participant's response (repeating the word), as measured by a voice key. The advantage of using offset-to-onset latency (e.g., Onishi et al. 2002), as opposed to measuring response latency from the onset of the stimulus word (e.g., Mullenix et al. 1989), is that this measure does not include the duration of the stimulus word and, therefore, is insensitive to differences in how quickly or slowly stimulus words are pronounced. Response latencies were excluded from further analyses if the voice key was not properly triggered (1.5% of the trials) or if the target word was repeated erroneously (5.3% of the trials). Erroneous word repetitions included all incorrect productions of the intended word that were not simply its mispronunciations traceable to the participants' L1 background (e.g., substituting /l/ for /r/, as in *afraid* repeated as *aflaid*). As described above, also excluded were response latencies to unknown and unfamiliar words (i.e., words the participants rated low in familiarity and were unable to translate), which accounted for 3.0% of the trials. To remove outlier response latencies within each participant's data set, the data were winsorized by replacing the slowest and fastest 10% of the individual's response latencies by the next slowest or fastest response latency, separately for each of the eight conditions formed by crossing the processing orientation (semantic, control) and test list factors.

The response latencies were tabulated for each participant separately for the 18 repeated words spoken by the same speaker (i.e., in a familiar voice) between the study and the test phases of the experiment, for the 18 repeated words spoken by a different speaker of the opposite gender (i.e., in an unfamiliar voice) between the study and the test phases of the experiment, as well as for the 36 unrepeated words (i.e., words appearing only in the test phase of the experiment). For all statistical tests reported below, the alpha level for significance was set at .05. The effect sizes reported below are partial eta squared (η_p^2), calculated by dividing the effect sum of squares by the effect sum of squares plus the error sum of squares. For t-tests, effect sizes are reported as r. A Bonferroni procedure was applied to adjust the level of significance for all tests of simple main effects.

Results

Baseline Performance

Previous research has demonstrated that drawing inferences about the strength of priming effects among several treatment conditions may depend, at least in part, on having comparable baseline performance rates (Chapman et al. 1994). A baseline performance rate refers here to response latencies for unprimed (unrepeated) materials. Therefore, in the first analysis, the participants' response latencies for unrepeated words were submitted to a 2 × 2 (orientation × LOR) analysis of covariance (ANCOVA) with AOA used as a covariate. An ANCOVA was used to analyze these data because LOR, the variable of primary interest here, was found to be confounded with AOA. A test of homogeneity of regression slopes, carried out prior to conducting this ANCOVA, revealed that AOA did not interact with LOR or orientation, Fs(1, 48) < .77, ps > .39, $\eta_p^2 < .02$. Thus, the effect of AOA was comparable at both levels of these factors, an important ANCOVA assumption. This ANCOVA yielded no significant main effects or interactions, Fs(1, 47) < 1.35, ps > .25, $\eta_p^2 < .03$, suggesting that baseline performance rates were comparable (see Table 3 for response latencies for unrepeated words). It appears, then, that the participants were equally fast at responding to well-known unrepeated L2 words, regardless of LOR or processing orientation. The following analyses were used to determine whether these two factors influenced the extent to which the participants benefit from repeated L2 words, those spoken in familiar and unfamiliar voices.

Familiar Voice

Mean response latencies for unrepeated words and for repeated words spoken in a familiar voice were submitted to a 2 × 2 × 2 (repetition × orientation × LOR) ANCOVA with AOA used as a covariate. A test of homogeneity of regression slopes, carried out prior to conducting this ANCOVA, revealed that AOA did not interact with LOR or orientation, Fs(1, 48) < .43, ps > .52, $\eta_p^2 < .01$. This ANCOVA yielded only a significant repetition × LOR interaction, $F(1, 47) = 3.87$, $p = .05$, $\eta_p^2 = .08$, with no other significant effects,

Table 3 Mean response latencies in milliseconds (standard error) for spoken words as a function of length of residence (LOR) and condition

Words	Condition			
	Control		Semantic	
	Short LOR	Long LOR	Short LOR	Long LOR
Unrepeated	571.2 (35.8)	693.8 (52.0)	631.4 (55.2)	631.3 (62.7)
Repeated in familiar voice	559.2 (35.8)	639.4 (49.8)	602.8 (50.9)	591.8 (58.9)
Repeated in unfamiliar voice	568.3 (36.7)	629.9 (50.8)	613.9 (56.7)	616.0 (63.9)

