THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2010, 63 (3), 544–554

Psychology Press
Taylor & Francis Group

Segmentation by lexical subtraction in Hungarian speakers of second-language English

Laurence White
Department of Experimental Psychology, University of Bristol, Bristol, UK

James F. Melhorn
School of Medicine, Cardiff University, Cardiff, UK

Sven L. Mattys
Department of Experimental Psychology, University of Bristol, Bristol, UK

Using cross-modal form priming, we compared the use of stress and lexicality in the segmentation of spoken English by native English speakers (L1) and by native Hungarian speakers of second-language English (L2). For both language groups, lexicality was found to be an effective segmentation cue. That is, spoken disyllabic word fragments were stronger primes in a subsequent visual word recognition task when preceded by meaningful words than when preceded by nonwords: For example, the first two syllables of *corridor* were a more effective prime for visually presented *corridor* when heard in the phrase *anythingcorri* than in *imoshingcorri*. The stress pattern of the prime (strong–weak vs. weak–strong) did not affect the degree of priming. For L1 speakers, this supports previous findings about the preferential use of high-level segmentation strategies in clear speech. For L2 speakers, the lexical strategy was employed regardless of L2 proficiency level and instead of exploiting the consistent stress pattern of their native language. This is clear evidence for the primacy and robustness of segmentation by lexical subtraction even in individuals whose lexical knowledge is limited.

Keywords: Speech segmentation; Lexical information; Second language; Prosody; Cross-modal priming.

There are multiple potential cues available to listeners for the detection of boundaries between spoken words. Mattys, White, and Melhorn (2005) suggest that the relative weighting of segmentation cues depends on the interpretive conditions in which speech is heard. In clear, contextualized speech, English listeners rely primarily on high-level information such as lexical identity and preceding syntactic/semantic structure, as available. In poorer interpretive conditions or where lexical information is ambiguous or unhelpful, Mattys et al. found that sublexical

Correspondence should be addressed to Laurence White, Department of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, UK. E-mail: laurence.white@bristol.ac.uk

This research was supported by a British Academy grant to Laurence White (SG-46958). Thanks are due to Anna Babarczy, Eva Deliaga, and staff and students of the Department of Cognitive Science at Budapest University of Technology and Economics for assistance with the experiment and to Maria Gósy and Katalin Mády for advice on Hungarian. We also thank Kathy Rastle and Arthur Samuel for helpful comments on earlier drafts of this paper.

http://www.psypress.com/qjep
DOI:10.1080/17470210903006971

cues (e.g., acoustic–segmental cues such as phonotactic probabilities, degree of coarticulation, and allophonic variation) become relatively more important for segmentation. The use of word-initial stress as a segmentation cue appears to be a fall-back strategy, in the absence of other cues. Thus, the overall picture for English is of a hierarchy of cues with decreasing weights, from lexical/contextual to acoustic/segmental to word stress.

However, native speakers of other languages may employ different sublexical segmentation strategies, according to language-specific phonotactic, allophonic, and metrical regularities. This potentially poses problems for second-language (L2) learners, who may persevere with native-language (L1)-appropriate cues, such as predominant stress patterns (Sanders, Neville, & Woldorff, 2002) and phonotactic regularities (Weber, 2000). As L2 competence develops, segmentation strategies may adapt to the L2 input, either by adoption of L2-appropriate strategies or simply by suppression of L1-appropriate strategies (Cutler, Mehler, Norris, & Segui, 1992).

There is, however, no a priori language-specific restriction on the use of lexical, syntactic, and semantic information for segmentation. L1 speakers of all languages presumably use such information, where available, in preference to potentially unreliable sublexical cues. Importantly, such a strategy is also applicable in L2 listening, once the learner has sufficient lexical, syntactic, and semantic knowledge. Thus, where clear lexical cues to word boundaries are available, even L2 speakers of relatively moderate proficiency may use them in preference to L1-appropriate sublexical strategies. We consider this here by testing the use of two segmentation strategies by L1 and L2 speakers of English: "segmentation by lexical subtraction" and metrical segmentation, the former based on word knowledge and the latter on stress distributional cues.

