

# First-language phonotactics in second-language listening

Andrea Weber<sup>a)</sup>

Saarland University, Saarbrücken, Germany

Anne Cutler

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

(Received 25 July 2005; revised 27 October 2005; accepted 31 October 2005)

Highly proficient German users of English as a second language, and native speakers of American English, listened to nonsense sequences and responded whenever they detected an embedded English word. The responses of both groups were equivalently facilitated by preceding context that both by English and by German phonotactic constraints forced a boundary at word onset (e.g., *lecture* was easier to detect in *moinlecture* than in *gorklecture*, and *wish* in *yarlwish* than in *plookwish*). The American L1 speakers' responses were strongly facilitated, and the German listeners' responses almost as strongly facilitated, by contexts that forced a boundary in English but not in German (*thrarshlecture*, *glarshwish*). The German listeners' responses were significantly facilitated also by contexts that forced a boundary in German but not in English (*moycelecture*, *loitwish*), while L1 listeners were sensitive to acoustic boundary cues in these materials but not to the phonotactic sequences. The pattern of results suggests that proficient L2 listeners can acquire the phonotactic probabilities of an L2 and use them to good effect in segmenting continuous speech, but at the same time they may not be able to prevent interference from L1 constraints in their L2 listening. © 2006 Acoustical Society of America. [DOI: 10.1121/1.2141003]

PACS number(s): 43.71.Hw, 43.71.Es [ARB]

Pages: 597–607

## I. INTRODUCTION

A speaker of English who does not know the word *lope* is nonetheless likely to parse the sequence *lope for* correctly, because any other boundary assignment would give rise to an un-English syllable: *lowpf or* would yield a postvocalic /pf/ cluster, and *low pfor* would yield a prevocalic /pf/ cluster. English speakers know from experience that the cluster /pf/ does not occur in either position in their language. Phonotactic constraints on the possible structures that syllables may assume constitute part of the implicit knowledge that listeners have amassed about the phonology of their language. Such knowledge can be usefully exploited when new words are encountered.

In fact, phonotactic knowledge is useful in all listening situations, even if no unfamiliar words are presented. The rules and probabilities applying to the sequencing of phonemes provide a rich source of information for listeners. They can tell listeners where the speech signal definitely contains word or syllable boundaries, as in the above-noted example; but they can also tell where such boundaries are more versus less likely to occur. This potential source of information in the recognition of continuous speech (by human listeners or by machines) was pointed out many years ago (e.g., Church, 1987; Lamel and Zue, 1984; Harrington, Watson, and Cooper, 1989). Its use by human listeners has been amply attested in perceptual experiments. Several such experiments have used the wordspotting task, which is particularly suited for examining detection of word boundaries.

In a wordspotting experiment, listeners attempt to find real words embedded in nonsense input. For instance, in the string *obzel crivthish dullfim* they should detect the word *dull*. The relative difficulty of detection associated with a particular boundary can affect the speed with which listeners correctly detect the embedded word, and indeed the likelihood of the word being detected at all.

Sequencing constraints are easy to manipulate in a wordspotting experiment. If *dull* were to be presented in *dullfbim* instead of *dullfim*, for instance, we would expect detection to be harder, since the sequence /fb/ should contain a word boundary (as in: *we found the gulf bigger than expected*) and thus there would be no natural boundary aligned with the edge of the word *dull*. In consequence, we would expect listeners to take longer finding *dull*, or perhaps not to find it at all. The study that first clearly showed this effect was carried out in Dutch. McQueen (1998) showed that listeners detected *vel* “rim” faster and more accurately in *velbroel* versus *velmroel*, and *rok* “skirt” likewise faster and more accurately in *fimrok* versus *fidrok*. The sequence /mr/ in each of these pairs must contain a boundary in Dutch, just as in English; this information helps listeners find *rok* in *fimrok*, but hinders them with *vel* in *velmroel*.

Dumay, Frauenfelder, and Content (2002) established that constraints of the same type affect wordspotting in French; thus *lac* “lake” is easier to spot in *zunlac*, in which its edge is aligned with a phonotactically obligatory syllable boundary, than in *zuglac*, in which this is not the case.

Probabilistic information can be used in the same way: van der Lugt (2001), in an analogous experiment in Dutch, showed that words with common onsets are easier to spot than words with rare onsets—e.g., it is easier to find *galg* “gallows” in *piengalg* than *geur* “aroma” in *piengeur*. In

<sup>a)</sup>Electronic mail: aweber@coli.uni.sb.de

Dutch, many words begin *ga-* but very few begin *geu-*; van der Lugt argued that listeners translate their experience of this pattern in the vocabulary into probabilistic heuristics for finding word boundaries in speech.

Both kinds of information are also used by Korean listeners; word boundaries are prohibited in Korean in the sequence [dʒi], legal but not highly likely in the sequence [di], and highly likely in the sequence [nni]; correspondingly, Korean listeners detect words like *imin* “migration” more easily in [pjodimin] than in [pjodʒimin] and easiest of all in [pjonnimin] (Warner *et al.*, 2005).

Sensitivity to legality or probability of phoneme sequences also appears in phoneme detection in many languages (e.g., Gow, 2001, for English; Hallé *et al.* 1998, for French; Otake *et al.* 1996, for Japanese; Weber, 2001, for German; Yip, 2000, for Cantonese), as well as in artificial language learning (Saffran, Newport, and Aslin, 1996), goodness ratings of nonsense words (Vitevitch *et al.* 1997), lexical decision and word repetition (Vitevitch and Luce, 1999). Note that the nature of transitional probability effects depends upon the task that listeners have to carry out—thus although van der Lugt (2001) found that many words beginning in the same way makes spotting such words easier, lexical decision and repetition responses are slower for words with many onset competitors (Vitevitch, 2002).

The ability to exploit such rules and probabilities comes, of course, from linguistic experience. What is relevant is experience with a particular language, because phonotactic syllable boundary constraints and probabilities are specific to particular vocabularies. Thus English may proscribe both pre- and postvocalic /pf/ clusters, but other languages happily allow them in both positions. German is one such language—*Kopf* “head” and *Pferd* “horse” are both German words. Similarly, for McQueen’s (1998) Dutch listeners, *vel* was particularly easy to find in *velbroel* and *rok* particularly hard in *fidrok* because /b/ and /d/ (and all voiced obstruents) do not occur in syllable-final position in Dutch. Therefore a syllable boundary had to precede these sounds. This obviously is not true of English, in which voiced obstruents may occur syllable-finally, so that words like *bulb* and *feed* are perfectly legal.

The language-specificity of phonotactic sequencing constraints has two consequences for the relevance of such information in second-language listening. First, native (L1) listeners’ greater experience with the language should better enable them to make use of this kind of information than second-language (L2) listeners, who will have enjoyed much less relevant experience. Experiments with listening to totally unfamiliar languages have shown that phoneme sequence effects shown by L1 listeners do not appear with naive listeners (Otake *et al.*, 1996; Weber, 2001). Thus it is not the case that these effects are primarily signal-driven; they are experience-driven, and L2 listeners will have had less of the relevant experience than L1 listeners.

A second, and potentially more dangerous, consequence of the language-specificity of phonotactics is the potential transfer of the L1 phonotactic knowledge to L2 listening. Just as L2 listeners show effects of the L1 phonemic inventory in L2 phonetic processing (see Strange, 1995, for re-

views), so might L1 phonotactics interfere and in consequence hinder L2 listening, when L2 listeners hear sequences that are prohibited syllable-internally in the L1 but not in the L2, or are allowed syllable-internally in the L1 but prohibited in the L2. Since the smallest viable word in any language is a syllable, segmentation of continuous speech into its component words necessarily makes reference to syllable boundaries (Norris *et al.*, 1997). Thus a German speaker of English might not readily perceive that *lope for* has to contain a syllable boundary between /p/ and /f/, because the /pf/ sequence can occur syllable-internally in German. An English speaker of German, in contrast, might falsely posit a boundary between /p/ and /f/ in *die Pferde* or *Kopf auf*, on the grounds that a boundary must occur in a /pf/ sequence in English.