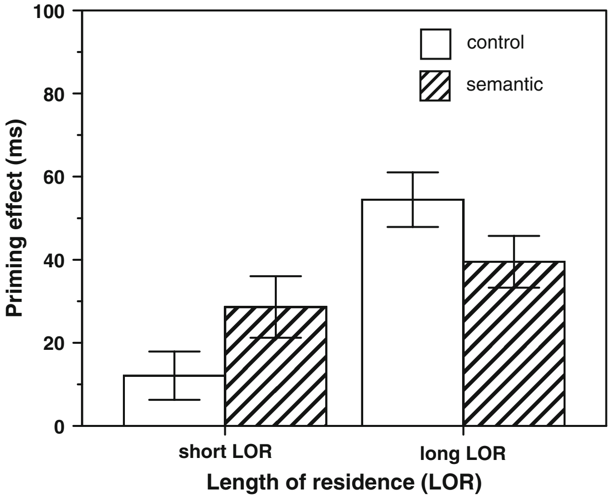

Fig. 1 Mean priming effects in the semantic and control condition for L2 words spoken in a familiar voice. Brackets enclose ±1 SE

Fs(1, 47) < 1.54, ps > .22, η_p^2 < .03. Tests of simple main effects were used to explore the significant repetition × LOR interaction ($\alpha = .0125$). The participants with a longer LOR demonstrated a significant processing facilitation in both conditions (47 ms on average), responding significantly faster to repeated words spoken in a familiar voice than to unrepeated words, $t(25) = 5.21$, $p = .0001$, $r = .72$. By contrast, the participants with a shorter LOR did not show a significant processing facilitation in either condition (20 ms on average); their facilitation effect failed to reach statistical significance after the Bonferroni correction, $t(25) = 2.17$, $p = .04$, $r = .39$ (see Table 3). For ease of graphical presentation of the results, mean priming effects (mean response latency for repeated words subtracted from mean response latency for unrepeated words), which represent the amount of processing facilitation obtained, were calculated for each group in each condition and plotted in Fig. 1 (the higher the bar, the more processing facilitation). In sum, processing facilitation effects for words spoken in a familiar voice were found only for the participants with a longer LOR in both the control and the semantic conditions.

Unfamiliar Voice

Mean response latencies for unrepeated words and for repeated words spoken in an unfamiliar voice were submitted to a similar 2 × 2 × 2 (repetition × orientation × LOR) ANCOVA

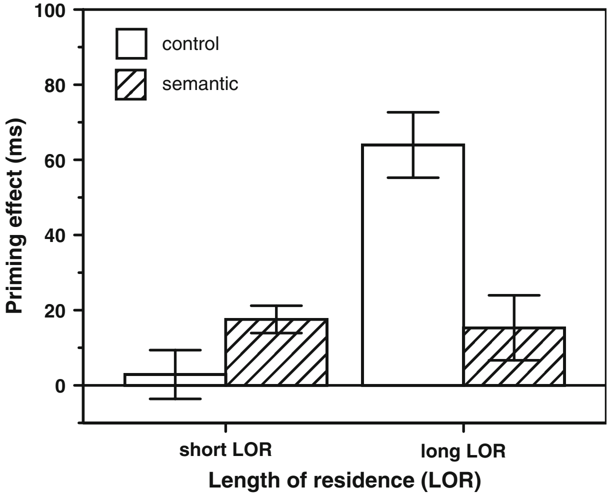

Fig. 2 Mean priming effects in the semantic and control condition for L2 words spoken in an unfamiliar voice. Brackets enclose ±1 SE