Segmentation by lexical subtraction

By "segmentation by lexical subtraction", we refer to the use of lexical knowledge to impose a segmentation structure on the speech input. Thus, boundaries of a recognized word allow a listener to hypothesize that the immediately preceding or following phonetic material ought to be the end or the beginning of new words. For example, assuming knowledge of the word *enormous*, hearing *enormousolister* should lead the listener to infer that *olister* might be a new word, or at least that /ol/ is the start of a new word. Consistent with this logic, Mattys et al. (2005) found that priming of a visual word by a spoken two-syllable fragment of that word was greater when the fragment was preceded by a real word than when it was preceded by a nonsense word.

This lexically derived approach to segmentation appears to be learnt very early in life. Six-month-old infants are more likely to extract new words from connected speech when these words are adjacent to known words, such as *Mommy*, than when they are adjacent to unfamiliar words (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005). Adults faced with an artificial language likewise use knowledge of words they have already learnt in that language to extract adjacent words (Dahan & Brent, 1999).

Given that developmental and artificial-language studies suggest a tendency for the adoption of such as strategy at a fairly minimal stage of vocabulary development, the same approach can potentially be taken in L2 segmentation, assuming sufficient lexical knowledge exists. However, it is possible that L2 speakers of limited proficiency may prefer to rely on the low-level acoustic–segmental or stress-based cues typical to their first language.

Metrical segmentation

Behavioural studies indicate that English speakers use word-initial stress for segmentation primarily where other cues are less accessible, such as in faint speech (Cutler & Butterfield, 1992), in dysarthric speech (Liss, Spitzer, Caviness, Adler, & Edwards, 1998), and in noise (Mattys, 2004; Smith, Cutler, Butterfield, & Nimmo-Smith, 1989). The avoidance of metrical segmentation where alternative cues are available reflects the fact that word-initial stress is a matter of statistical predominance in English. Only 85–90% of English content words begin with a stressed syllable (Cutler & Carter, 1987), and many polysyllables contain two or more stresses, so a strict metrical strategy would lead to frequent errors.

Hungarian, in contrast, has wholly consistent word-initial stress placement: All content words, even loan words, must begin with a stressed syllable (e.g., Siptár & Törkenczy, 2000). Stressed vowels and consonants are longer than unstressed vowels, but the durational difference is attenuated compared to English (e.g., Fónagy, 1958; Mády, Bombien, & Reichel, 2008), and, unlike in English, unstressed Hungarian vowels are not reduced. Although stress-based segmentation has not been investigated in Hungarian, electrophysiological evidence shows that Hungarian speakers are highly sensitive to violations of their regular stress-initial prosodic pattern (Honbolygó, Csépe, & Ragó, 2004). Furthermore, in Finnish, another language with fixed word-initial stress, rhythm is used as a segmentation cue and dominates vowel harmony when the two cues are pitted against one another (Vroomen, Tuomainen, & de Gelder, 1998). This suggests that stress has higher cue weighting in Finnish than in English, where it is only used in the absence of alternative lexical or segmental cues.

Although the use of stress to segment Hungarian has never been tested directly, it is reasonable to predict that—given the findings regarding Finnish segmentation and the similarity of stress distribution in Hungarian and Finnish—the consistent stress-initial pattern in Hungarian will likewise lend a relatively high weighting to metrical segmentation in Hungarian, potentially even competing with lexical information.

Experimental purpose

Second-language learners are comparable to infants in lacking lexical knowledge with which to interpret the speech stream. They differ, however, in having a repertoire of sublexical cues to word boundaries available from their L1. Thus, one possible strategy for L2 speakers is to exploit L1-appropriate sublexical cues when any are found in the L2 input. An alternative is to exploit their limited L2 lexical knowledge almost from the outset of L2 acquisition. Here, adopting a cross-modal form priming paradigm previously used in studies of L1 segmentation (e.g., Mattys et al., 2005), we examined the use of lexical knowledge and stress placement for segmentation of English by native Hungarian speakers of L2 English. As the strategy exploited may alter with L2 experience, we grouped Hungarian participants according to an independent measure of English proficiency. Attempting to replicate previous findings, we also used the same experimental design with a group of native English speakers.