The present study tests whether L1 phonotactic effects in L2 listening do indeed occur. We confronted L2 listeners—Germans listening to English—with phonotactic sequences that in their L1, German, would signal a boundary (but that should not be informative in the L2, English). The listeners were highly proficient in English, though their accumulated experience with the language would not match that of an L1 speaker of the same age. We asked—using the task that McQueen (1998), van der Lugt (2001), Dumay *et al.* (2002), and Warner *et al.* (2005) had shown to be sensitive to lexical segmentation via phonotactic information—whether the listeners would be influenced by the German constraints when listening to English. We compared their responses in this condition with conditions involving sequences that would indeed signal a boundary in English, and sequences that would signal a boundary in both languages or in neither language, and we also compared their responses with those of American English L1 listeners. We also ascertained what acoustic information the materials contained that might potentially signal the presence of a word or syllable boundary, and we determined whether the presence of this information was correlated with the pattern of each listener group’s responses.

## II. MATERIALS CONSTRUCTION

With the wordspotting paradigm it is possible to compare detection of the same word in different contexts. This is a particularly valuable feature of this paradigm in the present case, because with L2 listening it is hard to control the properties of words that experimenters usually attempt to keep constant in lexical processing experiments. Word frequency counts, for example, may not reflect the experience of L1 and L2 listeners in the same way. Such controls are unnecessary, however, if the manipulations of interest can be accomplished within items.

To make this possible, we required English words that could be placed in four different preceding contexts: one in which the preceding phoneme provided a phonotactically clear boundary in both English and German, one in which the preceding phoneme provided a clear boundary in English but not in German, one in which the preceding phoneme provided a clear boundary in German but not in English, and one in which the preceding phoneme provided a clear bound-

ary in neither language. For the latter two cases it was thus necessary to select words beginning with phonemes that could be the second element in more than one kind of English cluster. This severely constrains the range of available words. It rules out exploiting the pf- case, for example, because /f/ hardly occurs as second element in a cluster in English words (*sphere, sphinx*). We chose to use words beginning with /l/ and /w/. These phonemes make it possible to construct all four types of context.

In both English and German, words can begin with the singleton consonant /l/, and its pronunciation in syllable onset position is analogous in the two languages. Such close parallels cannot be achieved with /w/, since this labiovelar approximant does not occur in German; instead, German has the voiced fricative /v/, e.g., in cognate words (E: *winter, wool, warm*; G: *Winter, Wolle, warm*). Substitution of /v/ for /w/ is perhaps the most noticeable marker of a German accent in English (*ve have vays...*), suggesting that the English sound is assimilated to the German category. These two onset phonemes thus provide a further interesting contrast regarding the application of phonotactic constraints, since one can be directly mapped to L1 constraints, while the other can only be mapped via prior assimilation of the L2 sound to a L1 category.

English and German phonotactic constraints differ (Giegerich, 1992; Wiese, 1996). In English, words may begin with /s/, and there are many /s/-initial clusters (*sleep, swim, sport, stand, script*, etc.). Thus a word preceded by a context ending with /s/ is not aligned with a phonotactically clear boundary. The situation is quite different in German. Words in standard German may not begin with /s/. Thus a word beginning with a consonant is aligned with a clear boundary if the preceding context ends with /s/. In contrast, German words may begin with /ʃ/, and there are many /ʃ/-initial clusters (*Schlaf, schwimmen, Sport, stehen, Schrift*, etc.). A word preceded by a context ending with /ʃ/ is thus not aligned with a clear boundary. Again, there is a contrast with English, that has no native words beginning with /ʃ/-initial clusters (only a couple of Yiddish loanwords in American English, such as *schlep*). Thus a word beginning with a consonant has a very strong probability of being aligned with a clear boundary if the preceding context ends with /ʃ/.

The phonemes /s/ and /ʃ/, plus /t/ for some w-words, thus provided our cross-linguistically mismatching boundaries: /s/l/, /sw/, and /tw/ are possible onset clusters in English but not in German, /ʃ/l/ and /ʃw/ are impossible onsets in native English words but are possible in analogous German words. The two other boundary conditions were: phonotactically clear in both languages (/nl, ml, nw, mw, lw/ are impossible onsets in English, and impossible either directly or in analogue in German), and phonotactically clear in neither language (/pl, kl, fl, kw/ are all possible onsets in English and, directly or analogously, in German). Again, cognates abound: E: *plan, clear, flame, quality*; G: *Plan, klar, Flamme, Qualität*). All our selected words were thus preceded by four contexts, supporting No boundary, an English boundary, a German boundary, and a Common boundary, respectively. For *lecture*, this produced *gorklecture* (No), *thrarshlecture* (E), *moycelecture* (G), *moinlecture* (C); for *wish*, *plookwish* (No),

*glarshwish* (E), *loitwish* (G), *yarlwish* (C).

From the CELEX lexical database (Baayen, Piepenbrock, and van Rijn, 1993) we selected 68 words, 36 1-initial and 32 w-initial, mono- or bisyllabic (e.g., *lunch, lecture, wish, weapon*); all words and their contexts are listed in the Appendix. As far as possible, these words contained no other embedded English or German words, though it was not possible to avoid some embeddings leaving a single-consonant residue, such as *lea* in *league*; such embeddings are usually overlooked in wordspotting: Norris *et al.*, 1997). Each word was provided with four monosyllabic preceding contexts as described earlier; some contexts were used for several words. Contexts could contain /l/ and /w/. All context monosyllables contained long vowels, diphthongs or vowels plus /r/, so that they would be phonotactically legal syllables without their coda; thus the internal structure of the context syllable did not itself force a particular segmentation (e.g., *moyce lecture* and *moy slecture* are both permissible segmentations in English).

A further 55 nonsense sequences were constructed that began with context syllables similar to those in the experimental items, and contained, embedded in final position, an English word beginning with a consonant other than /l/ or /w/ (e.g., *hooshdonkey*). A further 251 bi- or trisyllabic nonsense sequences without embedded word were also constructed (e.g., *crivthish*); these sequences could contain /l/ or /w/, and the first syllable could end with any of the sounds in the boundary sequences in the experimental items. These were filler materials for the wordspotting study.

All 272 (68 × 4) experimental items and 306 fillers were recorded via a Sennheiser ME64 microphone onto Digital Audio Tape in a sound-attenuated booth by a phonetically trained female native speaker of American English (born in the Midwest, who had lived as a child and teenager in four different states). The speaker was instructed to avoid any clear syllable boundaries in her productions. Subsequently, the speaker also recorded the 136 (68 × 2) experimental sequences potentially ambiguous in English twice each, without pause but with different intended syllable boundaries (e.g., *gor clecture, gork lecture, moy slecture, moyce lecture, ploo kwish, plook wish, loy twish, loit wish*). These additional productions were for acoustic analysis to establish the dimensions of variation that might potentially cue boundaries in the experimental materials.

The recordings, initially sampled at 44.1 kHz, were downsampled to 16 kHz and stored on disk. Using XWAVES software, measurements were made of segmental durations in each target word and its context, as well as in the additional clear-boundary productions of the 136 potentially ambiguous sequences. Also using XWAVES software, the initial syllable was removed from each of the 272 experimental sequences and 306 fillers. This manipulation produced the stimuli for a control experiment using the lexical decision task.

### III. ACOUSTIC ANALYSES

Existing literature on acoustic cues to juncture motivated five measurements in the clear-boundary productions of the

potentially ambiguous (No boundary, German boundary) sequences. The aim of these measurements was to identify features of our speaker's production that might vary across items in the experimental materials, and thereby affect the ease of detection of the embedded words. Not all measurements applied to all stimuli.