with AOA used as a covariate. A test of homogeneity of regression slopes, carried out prior to conducting this ANCOVA, revealed that AOA again did not interact with LOR or orientation,Fs(1, 48) < .34, ps > .56, $\eta_p^2 < .01$. This ANCOVA yielded a significant repetition × LOR interaction, $F(1, 47) = 5.87$, $p = .019$, $\eta_p^2 = .11$, and a significant repetition × orientation × LOR interaction, $F(1, 47) = 4.89$, $p = .032$, $\eta_p^2 = .09$, with no other significant effects, Fs(1, 47) < 3.27, ps > .08, $\eta_p^2 < .06$. The significant three-way interaction was explored further using tests of simple main effects ($\alpha = .0125$). These tests yielded only one significant finding. The participants with a longer LOR demonstrated a significant processing facilitation in the control condition only (64 ms), responding faster to repeated words spoken in an unfamiliar voice than to unrepeated words, $t(12) = 3.68$, $p = .003$, $r = .73$. The participants with a shorter LOR in the control condition and the participants with both shorter and longer LORs in the semantic condition showed no significant processing facilitation (12 ms on average), ts(12) < 2.34, ps > .03, rs < .55 (see Table 3). As in the previous analysis, the mean priming effects, calculated to graphically depict the amount of processing facilitation for each group in each condition, are plotted in Fig. 2. In sum, processing facilitation effects for words spoken in an unfamiliar voice were found only in the control condition and only for the participants with a longer LOR.

Processing Benefits and Language Proficiency and Use

To explore the relationship between the amount of processing facilitation obtained and the participants' L2 proficiency and use, zero-order correlations were computed between the priming scores for words spoken in a familiar and unfamiliar voice (mean response latency for repeated words subtracted from mean response latency for unrepeated words), on the one hand, and measures of the participants' proficiency (speech rate ratio, error rate, accentedness, comprehensibility, fluency) and their self-rated amount of English use in reading, writing, listening, and speaking, on the other. Because the preceding analyses established that the semantic processing orientation had a measurable effect on the amount of processing facilitation obtained, these correlations were computed separately for the participants in the control and in the semantic condition. A Bonferroni adjusted alpha level for significance was used because of a large number of correlations computed ($\alpha = .006$). For the participants in the control condition ($n = 26$), these analyses yielded only one association that approached significance. The amount of processing facilitation for words spoken in a familiar voice was positively associated with the participants' self-estimated amount of time spent listening to English on a daily basis (i.e., listening to English media), $r(25) = .45$, $p = .01$. For the participants in the semantic condition ($n = 26$), by contrast, these analyses yielded no significant associations.

Discussion

Summary of Findings

The objective of this study was to determine how LOR, used here as a general index of the participants' L2 experience, and an explicit semantic processing orientation influence the amount of processing facilitation for spoken L2 words. For L2 words spoken in a familiar voice (repeated word forms and voices), the amount of processing facilitation appeared to depend only on LOR, such that the participants with a longer LOR showed greater processing facilitation than the participants with a shorter LOR, regardless of the processing orientation condition. This finding suggests that learners with more extensive L2 experience, as compared to those whose L2 experience is less extensive, are more likely to demonstrate processing benefits that are due (at a minimum) to context-specific phonological information in spoken L2 words. One possible implication of this finding is that L2 learners appear to be able, first, to encode a certain amount of acoustic–phonetic detail about spoken L2 words and, second, to use this information to aid in future re-processing of the same words. Another implication is that these processing benefits might accrue, as the findings of correlational analyses implied, as a result of learners' perceptual (listening) experience in the L2. These implications however must remain speculative until investigated further (e.g., in a detailed, perhaps longitudinal, study of L2 learners' sensitivity to various acoustic–phonetic detail in spoken L2 words).

For L2 words spoken in an unfamiliar voice (repeated word forms only), however, the amount of processing facilitation appeared to depend on both LOR and processing orientation. In the control condition (where no explicit attentional focus was imposed), the participants with a longer LOR showed greater processing facilitation than the participants with a shorter LOR, a finding identical to that obtained for words spoken in a familiar voice. By contrast, in the semantic condition, the participants with shorter and longer LORs alike demonstrated negligible processing benefits. It appears that semantic processing might obliterate any benefits, at least for the participants with a longer LOR, for words spoken in an unfamiliar (i.e., "physically" different) voice. One possible implication of this finding is that learners with more extensive L2 experience, as compared to those whose L2 experience is less extensive, seem to demonstrate processing benefits that are also due to abstract phonological information available in spoken L2 words. With more L2 experience, learners likely appear to make speaker-independent comparisons among instances of spoken words heard, being able to make abstract word-identity matches across them. However, this ability might depend on particular circumstances of learners' experience with spoken words. That is, learners may be less likely to make abstract, speaker-independent comparisons across spoken L2 words when they engage in their meaningful, semantic processing. Clearly, these findings must be replicated in future research before any strong conclusions regarding negative effects of meaningful, semantic processing can be drawn.