Method

Participants

We tested 77 L1 speakers of British English at the University of Bristol and 104 Hungarian speakers of L2 English at Budapest University of Technology and Economics. The native Hungarian speakers were self-reported as speaking L2 English well, but not having learnt English or any other foreign language before the age of 10. Formal assessment of the English level of Hungarian participants was carried out using the Dialang online diagnostic language testing system (see below). None of the participants reported speech or hearing difficulties. They all received course credit or a small honorarium for their participation.

Materials and experimental design

We used a cross-modal form priming paradigm. On each trial, participants heard a five-syllable phrase (e.g., *anythingcorri*) over their headphones, with visual presentation of a three-syllable letter string (the "target", e.g., *corridor*) during or after the auditory phrase. On experimental trials, the first three syllables of the auditorily-presented phrase were referred to as the "context" (e.g., *anything*) and the final two syllables as the "prime" (e.g., *corri*)—that is, the prime comprised the first two syllables of the three-syllable visual target. On these experimental trials, the visual target appeared 100 ms after the offset of the auditory phrase. On filler trials, the position of the two-syllable prime within the five-syllable auditory phrase was varied, as was the temporal synchronization between the auditory phrase and the visual target, as detailed below.

To ensure that all materials would be processed reasonably quickly and accurately by all participants, we ran a pilot lexical-decision test. For the potential primes, we recorded 110 trisyllabic words (55 stress-initial, 55 stress-medial) matched between stress conditions on initial segment, uniqueness point, and familiarity rating, using statistics from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995). For the auditory context preceding the prime, we recorded another 110 trisyllabic words (55 stress-initial, 55 stress-medial) and 110 nonwords, matched pairwise with the words on stress pattern, average diphone frequency, and final syllable rhyme. There were a further 110 nonword fillers to balance the number of words and nonwords. The stimuli were read in a sound-attenuated room by a male native speaker of standard southern British English (the second author) and were recorded directly to disk at 32 kHz. All words were presented in random order to 36 native English speakers and 34 non-native speakers from a variety of L1 backgrounds, all students at the University of Bristol. Participants used the left and right shift keys of a laptop (labelled "yes" and "no") to indicate whether each token was an English word or not. Both speed and accuracy were emphasized.

Based on the pilot lexical-decision results, we excluded words and nonwords with less than 90% accuracy for both natives and non-natives and with latencies more than 1.5 standard deviations above the stimulus group mean. From the remaining stimuli, we selected 16 sets of contexts and primes, as shown in Table 1. Word and nonword contexts were matched overall on average diphone frequency and pairwise on stress pattern and final-syllable rhyme. The word targets, of which the strong–weak (SW) and weak–strong (WS) primes comprised the first two syllables, were matched, as in the pilot experiment, on initial segment, uniqueness point, and familiarity rating.

For each of the 16 sets of contexts and primes, we constructed eight possible test phrases: each of the four contexts (word/nonword by stress-initial/stress-medial) followed by each of the two primes. An example set of eight possible test phrases is

Table 1. *Experimental materials, arranged in the 16 context + prime sets*

Word contexts		*Nonword contexts*		*Primes*	
SW	*WS*	*SW*	*WS*	*SW*	*WS*
anything	another	imoshing	imather	corridor	confusion
holiday	excitement	faluday	adjitement	critical	creative
injury	expensive	amjury	lastensive	difficult	decision
yesterday	umbrella	mosterday	elbrella	family	familiar
tropical	abnormal	broginal	perdomal	motivate	memorial
elegant	department	anigant	clipartment	mystery	mistaken
animal	already	erromal	elgeady	national	November
elephant	discussion	aniphant	sartussion	penalty	performance
library	important	lidrury	envortant	popular	potato
citizen	December	subizen	gosember	principle	proportion
argument	apartment	urdoment	eldartment	recipe	remember
photograph	opinion	saltograph	attrenion	register	republic
memory	impression	narmory	antression	similar	suggestion
character	condition	jelicter	temdition	surgery	suspicion
capital	arrival	dokital	fellival	temperature	together
interview	intention	entomview	amsention	typical	tradition

Note: Only the first two syllables of the prime were included in the test phrases. The stress pattern of the context was used only as a counterbalancing factor. S = strong. W = weak.

illustrated in Table 2. In order to reduce the number of exposures to each prime, we selected four test phrases from each set of eight, as shown, with context and prime stress patterns, and context lexicality systematically counterbalanced between sets. All participants heard the same 64 experimental utterances (16 sets × 4 utterances per set).