- (a) *Glottalization*. Spectrograms of the boundary sequences /kl, pl, tw, kw/ (63 pairs) <sup>1</sup> were examined for signs of glottalization of the stop consonant, a potential marker of syllable-finality (e.g., *loit wish*; Pierrehumbert, 1994).
- (b) *Stop VOT*. For the same sequences, we measured the duration from stop burst onset to onset of voicing; this was expected to be longer in sequences with no boundary, i.e., with a string-initial stop (*loy twish*; Keating, Wright, and Zhang, 1999).
- (c) *Fricative duration*. For the sequences /fl, sl, sw/ (58 pairs) we measured fricative duration from the offset of voicing of the preceding vowel to the offset of high-frequency fricative noise. Conflicting predictions were motivated for this measure; it might be longer when the sequence contained no boundary so that the fricative was string-initial (e.g., *moy slecture*; Klatt, 1974), or shorter in the same sequences containing a boundary due to shortening of phonemes within a cluster (Haggard, 1973).
- (d) *Duration of /l/*. For the sequences /fl, sl, pl, kl/ (68 pairs) we measured duration of the voiced portion of /l/, from the onset of voicing to a rise in the second formant and an increase in amplitude of the third formant. This measure was expected to be longer when the sequence contained a boundary so that the /l/ was in a singleton onset rather than a cluster (*moyce lecture*; Haggard, 1973).
- (e) *Vowel duration*. For all items (120 pairs), we measured duration of the vowel in the first syllable. Vowels were defined as beginning at the onset of voicing after voiceless obstruents, and as beginning at the onset of the second and third formant after voiced segments. When the vowel in the context syllable was followed by /r/, /r/ was included in the vowel duration measurement. Vowels were expected to be longer when the two-consonant boundary sequence did not contain a boundary, i.e., when the measured vowel was syllable-final (e.g., *moy slecture*; Lehiste, 1960).

#### Results and discussion.

- (a) *Glottalization* is apparently not a feature of our speaker's productions; it was observed on only two productions (both /tw/ sequences containing a boundary).

The mean durations for measures (b) to (e) are given in Table I. Two-factor analyses of variance (ANOVA) were conducted on each value set, with consonant sequence as between-items factor and boundary placement as within-items factor.

- (b) *Stop VOT* showed no boundary placement effect ( $F[1, 60]=2.86, p>0.09$ ).
- (c) *Fricative duration* showed an interaction between

TABLE I. Mean durations in milliseconds for the two intended syllabifications.

Measure	Boundary in consonant sequence ( <i>moyce lecture</i> )	No boundary in consonant sequence ( <i>moy slecture</i> )
VOT (stops; e.g., [t] in <i>loitwish</i> )	85	79
fricative duration (e.g., [s] in <i>moy slecture</i> )	153	136
voiced duration of /l/ (e.g., [l] in <i>moy slecture</i> )	44	24
First syllable vowel duration (all items; e.g., [ɔ] in <i>moyce, loit</i> )	150	177

boundary placement and consonant sequence (the only interaction of these factors); boundary placement was therefore analyzed in each consonant sequence separately. Fricatives were significantly longer before a boundary in the case of /fl/ ( $F[1, 10]=24.89, p<0.001$ ) and /sl/ ( $F[1, 33]=47.77, p<0.001$ ), but not in the case of /sw/ ( $F[1, 12]=1.13$ ).

- (d) *Duration of /l/* was increased by a boundary in the consonant sequence ( $F[1, 64]=126.92, p<0.001$ ).
- (e) *Vowel duration* was increased when there was no boundary in the consonant sequence ( $F[1, 115]=117.03, p<0.001$ ).

Duration of the vowel in the context syllable, and, for /l/-initial targets, duration of /l/ and duration of a preceding fricative, are therefore factors that, if they vary in our speaker's productions of the wordspotting materials, might produce effects on listeners' responses. Recall that the speaker, a trained phonetician, had tried in producing those materials to avoid durational differences between the conditions. Inspection of the acoustic measures of the experimental materials (performed for the English, German, and No boundary conditions only) showed that to a considerable degree she was successful in this. The measured duration of the context vowel, which in the clear-boundary productions was shorter when the word was aligned with a boundary (*moyce lecture*) than when it was not (*moy slecture*), was not shorter in the English or German boundary constraint conditions than in the No boundary condition, for either word type; it was slightly longer. Thus if listeners attended to this information, it could counteract the effects of the phonotactic boundary constraints.

For /l/-initial targets, the duration of the voiced portion of /l/ (in the clear-boundary productions, longer when the word was aligned with a boundary) also patterned in the wrong direction for the English boundary condition (at 22.6 ms it was 10% shorter than the 26.8 ms of the No boundary productions of the same items). However /l/ was on average 32% longer (35.5 ms) in the German boundary condition than in the No boundary condition. The duration of a preceding fricative in the boundary constraint conditions could not be compared with the No boundary condition, which had no fricative, but the average duration differed by

TABLE II. Control lexical decision experiment with American English participants: (a) mean RTs in milliseconds, measured from item offset, and (b) mean percentage of missed responses, with standard errors of each mean, as a function of word-initial phoneme and boundary context in the wordspotting materials from which the items were excised.

Measure	Initial sound	Common boundary	English boundary	German boundary	No boundary
(a) RT	l-words	401 (18.4)	382 (18.5)	372 (18.3)	361 (17.2)
	w-words	354 (22.4)	394 (21.4)	399 (22.2)	348 (15.9)
(b) Miss	l-words	14.7 (1.9)	16.5 (2.1)	11.6 (2.1)	12.5 (1.8)
	w-words	11.2 (2.2)	14.6 (2.5)	16.4 (3.3)	11.0 (2.0)

only 4% in the German (137.2 ms) versus the English boundary condition (142.3 ms). These analyses suggest that for some words in some conditions, acoustic variation in the wordspotting materials may provide listeners with useful information.

#### IV. CONTROL DATA FROM LEXICAL DECISION

The wordspotting task resembles lexical decision (in which listeners decide whether a spoken form is a real word), but the focus is on the effects of the adjacent context rather than on properties of the word itself. Although in a typical wordspotting experiment the same word is presented in different contexts, natural productions of the items can lead to differences in the acoustic form of the word across contexts. A standard way of controlling for such effects is to excise the words from their contexts and present them for lexical decision. The task is the same—respond whenever a real word is heard—but the context is removed while the acoustic form of the word remains. If the different contexts had different effects on the acoustic goodness of the word itself, in a way that would affect ease of recognition and so confound the wordspotting data, this should be apparent in lexical decision responses to the words alone.

*Participants.* Forty native speakers of American English, students at the University of South Florida, took part in the experiment in return for course credit or a small monetary compensation.

*Materials and Procedure.* The excised real words (*lecture*, *wish*, *donkey*, etc.) and nonwords (*thish*, etc.) formed four lists, each list containing one truncated version of each of the 68 experimental words and all 306 truncated filler items. The first 14 items in each list formed a practice set. No version of an experimental word occurred in more than one list; the four types of context were counterbalanced across lists. Ten listeners heard each list over headphones. They were asked to press a response button whenever they spotted a real English word, and to repeat aloud each word that they had found. This form of the lexical decision task mimicked the word-spotting task as closely as possible. Stimulus presentation and response collection was controlled by a portable computer running NESU experimental control software. Both response times (RT; measured from item onset) and accuracy were recorded.

*Results and discussion.* In any lexical decision task responses cannot be made until word offset (*donkey* could always turn out to be *donkeef*). Hence the measured whole-

word duration was subtracted from the recorded RTs to give RTs from that point.<sup>2</sup> Table II shows the mean RT from word offset for detections which were accompanied by a correct oral response, and the mean miss rate (no response, response without correct oral report), as a function of the original context from which the word had been excised. ANOVAs were conducted across participants and across items to test for effect of original context on these responses. Across items, all main effects and interactions were insignificant in RTs and miss rates; across participants, the only significant effect was Condition ( $F [3, 108]=3.26, p < 0.03$  for RTs,  $F [3, 108]=3.22, p < 0.03$  for miss rates). Subsequent t-tests showed that the words excised from the No boundary condition (predicted to be hardest in wordspotting) produced faster RTs than words in the other three conditions, and fewer misses than words in the English boundary condition. The acoustic form of the words thus did not differ across contexts in any way likely to produce the pattern of results we predict in wordspotting; on the contrary, acoustic effects, if operative, would tend to work against our predictions.

#### V. WORDSPOTTING EXPERIMENTS

*Participants.* Ninety-six participants were tested. Forty-eight were L1 listeners, native speakers of American English, mostly students at the University of South Florida; they took part in the experiment in return for course credit or a small monetary compensation. These participants had no knowledge of German. The other forty-eight were native speakers of German, students of English translation and interpretation at the University of Heidelberg, Germany; they received a small monetary compensation for taking part. These participants had received an average of 15 years of training in English as a foreign language, beginning at a mean age of 11. Their knowledge of English was in consequence excellent.