Taken together, the findings of the present study yielded preliminary evidence of the positive effects of L2 experience on learners' ability to benefit from their exposure to spoken L2 words and revealed at least one factor (semantic processing orientation) that might negatively influence this ability. The implications of these findings for L2 processing and learning are discussed next.

L2 Experience and Spoken Word Processing

The present study provided evidence that learners, with an increasing amount of L2 experience, might become more likely to encode both abstract (word identity) and specific (voice identity) characteristics of spoken L2 words and to use this information to aid in subsequent re-processing of the same words. With more L2 experience, learners thus may learn to generalize across multiple, non-identical instances of spoken L2 words (e.g., words spoken in a variety of L2 dialects, voices, intonation patterns, or speaking rates) and, therefore, to engage in a processing strategy that, with time, should help learners perceive and produce L2 words accurately (Barcroft and Sommers 2005; Bradlow and Pisoni 1999; Bradlow et al. 1997). These findings underscore similarities in L1 and L2 speech processing (Goldinger 1996; Sheffert 1998; Trofimovich 2005; Trofimovich and Gatbonton 2006) and, more importantly, yield insights into the nature of spoken word processing in L2 learning.

As the results of the present study suggest, the initial stages of L2 learning (at least within 3–4 years of learners' residence in the target country) are likely characterized by a shift in learners' spoken word processing strategy. Early in the learning process, learners first appear to engage in a processing that is speaker- and situation-specific, characterized by learners' inability to disregard variability in a speaker's voice and perhaps in other context-specific details of speech (Lively et al. 1994; Trofimovich 2005). As a result, learners most likely encode, store, and access spoken L2 words episodically, in the form of discrete instances or individual records of experience in a memory system that is not specialized for the storage of linguistic information (Goldinger 1996, Goldinger 1998; Nosofsky and Johansen 2000). Put differently, learners might initially over-rely, in their spoken word processing, on speaker- and context-specific details of spoken L2 words, encoding these details fully, as entire episodes of perceptual experience. These details likely act as perceptual "anchors", offering learners a degree of consistency in the otherwise variable (and, therefore, unreliable) spoken input.

With more experience, however, learners become sensitive to abstract information available in spoken L2 words, being able to *both* generalize across their multiple context-specific instances and, in doing so, to encode (at a minimum) a sufficient amount of perceptual detail. At this stage of the learning process, learners, therefore, begin to create abstract phonological representations for spoken L2 words, which likely develop as generalizations over previously stored context-specific (episodic) instances of words or longer sequences previously heard (Bybee 2001; Nosofsky et al. 1992; see Pierrehumbert 2003, for a related view of phonological learning). As a result, learners begin to encode, store, and access spoken words as lexical entities (Pallier et al. 2001), using a semantic (lexical) memory system specialized for the storage of linguistic information (see Forster 1985; Jiang and Forster 2001; McKoon et al. 1986, for further discussion). With an increasing amount of L2 experience, learners may thus eventually come to rely on abstract phonological representations in order to perceive and produce L2 speech, although these representations ultimately may not evolve to be entirely native-like (Pallier et al. 2001). In future research, it is important to determine exactly how learners create phonological representations for spoken L2 words, especially when they encounter such words for the first time, to describe the nature of these representations, and to identify the conditions (e.g., in terms of the nature and amount of L2 input or demands of a learning task) that facilitate their development.

Processing Orientation and Spoken Word Processing

The present study provided preliminary evidence that L2 learners, regardless of the amount of their L2 experience (1.3 vs. 3.5 years of residence in the target country, on average) may not be able to make abstract, speaker-independent comparisons across spoken L2 words when they engage in their meaningful, semantic processing. This negative effect of semantic attentional orientation on spoken word processing in the L2 is not surprising, given that similar semantic influences have been documented previously in the L1, albeit in visual word processing (Blaxton 1989; Jacoby 1983; Levy and Kirsner 1989; see also Talamas et al. 1999). One common finding of these and other investigations is that the amount of processing facilitation obtained crucially depends on the degree to which a learning task requires extensive processing of auditory (perceptual) information available in spoken words (Sheffert 1998) and, consequently, on the degree to which individuals attend to and notice these words (Crabb and Dark 1999; Hawley and Johnston 1991; Mulligan and Hartman 1996). For example, Hawley and Johnston (1991) reported processing benefits only for those visually presented L1 words that had been correctly identified (and, therefore, attended to and noticed) in an earlier phase of their experiment. If this explanation is valid, then the processing benefits for repeated spoken L2 words in the present study were reduced to the extent that the semantic processing orientation interfered with spoken word identification, leading to a less extensive processing of auditory/perceptual information available in spoken words.[1]

At least one way to conceptualize the relationship between spoken word processing, on the one hand, and a semantic attentional orientation, on the other, is to suggest that both tasks (one requiring learners, at least in part, to encode spoken word *forms*, the other directing their attention to spoken word *meanings*) compete for limited memory resources (Mulligan and Hornstein 2000; Stone et al. 1998). Research on dual-task performance and, more specifically, "bottleneck" (task-switching) models of memory (e.g., Pashler 1994, Pashler 1998) offer a useful theoretical framework for understanding this relationship. Bottleneck models assume that certain processing operations or tasks cannot proceed simultaneously if they require a single mechanism for their operation. That is, when two tasks compete, a processing bottleneck may result such that one or both tasks are stalled, delayed, or impaired (see Barcroft's TOPRA model for a related conceptualization in L2 lexical learning; Barcroft 2002). If lexical access (access to a spoken word's meaning) and the encoding of a spoken word form rely on the same processing mechanism or draw on shared attentional resources, then at least one of these tasks should be selected at the expense of the other when both compete for processing resources simultaneously. In the experiment reported here, the rating of word pleasantness in the semantic condition, the task that required the learners to access spoken word meanings, might have been prioritized over the encoding of word forms, thus constraining the amount of phonological detail the learners encoded and thereby reducing the

[1] An anonymous reviewer suggested that the use of phonological information in word processing might be investigated more directly under a processing condition that encourages learners' attention to form-related, auditory (acoustic) properties of speech. An experimental condition of this kind was not employed in this study because previous research (Trofimovich 2005) revealed that the amount of processing benefits under such circumstances does not differ from the benefits found in a baseline (control) condition (e.g., one in which learners simply listen to spoken L2 words).

extent of processing facilitation obtained. This interpretation of this study's findings clearly needs to be tested in future research.

Several factors may explain why learners, at least relatively early in L2 learning, tend to prioritize meaning-based, semantic processing of spoken words over their form-based, perceptual encoding, particularly when task demands are elevated and processing resources are limited. One factor may be related to adult L2 learners' processing strategies. For example, young children learning their L1 extract their first words based on acoustic–phonetic properties of spoken words given ample native-speaker input (Gleitman et al. 1984). Adult L2 learners, by contrast, typically add acoustic–phonetic labels (often without sufficient native-speaker input) to the words whose meanings they already know. These effects of prior learning, compounded with learners' propensity to access word meanings automatically in L1 spoken discourse (Swinney and Love 2002), thus, make it likely that learners would tend to process the available L2 input for meaning before encoding the perceptual details of its form (VanPatten 1996).

Another factor may be related to the inherent complexity of L2 input and its elevated demands on processing resources. It is likely that learners, at least early in L2 learning, may have difficulty resolving subtle perceptual cues related to spoken word form, especially in situations where input is complex and learning conditions are not ideal: for example, when listening to words presented in noise (Rosenberg and Jarvella 1970) or when concurrently completing a secondary task (as in this experiment). In these situations, learners might, therefore, be expected to rely on their already predominant strategy of processing L2 input for meaning at the expense of form.

Yet another factor explaining why meaning-based processing of spoken words is preferred over their form-based encoding may be related to learners' short-term phonological memory capacity. Learners' short-term memory capacity is frequently limited, at least early in L2 learning (Brown and Hulme 1992; Cheung et al. 2000). Because there seems to exist a causal relationship between short-term phonological memory and learners' ability to repeat (Papagno et al. 1991) and eventually learn (Gathercole et al. 1991) unfamiliar spoken words, learners' limited short-term phonological memory capacity may constrain the amount of perceptual detail they comprehend, promoting their reliance on semantic and conceptual processing of spoken L2 words. In future research, it is important to examine the individual and combined influence of all these factors to determine the extent to which L2 learners benefit from their experience with spoken L2 words.