*Materials and Procedure.* The wordspotting materials, comprising four versions of each of the 68 experimental words with appended context, and 306 filler items, were arranged in four lists, each list containing one version of each of the 68 experimental words and all 306 filler items. The first 14 items in each list formed a practice set. No version of an experimental word occurred in more than one list; the four types of context were counterbalanced across lists. The order of the lists was the same as in the lexical decision experiment; the order was pseudo-random, with each experimental item occurring after at least one filler item containing no real word.

TABLE III. Wordspotting experiments: (a) mean RTs in milliseconds, measured from item offset, and (b) mean percentage of missed responses, with standard errors of each mean, for American English L1 listeners vs German listeners as a function of word-initial phoneme and boundary constraint exercised by the context.

		Boundary constraint			
		Common boundary ( <i>moinlecture</i> )	English boundary ( <i>thrarshlecture</i> )	German boundary ( <i>moycelecture</i> )	No boundary ( <i>gorklecture</i> )
(a) RT					
English	l-words	482 (22.5)	543 (24.3)	516 (23.6)	596 (26.1)
L1 listeners	w-words	466 (25.9)	468 (21.2)	593 (31.8)	528 (30.9)
German	l-words	608 (33.6)	662 (33.5)	651 (31.9)	676 (34.1)
listeners	w-words	542 (27.5)	589 (29.2)	570 (26.7)	678 (36.8)
(b) Miss					
English	l-words	11.9 (1.6)	17.2 (2.2)	18.3 (1.9)	26.0 (2.8)
L1 listeners	w-words	16.9 (2.8)	17.3 (2.4)	24.3 (2.6)	24.6 (3.0)
German	l-words	18.6 (2.1)	24.2 (2.6)	24.5 (2.5)	33.1 (3.0)
listeners	w-words	19.9 (2.5)	20.4 (2.6)	22.2 (2.5)	32.1 (3.0)

In each participant group, 12 listeners heard each list over headphones. They were asked to press a response button whenever they spotted a real English word at the end of one of the nonsense sequences, and to repeat aloud each word that they had found. Their response times and accuracy were recorded. Except for the brand of portable computer used, all aspects of the testing setup were the same in each location. Stimulus presentation and timing were as for the lexical decision experiment.

The crucial predictions for the wordspotting results concern the relative effect of the different boundary constraints across the listener groups. First, we predict that both listener groups should show significant effects of common boundary constraints. We test this by comparing RT and accuracy in the Common boundary condition (*lecture* in *moinlecture*) with the No boundary condition (*lecture* in *gorklecture*); responses in the Common condition should be faster and more accurate. Second, English-speaking L1 listeners should show an effect of English constraints, and our highly proficient L2 listeners may do so. We test this by comparing RT and accuracy in the English boundary condition (*lecture* in *thrarshlecture*) with the No boundary condition; responses in the English condition should be faster and more accurate. Third, German boundary constraints may affect German listeners (but should not affect L1 listeners). We test this by comparing RT and accuracy in the German boundary condition (*lecture* in *moycelecture*) with the No boundary condition; any effect should show up as faster and more accurate responses in the German condition.

*Results and discussion.* Table III shows the mean RTs (again adjusted to measure from word offset) for detections accompanied by a correct oral response, and mean miss rates (no response, response before word offset, or response without correct oral response), for each listener group in the four conditions. Seven words have been removed from these results (six missed by 50% or more of the German listeners, and one, *leisure*, with multiple pronunciations—rhyming with *measure* or *seizure*—known to our listener group), leav-

ing 61 experimental words. It can be seen that, unsurprisingly, the L1 participants produced overall faster RTs and lower miss rates than the L2 participants. It is noteworthy, though, that both the RT and miss rate distributions for the two listener groups overlap; that is, these German listeners had, with the exception of the seven excluded words, little trouble performing the wordspotting task in their L2.

ANOVAs were conducted across participants (F1) and across items (F2) to examine the effects of participant language, context and word-initial phoneme. The first two comparisons were within-item and the last two within-participant. There was a main effect of language (L1 listeners were faster and more accurate: F1 [1,88]=9.8,  $p < 0.005$ , F2 [1,59]=82.1,  $p < 0.001$  for RTs, F1 [1,88]=5.94,  $p < 0.02$ , F2 [1,59]=9.65,  $p < 0.005$  for miss rates) and of context (F1 [3,264]=19.33,  $p < 0.005$ , F2 [3,177]=82.1,  $p < 0.001$  for RTs, F1 [3,264]=27.12,  $p < 0.001$ , F2 [3,177]=8.02,  $p < 0.001$  for miss rates), but not of word-initial phoneme. Initial phoneme also did not interact with either of the other main effects, but language and context did interact (F1 [3,264]=2.63,  $p = 0.05$ , F2 [3,177]=2.63,  $p = 0.05$ , RTs only).

The crucial predictions, as described earlier, concern the relative effects of context in each listener group, i.e., the components of the language by context interaction, and we test the predictions for each context type in turn via separate t-tests, across the mean RTs and miss rates for participants (t1) and items (t2), for each listener group. Because the ANOVA also revealed a three-way interaction between the effects (F1 [3,264]=6.47,  $p < 0.001$ , F2 [3,177]=8.05,  $p < 0.001$ , RTs only), separate ANOVAs were additionally conducted for each listener group. The main effect of context remained significant and the main effect of initial phoneme remained insignificant in each of these analyses, but the interaction between context and initial phoneme differed: it was insignificant for the German listeners but significant for the L1 listeners (F1 [3,132]=8.43,  $p < 0.001$ , F2 [3,177]

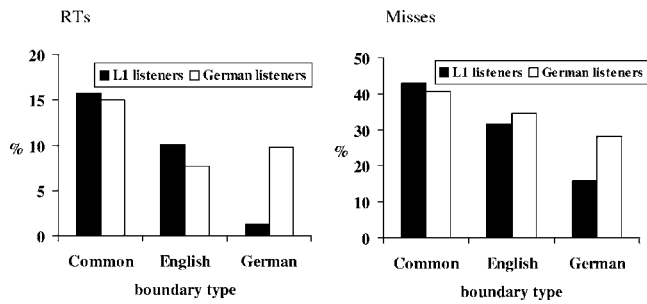


FIG. 1. Wordspotting results: effect of boundary type, expressed as mean percentage reduction in RT and miss rate, for each listener group in each boundary constraint condition (Common, English, German) compared with the No boundary condition.

$=7.42, p < 0.001$ ). To understand this interaction we carried out separate t-tests examining the context effects separately for each word type for this listener group.

The effects for each listener group of each constraint type (Common, English, German) are displayed in Fig. 1 as the percentage reduction in RT and miss rate compared with the No boundary condition. Because the average RT and average miss rate of the two listener groups were significantly different, as can only be expected in a comparison of L1 vs L2, displaying the reduction as a percentage better allows comparison of the effect size across the two groups. The higher the bar in Fig. 1, the larger the effect size; it can be seen that a common boundary has a large effect for both groups (i.e., responses are significantly faster and more accurate when there is a common boundary than when there is no boundary), an English boundary also has an effect for both groups, but a German boundary has more effect for German listeners.

**Common constraints.** The first set of t-tests compared responses in the Common boundary condition and the No boundary condition; as predicted, for both listener groups, responses in the Common condition were faster and more accurate (for L1 listeners,  $t_1[47]=5.25, p < 0.001, t_2[60]=5.06, p < 0.001$  for RTs and  $t_1[47]=4.32, p < 0.001, t_2[60]=3.51, p < 0.001$  for miss rates; for German listeners,  $t_1[47]=4.53, p < 0.001, t_2[60]=3.77, p < 0.001$  for RTs and  $t_1[47]=5.73, p < 0.001, t_2[60]=3.98, p < 0.001$  for miss rates). The difference between these two conditions was also highly significant for both the l-words and the w-words in the separate t-tests on the L1 listeners' RTs.

**English constraints.** The second set of t-tests compared responses in the English boundary condition versus the No boundary condition; as predicted, L1 listeners' responses were faster and more accurate in the English boundary condition ( $t_1[47]=2.81, p < 0.01, t_2[60]=3.01, p < 0.005$  for RTs and  $t_1[47]=3.58, p < 0.001, t_2[60]=2.36, p < 0.025$  for miss rates). This effect also held for RTs to l-words, and for w-words was significant across items ( $p < 0.05$ ) although not ( $p=0.067$ ) across participants. The German listeners' responses also were faster and more accurate in the English boundary condition, though the RT effect across items narrowly missed significance ( $t_1[47]=2.08, p < 0.05, t_2[60]=1.81, p=0.075$  for RTs and  $t_1[47]=4.19, p < 0.001, t_2[60]=3.08, p < 0.003$  for miss rates). Thus these highly

proficient L2 listeners were able to make use of boundary constraints in their L2.

**German constraints.** The third set of t-tests compared the German boundary condition to the No boundary condition. The German listeners responded faster and more accurately in the German boundary condition ( $t_1[47]=2.83, p < 0.01, t_2[60]=2.25, p=0.03$  for RTs and  $t_1[47]=4.13, p < 0.001, t_2[60]=3.1, p < 0.003$  for miss rates). For L1 listeners, this overall comparison was insignificant across both items and participants for RTs, and insignificant across items though significant across participants for miss rates ( $t_1[47]=2.07, p < 0.05$ ). Separate t-tests across word type revealed the L1 listeners' RTs to l-words to be faster in the German boundary condition than in the No boundary condition ( $t_1 p < 0.005, t_2 p < 0.005$ ), but their RTs to w-words to be slower in the German boundary condition than in the No boundary condition ( $t_1 p=0.068, t_2 p < 0.04$ ).<sup>3</sup> This latter difference between l-words and w-words is presumably the source of the significant interaction of context and initial phoneme for these listeners' RTs, noted above.

The wordspotting results thus showed that common boundary constraints affected both listener groups' responses equivalently, English constraints affected L1 listeners' responses strongly and German listeners' responses almost as strongly, while German constraints affected only the German listeners' responses.

The fact that the L1 listeners' response pattern in the latter comparison was exactly opposite for l-words versus w-words suggests that these differences may not reflect the phonotactic information but some other factor. In search of such a factor, we conducted correlations of the acoustic measures with listeners' RTs. The only significant correlations were with // duration: the longer the //, the faster l-words were spotted, both by L1 listeners ( $r[101]=-0.24, p < 0.015$ ) and German listeners ( $r[101]=-0.33, p < 0.001$ ). Since segmental lengthening is known to be a cue to word boundaries in running speech (Shatzman and McQueen, in press), we would predict that this correlation in the results of wordspotting (which measures segmentation) would have no counterpart in the results of the control lexical decision experiment (which measures recognition of the same tokens without segmentation being needed). Indeed, the lexical decision RTs showed no significant correlation with // duration ( $r[101]=-0.08, p > 0.4$ ).

It appears that both listener groups could take advantage of subtle acoustic cues to boundaries where these were available. As described earlier, however, the only comparison in which the // duration cue was usefully available was the comparison between the German boundary context and the No boundary context, with the cue here clearly favoring the words in the German boundary context. The result was that L1 listeners and L2 listeners alike detected the words more rapidly in this context. No acoustic cues favored detection of w-words in the German boundary context over the No boundary context, however. In the absence of acoustic cues, the phonotactic cues had a clear facilitatory effect for the German listeners, but not for the L1 listeners.

## VI. GENERAL DISCUSSION

Figure 1 says it all: The advantage of a common phonotactic boundary constraint in a spoken sequence is robust and almost identical for both our listener groups, the advantage of an English-only boundary constraint is again robust and not significantly different for the L1 versus the L2 listener groups, but the advantage of a German-only boundary constraint is robust only for the German-native L2 listeners. Thus first-language phonotactic constraints can influence second-language listening. German listeners to English found it easier to spot English words embedded in a nonsense context like *moycelecture*, where the word-initial phoneme could not be combined with the preceding phoneme to form a syllable-initial cluster in German, than when the preceding phoneme plus word-initial phoneme could be a syllable-initial cluster in either English or German (e.g., *gorklecture*). This facilitatory effect appeared for words beginning both with /l/ and with /w/. This pattern did not appear in the responses of American English listeners to the same materials.

Words aligned with a boundary signaled by a constraint common to both English and German (e.g., *moinlecture*) were easiest to detect, for both L1 and German listeners. Words aligned with a highly likely boundary in English only (e.g., *thrarshlecture*) were likewise easy for both L1 and German listeners to detect. These results replicate for American English the effects of phonotactic constraints on word segmentation already demonstrated with the same listening task in Dutch (McQueen, 1998), French (Dumay *et al.*, 2002) and Korean (Warner *et al.*, 2005). More significantly, the latter result shows that experienced L2 listeners can exploit constraints specific to their L2.

The finding that the common boundary constraints affected both listener groups' responses equivalently is open to alternative interpretations. It could mean that each listener group was making use of English boundary constraints or, in principle, it could mean that each listener group was making use of the boundary constraints appropriate for their L1. However, the finding that English boundary constraints affected German listeners' responses almost as strongly as they affected the responses of L1 listeners is unambiguous evidence that with those items the German listeners were able to exploit their knowledge of the probabilities relevant to the L2.

This is good news for L2 listeners: in this respect it is possible for L2 listening to approach the level of L1 listening. These listeners, it is true, were very proficient indeed in their L2; we cannot say on the basis of the present results exactly how much L2 experience is necessary for the efficient exploitation of L2 boundary constraints which do not apply in the L1. But note that the constraints in the English-boundary condition not only did not apply in German, they were equivalent to the No boundary condition if interpreted in terms of German phonotactics; in *thrarshlecture*, *lecture* is preceded by /ʃ/, and /ʃl/ is just as good a possible onset in German as /pl/ or /kl/ -*Schlaf*, *Plan*, *klar*. In this case the German listeners were thus able to suppress or ignore the probabilities of their L1 while listening to their L2. More-

over, they applied this English-only constraint not only to the set of w-words such as *weapon* and *wish*, that putatively show less phonetic resemblance to their L1, but also to the l-words such as *lunch* and *lecture*, even though the /ʃl/ boundary sequence was in this case phonetically very close indeed to potential L1 sequences.

Much less good news for L2 listeners emerges from the results from the German-only boundary condition. Even though these listeners are highly proficient in the L2, and even though they are clearly sensitive to the L2 probabilities, nevertheless they also show sensitivity to a boundary cue from their L1 which is absolutely not informative in the L2 they are listening to. Again, they applied this boundary constraint in the detection of both l-words—where it was putatively the case that the constraint could be translated directly from the L1—and w-words, where, as we argued earlier, application of the constraint can only proceed via a prior assimilation of the L2 category /w/ to the L1 category /v/. Evidence from German pronunciation of English suggests that such assimilation does occur; our results suggest that it does not inhibit further application to the L2 category of rules pertaining to the L1 category.

Acoustic correlates of the realization of word-initial /l/ were exploited by listeners of both groups in finding the words. The fact that both groups made similar use of these cues is again good news for L2 listening, though probably to be expected given that domain-initial strengthening effects appear consistently across languages (Keating *et al.*, 2003). As it happened, these acoustic cues provided useful evidence in the German-only boundary condition to a greater extent than in other conditions. The result of this was that the American English L1 listeners' detection of the l-words was facilitated in this condition. The fact that their responses to w-words were not facilitated in this condition—if anything, they showed the reverse pattern, although for w-words the difference did not reach our criterion of significance—strongly suggests that the German phonotactic constraints were having no facilitatory effect at all for this listener group. The facilitation in the l-words arose from the acoustic cues alone. Note that Kirk (2000), in a wordspotting study with American English, found that allophonic cues on the initial phoneme of an word affected how rapidly the word could be detected; *rock* in *voodrock*, for instance, was spotted more rapidly if the /dr/ sequence was pronounced without retroflexion than if it was pronounced with the retroflex articulation characteristic of an onset cluster. Our results with l-words provide further evidence that listeners are able to make good use of such cues in segmentation. We could find no such cues in the realization of the w-words. The German listeners' responses, significantly affected even for the w-words by the presence of the German-only boundary constraint, emphasize the relevance of this constraint in their L2 listening.