Conclusion

The present study was motivated by a general observation that the study of L2 lexical learning from a purely semantic perspective may overlook the significant role of spoken word processing in L2 learning. Experimental evidence was then provided suggesting that spoken word processing has measurable benefits in the L2 and that these benefits might depend on learners' amount of L2 experience (defined here as length of residence in L2 environment) and on attentional demands of a learning task. Results suggested that learners are sensitive to the "structure" of spoken L2 input (conceptualized here as form-based phonological properties of spoken words), implying that this sensitivity is a likely consequence of implicit learning driven by linguistic input (Church and Fisher 1998; Ellis 2002; see Nosofsky and Johansen 2000, for theoretical motivations). The findings overall provide evidence for the complex interaction of meaning-based semantic and form-based auditory/perceptual processing in L2 word learning and underscore the importance of investigating psycholinguistic (processing) bases of such learning.

Acknowledgements This research was supported by a grant from the Fonds québécois de la recherche sur la société et la culture (FQRSC). The author gratefully acknowledges the assistance of Randall Halter, Hyojin Song, Candace Farris, and Alin Zdrite with various aspects of this study, and wishes to thank Randall Halter, Sarita Kennedy, and two anonymous reviewers for their helpful suggestions on earlier drafts of this manuscript.

Appendix A

English Words

abroad	country	income	review
advice	demand	judgment	seldom
afraid	dirty	laughter	severe
again	discover	magic	sister
agreement	effort	manager	stomach
almost	even	many	story
alone	expensive	matter	suddenly
already	factory	mistake	summary
answer	failure	officer	supper
area	finger	often	today
attitude	follow	pattern	together
average	foreign	pencil	unique
avoid	forgive	people	water
become	government	properly	welcome
belong	happy	provide	wonderful
building	healthy	random	
ceiling	hungry	ready	
city	husband	remain	
corner	improve	remember	

References

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575–589.

Baker, W., & Trofimovich, P. (2006). Perceptual paths to accurate L2 production of L2 vowels: The role of individual differences. *International Review of Applied Linguistics in Language Teaching, 44*, 231–250.

Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning, 52*, 323–363.

Barcroft, J. (2003). Effects of questions about word meaning during L2 Spanish lexical learning. *Modern Language Journal, 87*, 546–561.

Barcroft, J., & Sommers, M. S. (2005). Effects of acoustic variability on second language vocabulary learning. *Studies in Second Language Acquisition, 27*, 387–414.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.

Beck, I. L., McKeown, M. G., & Omanson, R. C. (1987). The effects and use of diverse vocabulary instruction techniques. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 147–163). Hillsdale, NJ: Lawrence Erlbaum.

Beckman, M. E., & Edwards, J. (2000). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development, 71*, 240–249.

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 657–668.

Bogaards, P., & Laufer, B. (Eds.). (2004). *Vocabulary in a second language*. Amsterdam: John Benjamins.

Bower, G. H., & Winzenz, D. (1970). Comparison of associative learning strategies. *Psychonomic Science, 20*, 119–120.

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*, 2074–2085.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101*, 2299–2310.

Brand, N., & Jolles, J. (1985). Learning and retrieval rate of words presented auditorily and visually. *The Journal of General Psychology, 112*, 201–210.

Brown, G. D. A., & Hulme, C. (1992). Cognitive psychology and second-language processing: The role of short-term memory. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 105–122). North Holland: Elsevier.

Brown, T. S., & Perry, F. L., Jr. (1991). A comparison of three learning strategies of ESL vocabulary acquisition. *TESOL Quarterly, 25*, 17–32.

Bybee, J. (2001). Frequency effects on French liaison. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 337–359). Amsterdam: John Benjamins.

Channell, J. (1988). Psycholinguistic considerations in the study of L2 vocabulary acquisition. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 83–96). London: Longman.

Chapman, L. J., Chapman, J. P., Curran, T. E., & Miller, M. B. (1994). Do children and the elderly show heightened semantic priming? How to answer the question. *Developmental Review, 14*, 159–185.

Cheung, H., Kemper S., & Leung, E. (2000). A phonological account for the cross-language variation in working memory processing. *Psychological Record, 50*, 373–386.

Church, B., & Fisher, C. (1998). Long-term auditory word priming in preschoolers: Implicit memory support for language acquisition.*Journal of Memory and Language, 39*, 523–542.