It is well known that L2 listening is influenced by expectations from the L1. Thus a mismatch between L1 and L2 phoneme repertoires can lead to L2 categories being interpreted in terms of distinctions used in the L1 (see Strange, 1995, for an overview). These phoneme confusions can cause pseudo-homophony, as when no difference can be per-



ceived between, for instance, *write* and *light* (Cutler and Otake, 2004; Pallier, Colomé, and Sebastián-Gallés, 2001). They further cause an increase in word activation (Broersma, 2002; Sebastián-Gallés, Echeverría, and Bosch, 2005), which can multiply several times the amount of lexical competition involved in a given word recognition situation (Cutler, 2005). To add to the L2 listener's word recognition problems, the native vocabulary may even be activated during listening to the L1 (Spivey and Marian, 1999; Weber and Cutler, 2004).

The segmentation of continuous speech in a second language is similarly subject to inappropriate influence from the L1. Thus English listeners do not segment their native language syllable by syllable or mora by mora, but French listeners to English apply syllabic segmentation, just as they do in listening to their L1 (Cutler *et al.*, 1986), and Japanese listeners to English apply moraic segmentation, just as for their L1 (Cutler and Otake, 1994). Inappropriate application of L1 phonotactic constraints in L2 listening thus joins the range of interference effects that L2 listeners must attempt to overcome in segmenting spoken language.

It should perhaps not be surprising that phonotactic constraints are well-anchored in listening, given that language acquisition studies have demonstrated that these constraints belong to the earliest information acquired about the mother tongue. Late in the first year of life, before they begin to talk, babies show evidence of a preference for speech that meets the sequencing constraints of their language over speech that does not (Jusczyk *et al.*, 1993; Jusczyk, Luce, and Charles-Luce, 1994). They also prefer legal over illegal word boundary clusters within their own language (Friederici and Wessels, 1993). Adult listeners, as we saw, are sensitive not simply to gross differences between legal and illegal patterns (Gow, 2001; Halle *et al.*, 1998; Otake *et al.*, 1996; Weber, 2001), but also to finer-grained differences in probabilities of acceptable sequences (Saffran *et al.*, 1996; van der Lugt, 2001; Vitevitch *et al.*, 1997; Yip, 2000).

However, our results indicate that listeners' knowledge of phonotactic probabilities is not always translated into the ability to exploit the information to maximum effect in on-line listening. The German listeners' responses in the English boundary condition showed that they had good mastery of the English probabilities and also of the difference between the probabilities of English and German (since they did not treat the sequences in this condition in the way appropriate for German). Nevertheless they were apparently unable to suppress the German boundary constraint probabilities in sequences such as *moycelecture* or *loitwish*. One possible interpretation of this asymmetry is that there is a difference between the effect of a positive boundary constraint (/s/ in the L1 must always contain a boundary, /ʃ/ in the L2 should always contain a boundary) versus the effect of probabilities indicating that a boundary is unnecessary (/ʃ/ in the L1 and /s/ in the L2 may be word-initial). In the spoken-word recognition model Shortlist (Norris *et al.*, 1997), which explicitly incorporates a role for phonotactic segmentation information, such a stronger effect of positive boundary information would indeed be predicted: Only clear boundary cues play a role in the model, so that strings with ambiguous

boundaries (e.g., *moycelecture*) must be resolved via competition between alternative word candidates. A recent study by Altenberg (2005) indeed shows that in parsing potentially ambiguous sequences Spanish listeners to English exploit positive boundary cues (e.g., a glottal stop before the initial vowel in *grave at*) far better than negative cues (e.g., the absence of a glottal stop before the medial vowel in *grey vat*).

Further, there is converging evidence that L2 listeners do not always capitalize maximally upon their phonological knowledge of the L2 in on-line recognition of spoken language. Broersma (2005) found that Dutch listeners to English performed within the L1 range on forced-choice categorization of syllable-final voiced versus voiceless obstruents (/s,z/, /p,b/, etc.). Despite this evidence that Dutch listeners can indeed distinguish voicing contrasts in syllable-final position, Cutler *et al.* (2004) found that listeners from the same population produced many syllable-final voicing errors in an open phonemic identification task, and Broersma (2002) observed many false-alarm positive responses due to voicing misperception in the same listener group's lexical decision performance (such that, for instance, *cheece* or *glope* was accepted as a word of English). Thus proficiency in distinguishing L2 contrasts in a forced-choice task may not be applied in other listening situations.

We note that visual word recognition also can be affected by such asymmetric application of the same type of phonotactic knowledge. Altenberg and Cairns (1983) asked English monolinguals and English-German bilinguals to rate nonwords such as *smatt* (legal in English but illegal in German) as potential English words. The ratings of both the bilinguals and the monolinguals were only influenced by phonotactic legality in English, suggesting that their knowledge of phonotactic constraints was equivalent. But when asked to decide whether a visually presented item was an English word or not (lexical decision), the bilinguals did not show the same pattern of response times as the monolinguals did. Bilinguals' decisions were affected by the phonotactic legality of the stimuli in both German and English, while monolinguals' decisions were only affected by phonotactic legality in English.

Mastery of a second language, in short, requires not only knowledge of the sequencing probabilities of the L2, but also the ability to translate this knowledge into on-line segmentation and recognition decisions. Our results suggest that the knowledge can be adequately mastered by L2 users—at least, by proficient L2 users such as those in our German listener group. However, translation of the knowledge into useful application is not necessarily achieved even by listeners with high proficiency. Especially positive cues to the presence of an L1 boundary can still provide unwanted interference in L2 listening.

## ACKNOWLEDGMENTS

This research was supported by a doctoral fellowship from the Max Planck Society to A.W., with additional support from the NWO-SPINOZA project "Native and Non-native Listening" to A.C. We thank Winifred Strange and

Elaine Griffiths for making it possible to test the Florida and Heidelberg participants, respectively. We further thank Diane Bradley, Taehong Cho, James Jenkins, James McQueen, Holger Mitterer, Kanae Nishi, and Natasha Warner for assistance and helpful commentary. Parts of the research were presented to the workshop “Spoken Word Access Processes” in Nijmegen, May 2000, and to the Sixth International Conference on Spoken Language Processing, Beijing, October 2000.

## APPENDIX

Target-bearing materials used in the wordspotting study. The seven words excluded from the analysis are marked with two asterisks.

### Embedded words with initial /l/

Common boundary	English boundary	German boundary	No boundary	Embedded word
/dʒimləns/	/ðijləns/	/blɔɪsləns/	/θɪpləns/	lance
/fumlɔft/	/prɑːfɔlft/	/forslɔft/	/zɑːplɔft/	loft
/dʒɑnlɪrɪk/	/dijlɪrɪk/	/nɔslɪrɪk/	/nɑːplɪrɪk/	lyric
/rɪnlɛd/	/rɑːflɛd/	/bɪrslɛd/	/bɔklɛd/	lead**
/dɔɪnlɔs/	/nuːflɔs/	/tɪrslɔs/	/mɔklɔs/	loss
/jɔɪnlɑrdʒ/	/gɪjɪlɑrdʒ/	/ðɑʊslɑrdʒ/	/fʊklɑrdʒ/	large
/gɑːnlɛzɔr/	/frɑːflɛzɔr/	/fɔːslɛzɔr/	/gɑːflɛzɔr/	leisure**
/pɑːnlɛtɔr/	/ruːflɛtɔr/	/pʊslɛtɔr/	/jɪflɛtɔr/	letter
/mɑːnlʊp/	/grɪflʊp/	/hɔːslʊp/	/nɑːflʊp/	loop
/pʊmlɪft/	/nɑːflɪft/	/rɪrslɪft/	/nʊplɪft/	lift
/fʊmlɔrd/	/krɪflɔrd/	/mɔːslɔrd/	/jɑːplɔrd/	lord
/pɔːnlɔndri/	/pɑːflɔndri/	/bɔːslɔndri/	/fɑːplɔndri/	laundry
/gɔːnlɑf/	/prɪflɑf/	/krɪrslɑf/	/fɔːklɑf/	laugh
/fʊnlɛŋθ/	/zɑːflɛŋθ/	/jɔːslɛŋθ/	/gʊklɛŋθ/	length
/hɑːnlʊntʃ/	/rɪflʊntʃ/	/fɑːslʊntʃ/	/jɔːklʊntʃ/	lunch
/wɑːnlɪst/	/fɑːflɪst/	/gɔːslɪst/	/bʊflɪst/	list
/hɔːnlɛvəl/	/dʒɪflɛvəl/	/bʊslɛvəl/	/pɑːflɛvəl/	level
/pɑːnlɑst/	/grɑːflɑst/	/pɔːslɑst/	/dɑːflɑst/	lust
/dʒɪmlɔ/	/θɪflɔ/	/grɪrslɔ/	/mɑːplɔ/	law
/fʊmlɑdʒɪk/	/jɪflɑdʒɪk/	/hɔːslɑdʒɪk/	/dɑːplɑdʒɪk/	logic
/wɔːmlɔn/	/tɑːflɔn/	/θɪrslɔn/	/jɪplɔn/	lawn
/pʊnlɪdʒən/	/dɑːflɪdʒən/	/bʊslɪdʒən/	/rʊklɪdʒən/	legion
/zɑːnlʌv/	/brɪflʌv/	/pɑːslʌv/	/fʊklʌv/	love
/ðɑːnlɪnən/	/dʒɪflɪnən/	/kɔːslɪnən/	/dɪklɪnən/	linen
/rɑːnlɛft/	/trɑːflɛft/	/jɔːslɛft/	/rɑːflɛft/	left
/ʃɑːnlɛðər/	/bʊflɛðər/	/hɪrslɛðər/	/mɑːflɛðər/	leather
/mɑːnlɑʊndʒ/	/krɑːflɑʊndʒ/	/nɔːslɑʊndʒ/	/trɑːflɑʊndʒ/	lounge
/pʊmlɑk/	/gɑːflɑk/	/mɔːslɑk/	/fʊplɑk/	lock
/dʒɪmlɛnz/	/frɑːflɛnz/	/tɔːslɛnz/	/rɑːplɛnz/	lens
/zɔːnlæg/	/drɑːflæg/	/rɔːslæg/	/gɑːplæg/	lag
/fɑːnlʌk/	/brɪflʌk/	/dɔːslʌk/	/wʊklʌk/	luck
/jɑːnliɡ/	/kɑːflɪɡ/	/jɔːslɪɡ/	/nɔːkliɡ/	league
/mɔːnlɛkʃər/	/θrɑːflɛkʃər/	/mɔːslɛkʃər/	/gɔːklɛkʃər/	lecture
/rɔːnlɑdʒ/	/fʊflɑdʒ/	/dɔːslɑdʒ/	/krɑːflɑdʒ/	lodge
/pɑːnlɑʊs/	/drɪflɑʊs/	/fɔːslɑʊs/	/tɑːflɑʊs/	louse
/fɑːnlɛpɔrd/	/dʒɑːflɛpɔrd/	/gɑːslɛpɔrd/	/gɪflɛpɔrd/	leopard

### Embedded words with initial /w/

/pɪrlwɔrm/	/flɑːflwɔrm/	/gɔːtwɔrm/	/zɔːkwɔrm/	warm
/fɪrlwɛl/	/nɑːflwɛl/	/mɔːtwɛl/	/mɔːkwɛl/	well
/nɑːrlwɔr/	/fɑːflwɔr/	/mɪrtwɔr/	/plʊkwɔr/	war
/jɑːrlwɔsp/	/grʊflwɔsp/	/nʊtwɔsp/	/ðɪkwɔsp/	wasp

### Embedded words with initial /w/

/tɑːrlwʊmən/	/fʊflwʊmən/	/θɑːslwʊmən/	/hɔːkwʊmən/	woman
/pɔːnlwɛðər/	/dɑːflwɛðər/	/dɔːslwɛðər/	/jɪkwɛðər/	weather
/zɑːnlwɔrm/	/grɪflwɔrm/	/grɔːslwɔrm/	/rʊkwɔrm/	worm
/pʊmlwɑɪf/	/rɪflwɑɪf/	/blɪrslwɑɪf/	/nɔːkwɑɪf/	wife
/gɪrlwʊnd/	/lɑːflwʊnd/	/gɑːtlwʊnd/	/θɪkwʊnd/	wound
/nɑːrlwɔrd/	/tɑːflwɔrd/	/tʃʊtlwɔrd/	/fʊkwɔrd/	ward
/dɑːrlwɔk/	/frʊflwɔk/	/nɪrtwɔk/	/bɔːkwɔk/	walk
/jɑːrlwɪfl/	/glɑːflwɪfl/	/lɔːtlwɪfl/	/plʊkwɪfl/	wish
/mɪrlwʊl/	/ðɪflwʊl/	/gɔːrslwʊl/	/bʊkwʊl/	wool**
/hɔːrlwɔrp/	/ruːflwɔrp/	/drɪrslwɔrp/	/pʊkwɔrp/	warp
/fɔːnlwɔrd/	/gʊflwɔrd/	/glɔːslwɔrd/	/gʊkwɔrd/	word
/fʊmlwɪn/	/nuːflwɪn/	/fɔːrslwɪn/	/θɔːkwɪn/	win
/rɑːrlwɔrf/	/prɪflwɔrf/	/hɔːtlwɔrf/	/dʊkwɔrf/	wharf**
/pɑːrlwɛɪ/	/gʊflwɛɪ/	/putwɛɪ/	/nʊkwɛɪ/	way
/kɪrlwæg/	/gɪflwæg/	/lɪrtwæg/	/ðɪkwæg/	wag**
/rɑːrlwɪŋ/	/klɑːflwɪŋ/	/fɔːtlwɪŋ/	/jɔːkwɪŋ/	wing
/bɪrlwɔrs/	/mʊflwɔrs/	/glɪrslwɔrs/	/gɔːkwɔrs/	worse**
/jɔːnlwɛpən/	/rɑːflwɛpən/	/fɔːslwɛpən/	/hɔːkwɛpən/	weapon
/fʊnlwʊd/	/plɑːflwʊd/	/mɪrslwʊd/	/frʊkwʊd/	wood
/dʒɪmlwʊm/	/pʊflwʊm/	/plʊrslwʊm/	/jɔːkwʊm/	womb**
/nɪrlwɛt/	/klʊflwɛt/	/mɔːtlwɛt/	/bʊkwɛt/	wet
/hɪrlwɑtʃ/	/jɪflwɑtʃ/	/zɔːtlwɑtʃ/	/lʊkwɑtʃ/	watch
/dɪrlwɛdɪŋ/	/pɑːflwɛdɪŋ/	/hɪrtwɛdɪŋ/	/jɔːkwɛdɪŋ/	wedding
/mɑːrlwəri/	/bʊflwəri/	/pɪrtwəri/	/dɔːkwəri/	worry
/gɪrlwɪk/	/lʊflwɪk/	/krɔːrslwɪk/	/prʊkwɪk/	week
/dɔːrlwɪntər/	/prɑːflwɪntər/	/nɔːrslwɪntər/	/mʊkwɪntər/	winter
/θɔːnlwɪnd/	/pɑːflwɪnd/	/θɪrslwɪnd/	/zɔːkwɪnd/	wind
/pʊmlwɪðθ/	/prʊflwɪðθ/	/pɔːrslwɪðθ/	/jɪkwɪðθ/	width

<sup>1</sup>Seven items excluded from the wordspotting results have also been excluded from these acoustic analyses. One further item, *wing*, was omitted by mistake from the recording session. The acoustic analyses are thus based on 60 items.

<sup>2</sup>The seven items excluded from the wordspotting results have also been excluded from the lexical decision analyses.

<sup>3</sup>A reviewer of this paper suggested that /t/-final contexts of w-words might have been fricated, rendering these items particularly difficult for English L1 listeners but possibly particularly easy for German listeners who might have been unfamiliar with this allophonic variation, so that they would hear it as an affricate rather than a stop, and thus be more likely to postulate a boundary. Inspection of the mean responses for the 14 w-words that had /t/-final contexts in the German boundary condition revealed that, contrary to this prediction, they were not relatively more difficult in this condition than in other conditions for English L1 listeners, and they were not relatively easier in this condition than in other conditions for German listeners (in fact they were harder). However, there was indeed some frication in these items; measured duration and amplitude of the fricated portion did not correlate with English L1 listeners' responses, but German listeners responded more rapidly the longer the fricated portion ( $t[13]=0.54, p < 0.05$ ). An analysis of the German listeners' results in the German boundary condition for items with /s/-final contexts only (i.e., omitting the 14 items with /t/-final contexts) showed responses to be still significantly faster ( $t[46]=2.14, p < 0.04$ ) and more accurate ( $t[46]=3.12, p < 0.003$ ) in this condition than in the No boundary condition.