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intention and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 521–533.

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

Cooper, W. E., Blumstein, S. E., & Nigro, G. (1975). Articulatory effects on speech perception: A preliminary report. *Journal of Phonetics, 3*, 87–98.

Crabb, B. T., & Dark, V. J. (1999). Perceptual implicit memory requires attentional encoding. *Memory & Cognition, 27*, 267–275.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268–294.

Crow, J. T., & Quigley, J. R. (1985). A semantic field approach to passive vocabulary acquisition for reading comprehension. *TESOL Quarterly, 19*, 497–513.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143–188.

Engle, R. N., Clark, D. D., & Cathcart, J. S. (1980). The modality effect: Is it a result of different strategies? *Journal of Verbal Learning and Verbal Behavior, 19*, 226–239.

Engle, R. W., & Roberts, J. S. (1982). How long does the modality effect persist? *Bulletin of the Psychonomic Society, 19*, 343–346.

Fisher, C., Hunt, C., Chambers, K., & Church, B. (2001). Abstraction and specificity in preschoolers' representations of novel spoken words. *Journal of Memory and Language, 45*, 665–687.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition, 23*, 527–552.

Forster, K. I. (1985). Lexical acquisition and the modular lexicon. *Language and Cognitive Processes, 1*, 87–108.

Gathercole, S. E., & Conway, M. A. (1988). Exploring long-term modality effects: Vocalization leads to best retention. *Memory & Cognition, 16*, 110–119.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics, 12*, 349–367.

Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language, 11*, 43–79.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1166–1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251–279.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the Critical-Period Hypothesis for second language acquisition.*Psychological Science, 14*, 31–38.

Hall, J. W., Wilson, K. P., & Patterson, R. J. (1981). Mnemotechnics: Some limitation of the mnemonic keyword method for the study of foreign language vocabulary. *Journal of Educational Psychology, 73*, 345–357.

Hawley, K. J., & Johnston, W. A. (1991). Long-term perceptual memory for briefly exposed words as a function of awareness and attention. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 807–815.

Henning, G. H. (1974). Remembering foreign language vocabulary: Acoustic and semantic parameters. *Language Learning, 23*, 185–196.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*, B47–B57.

Jackson, A., & Morton, J. (1984). Facilitation of auditory word recognition. *Memory and Cognition, 12*, 568–574.

Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior, 22*, 485–508.

Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language, 44*, 32–51.

Kasper, L. F. (1983). The effect of linking sequence and interactive picture mnemonics on the acquisition of Spanish nouns by middle school children. *Journal of Practical Research and Applications, 2*, 141–156.

Kelly, P. (1992). Does the ear assist the eye in the long-term retention of lexis? *International Review of Applied Linguistics, 30*, 137–145.

Kirsner, K., Brown, H., Abrol, S., Chadha, N. K., & Sharma, N. K. (1980). Bilingualism and lexical representation. *Quarterly Journal of Experimental Psychology, 32*, 585–594.

Kroll, J. F., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals: The development of lexical and conceptual representations. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 104–129). Cambridge: Blackwell.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence: Brown University Press.

Lawson, M. J., & Hogben, D. (1996). The vocabulary-learning strategies of foreign-language students. *Language Learning, 46*, 101–135.

Levy, B. A., & Kirsner, K. (1989). Reprocessing text: Indirect measures of word and message level processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 407–417.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America, 96*, 2076–2087.

Masson, M. E. J., & MacLeod, C. M. (2000). Taking the "text" out of context effects in repetition priming of word identification. *Memory & Cognition, 28*, 1090–1097.

McKoon, G., Ratcliff, R., & Dell, G. S. (1986). A critical evaluation of the semantic-episodic distinction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 295–306.

Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics: Abstracts, 15*, 221–246.

Meara, P. (1982). Word associations in a foreign language. *Nottingham Linguistic Circular, 11*, 29–38.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, A. Howart, & C. Criper (Eds.), *Interlanguage* (pp. 225–235). Edinburgh: Edinburgh University Press.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519–533.

Mullenix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America, 85*, 365–378.

Mulligan, N. W., & Hartman, M. (1996). Divided attention and indirect memory tests. *Memory & Cognition, 24*, 453–465.

Mulligan, N. W., & Hornstein, S. L. (2000). Attention and perceptual priming in the perceptual identification task. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 626–637.

Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of “multiple-system” phenomena in perceptual categorization. *Psychonomic Bulletin and Review, 7*, 375–402.

Nosofsky, R. M., Kruschke, J. K., & McKinley, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules.*Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 211–233.

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition, 83*, B13–B23.

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based vs. abstract lexical entries. *Psychological Science, 12*, 445–449.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30*, 331–347.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220–244.

Pashler, H. (1998). *The psychology of attention*. Cambridge: MIT Press.

Pierrehumbert, J. B. (2003). Probabilistic phonology: Discrimination and robustness. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 177–228). Cambridge: MIT Press.

Pilotti, M., Bergman, E. T., Gallo, D. A., Sommers, M., & Roediger, H. L., III (2000). Direct comparison of auditory implicit memory tests. *Psychonomic Bulletin and Review, 7*, 347–353.

Pressley, M., Levin, J. R., & Delaney, H. D. (1982). The mnemonic keyword method. *Review of Educational Research, 52*, 61–91.

Pressley, M., Levin, J. R., Hall, J. W., Miller, G. E., & Berry, J. K. (1980). The keyword method and foreign word acquisition. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 163–173.

Pressley, M., Levin, J. R., & Miller, G. E. (1981). The keyword method and children’s learning of foreign vocabulary with abstract meanings. *Canadian Journal of Psychology, 35*, 283–287.

Rae, G. (1979). Effect of word meaningfulness on primary and secondary memory. *Journal of General Psychology, 101*, 175–181.

Rodgers, T. S. (1969). On measuring vocabulary difficulty: An analysis of item variables in learning Russian-English vocabulary pairs. *International Review of Applied Linguistics, 7*, 327–343.

Rosenberg, S., & Jarvella, R. J. (1970). Semantic integration and sentence perception. *Journal of Verbal Learning and Verbal Behavior, 9*, 548–553.

Sagarra, N., & Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *Modern Language Journal, 90*, 228–243.

Schacter, D. L., & Church, B. A. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 915–930.

Segalowitz, N., & Frenkiel-Fishman, S. (2005). Attention control and ability level in a complex cognitive skill: Attention-shifting and second language proficiency. *Memory and Cognition, 33*, 644–653.

Seibert, L. C. (1927). An experiment in leaning French vocabulary.*Journal of Educational Psychology, 18*, 294–309.

Sheffert, S. M. (1998). Voice-specificity effects on auditory word priming. *Memory & Cognition, 26*, 591–598.

Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics, 25*, 293–321.

Stark, C. E. L., & McClelland, J. L. (2000). Repetition priming of words, pseudowords, and nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 945–972.

Stevens, G. (2004). Using census data to test the critical-period hypothesis for second language acquisition. *Psychological Science, 15*, 215–216.

Stone, M., Ladd, S. L., Vaidya, C. J., & Gabrieli, J. D. E. (1998). Word identification priming for ignored and attended words. *Consciousness and Cognition, 7*, 238–258.

Swinney, D., & Love, T. (2002). Context effects on lexical processing during auditory sentence comprehension: On the time-course and neurological bases of a basic comprehension process. In E. Witruk & A. D. Friederici (Eds.), *Basic functions of language, reading and reading disability: Neuropsychology and cognition* (Vol. 20, pp. 25–40). Netherlands: Kluwer.

Talamas, A, Kroll, J. F., & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second-language vocabulary. *Bilingualism: Language and Cognition, 2*, 45–58.

Trofimovich, P. (2005). Spoken-word processing in native and second languages: An investigation of auditory word priming. *Applied Psycholinguistics, 26*, 479–504.

Trofimovich, P., & Baker, W. (2006). Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28*, 1–30.

Trofimovich, P., & Gatbonton, E. (2006). Repetition and focus on form in L2 Spanish word processing: Implications for pronunciation instruction. *Modern Language Journal, 90*, 519–535.

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

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex.

Vochatzer, K. G., & Blick, K. A. (1989). Levels of processing and the retention of paired-associates. *Perceptual and Motor Skills, 69*, 349–350.

Walsh, D. A., & Jenkins, J. J. (1973). Effects of orienting tasks on free recall in incidental learning: "Difficulty," "effort," and "process" explanations. *Journal of Verbal Learning and Verbal Behavior, 12*, 481–488.