Altenberg, E. P. (2005). "The perception of word boundaries in a second language." *Second Lang. Res.* **21**, 325–358

Altenberg, E. P., and Cairns, H. S. (1983). "The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects." *J. Verbal Learn. Verbal Behav.* **22**, 174–188.

Baayen, H., Piepenbrock, R., and van Rijn, H. (1993). "The CELEX Lexical database." Philadelphia: Linguistic Data Consortium, University of Pennsylvania (CD-ROM).

Broersma, M. (2002). "Comprehension of non-native speech: Inaccurate phoneme processing and activation of lexical competitors." in *Proceedings of the Seventh International Conference on Spoken Language Processing*.

- Denver, CO, pp. 261–264.
- Broersma, M. (2005). "Perception of familiar contrasts in unfamiliar positions." *J. Acoust. Soc. Am.* **117**, 3890–3901.
- Church, K. W. (1987). "Phonological parsing and lexical retrieval," *Cognition* **25**, 53–69.
- Cutler, A. (2005). "The lexical statistics of word recognition problems caused by L2 phonetic confusion," in Proceedings of INTERSPEECH 2005, Lisbon, pp. 413–416.
- Cutler, A., Mehler, J., Norris, D. G., and Seguí, J. (1986). "The syllable's differing role in the segmentation of French and English," *J. Mem. Lang.* **25**, 385–400.
- Cutler, A., and Otake, T. (1994). "Mora or phoneme? Further evidence for language-specific listening," *J. Mem. Lang.* **33**, 824–844.
- Cutler, A., and Otake, T. (2004). "Pseudo-homophony in non-native listening," 147th Meeting of the Acoustical Society of America (ASA), New York.
- Cutler, A., Weber, A., Smits, R., and Cooper, N. (2004). "Patterns of English phoneme confusions by native and non-native listeners," *J. Acoust. Soc. Am.* **116**, 3668–3678.
- Dumay, N., Frauenfelder, U. H., and Content, A. (2002). "The role of the syllable in lexical segmentation in French: Wordspotting data," *Brain Lang* **81**, 144–161.
- Friederici, A. D., and Wessels, J. M. I. (1993). "Phonotactic knowledge and its use in infant speech perception," *Percept. Psychophys.* **54**, 287–295.
- Giegerich, H. J. (1992). *English Phonology: An introduction* (Cambridge University Press, Cambridge).
- Gow, D. (2001). "Assimilation and anticipation in continuous spoken word recognition," *J. Mem. Lang.* **45**, 133–159.
- Haggard, M. (1973). "Abbreviation of consonants in English pre- and post-vocalic clusters," *J. Phonetics* **1**, 9–24.
- Hallé, P., Seguí, J., Frauenfelder, U., and Meunier, C. (1998). "Processing of illegal consonant clusters: A case of perceptual assimilation?," *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 592–608.
- Harrington, J., Watson, G., and Cooper, M. (1989). "Word boundary detection in broad class and phoneme strings," *Comput. Speech Lang.* **3**, 367–382.
- Jusczyk, P. W., Friederici, A. D., Wessels, J. M., Svenkerud, V. Y., and Jusczyk, A. M. (1993). "Infants' sensitivity to the sound patterns of native language words," *J. Mem. Lang.* **32**, 402–420.
- Jusczyk, P. W., Luce, P. A., and Charles-Luce, J. (1994). "Infants' sensitivity to phonotactic patterns in the native language," *J. Mem. Lang.* **33**, 630–645.
- Keating, P., Cho, T., Fougeron, C., and Hsu, C. (2003). "Domain-initial strengthening in four languages," in *Laboratory Phonology VI*, edited by J. Local, R. Odgen, and R. Temple (Cambridge University Press, Cambridge), pp. 145–163.
- Keating, P., Wright, R., and Zhang, J. (1999). "Word-level asymmetries in consonant articulation," *UCLA Working Papers in Phonetics* **97**, 157–173.
- Kirk, C. (2000). "Syllabic cues to word segmentation," in Proceedings of the Workshop on Spoken Word Access Processes (SWAP), edited by A. Cutler, J. M. McQueen, and R. Zondervan (Nijmegen, The Netherlands), pp. 131–134.
- Klatt, D. H. (1974). "The duration of [s] in English words," *J. Speech Hear. Res.* **17**, 51–63.
- Lamel, L., and Zue, V. W. (1984). "Properties of consonant sequences within words and across word boundaries," *Proceeding IEEE Int. Conf. Acoust. Sp., Sig. Processing (ICASSP)*, pp. 42.3.1–42.3.4.
- Lehiste, I. (1960). "An acoustic-phonetic study of internal open juncture," *Phonetica* **5**, 1–54.
- McQueen, J. M. (1998). "Segmentation of continuous speech using phonotactics," *J. Mem. Lang.* **39**, 21–46.
- Norris, D., McQueen, J. M., Cutler, A., and Butterfield, S. (1997). "The possible-word constraint in the segmentation of continuous speech," *Cogn. Psychol.* **34**, 191–243.
- Otake, T., Yoneyama, K., Cutler, A., and van der Lugt, A. H. (1996). "The representation of Japanese moraic nasals," *J. Acoust. Soc. Am.* **100**, 3831–3842.
- Pallier, C., Colomé, A., and Sebastián-Gallés, N. (2001). "The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries," *Psychol. Sci.* **12**, 445–449.
- Pierrehumbert, J. (1994). "Knowledge of variation," in Papers of the 30th Regional Meeting of the Chicago Linguistic Society, Chicago, University of Chicago, pp. 232–256.
- Saffran, J. R., Newport, E. L., and Aslin, R. N. (1996). "Word segmentation: The role of distributional cues," *J. Mem. Lang.* **35**, 606–621.
- Sebastián-Gallés, N., Echeverría, S., and Bosch, L. (2005). "The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals," *J. Mem. Lang.* **52**, 240–255.
- Shatzman, K. B., and McQueen, J. M. (in press). "Segment duration as a cue to word boundaries in spoken-word recognition," *Percept. Psychophys.*
- Spivey, M., and Marian, V. (1999). "Cross talk between native and second languages: Partial activation of an irrelevant lexicon," *Psychol. Sci.* **10**, 281–284.
- Strange, W. (1995). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (York, Timonium, MD).
- van der Lugt, A. H. (2001). "The use of sequential probabilities in the segmentation of speech," *Percept. Psychophys.* **63**, 811–823.
- Vitevitch, M. S. (2002). "Influence of onset density on spoken-word recognition," *J. Exp. Psychol. Hum. Percept. Perform.* **28**, 270–278.
- Vitevitch, M. S., and Luce, P. A. (1999). "Probabilistic phonotactics and neighborhood activation in spoken word recognition," *J. Mem. Lang.* **40**, 374–408.
- Vitevitch, M. S., Luce, P. A., Charles-Luce, J., and Kemmerer, D. (1997). "Phonotactics and syllable stress: Implications for the processing of spoken nonsense words," *Lang Speech* **40**, 47–62.
- Warner, N., Kim, J., Davis, C., and Cutler, A. (2005). "Use of complex phonological patterns in processing: Evidence from Korean," *J. Linguistics* **41**, 353–387.
- Weber, A. (2001). "Help or hindrance: How violation of different assimilation rules affects spoken-language processing," *Lang Speech* **44**, 95–118.
- Weber, A., and Cutler, A. (2004). "Lexical competition in non-native spoken-word recognition," *J. Mem. Lang.* **50**, 1–25.
- Wiese, R. (1996). *The Phonology of German* (Clarendon, Oxford).
- Yip, M. C. W. (2000). "Recognition of spoken words in continuous speech: Effects of transitional probability," in Proceedings of the Sixth International Conference on Spoken Language Processing, Beijing, China pp. 758–761.