The same speaker and recording conditions were used to construct the experimental stimuli as those used for the pilot study. During the recording, the speaker pronounced each full phrase without interruption (e.g., *anythingcorridor*). The last syllable of the prime was then edited out (e.g., *anythingcorri*, with *anything* being the word context and *corri* being the SW prime). Cross-splicing was then carried out to minimize acoustic idiosyncrasies across conditions: Thus, the beginning of the word context and the nonword context of a set (e.g., *anyth* and *imosh*) was spliced to a common remainder (e.g., *ingcorri*), yielding test phrases acoustically identical from a few segments prior to the end of the context onwards, but differing in the lexicality of the context (e.g., *anythingcorri* and *imoshingcorri*). The common remainder was taken equally often from the word and the nonword context phrases.

As in Mattys et al. (2005), a baseline phrase was matched to each test phrase by replacing the prime (i.e., the last two syllables of the test phrase) with distorted speech of matched duration. The distorted speech, created by superimposing several unrelated disyllables, sounded like babble noise with no identifiable segmental or suprasegmental characteristics.

In addition to the 64 experimental and 64 baseline trials, there were three different types of filler trials. There were 128 fillers in which nonword visual targets appeared, as with the word targets in the experimental trials, 100 ms after the offset of the auditory phrase. There were 80 fillers in which the presentation of the visual target (word or nonword) was aligned with the end of the second syllable of the auditory test phrase and 80 in which the presentation of the visual target (word or nonword) was aligned with the end of the third syllable, thus varying the point at which visual targets were expected. All three types of filler trials were divided between those in which the two auditory syllables immediately preceding the visual target (word or nonword) were congruent with the target and those in which they were not. Nonwords were trisyllabic orthographically and phonotactically legal letter strings matched with word targets on their average number of letters and expected stress pattern.

Procedure

Participants were tested using a laptop PC in a quiet room, with the experimental utterances played over good-quality headphones and the visual targets presented in the centre of the laptop screen in 24-point lower-case font. For each experimental and baseline trial (and for an equal number of fillers, as outlined above), the visual target appeared 100 ms after the offset of the

Table 2. *Example test utterances*

Stress pattern		*Context*	
Context	*Prime*	*Word*	*Nonword*
SWW	SW	anything<u>corri</u>	**imoshing<u>corri</u>**
	WS	**anything<u>confu</u>**	imoshing<u>confu</u>
WSW	SW	**another<u>corri</u>**	imather<u>corri</u>
	WS	another<u>confu</u>	**imather<u>confu</u>**

Note: In this set, the prime (underlined) originated from *corridor* (SW) or *confusion* (WS). The four utterances in bold were those selected as experimental stimuli for this particular set: Half the experimental sets had the alternative combination of context lexicality, context stress pattern, and prime stress pattern. S = strong. W = weak.

utterance, as Mattys (2004) showed maximum phonological priming for that delay compared to earlier alignments. Participants had 3,000 ms from the onset of target presentation to give their response. Following the response or at the end of the response window, there was a 1,500-ms pause before the onset of the next utterance.

Participants were told that, on each trial, an utterance would be played over the headphones and that, during or immediately following the utterance, a letter string would be displayed on the laptop screen. As in the pilot, participants were instructed to decide whether the letter string was an English word or not, responding with the two shift keys of a laptop. Both speed and accuracy were emphasized.

Trials were pseudorandomized so that repeated primes or targets were separated by at least 20 trials. There were no more than three visual words or non-words in a row and no more than five SW or WS auditory primes in a row, including fillers. Each participant received a different random order.

Language testing

We used the Dialang online diagnostic language testing system (software developed with the support of the European Commission and available from www.dialang.org) to obtain an independent assessment of Hungarian speakers' competence in English, based upon their scores in reading and listening tests. All Hungarian participants carried out the tests immediately following the experiment. They first took part in a written lexical decision placement test, in which they were presented with a series of 75 verbs and verb-like nonwords on the computer screen and had to decide whether each string was a word or not. The score from this placement test was used to select the appropriate level of the subsequent listening test, in which participants heard 30 short recordings of English speech and answered a multiple-choice comprehension question about each recording. Based on the outcome of the placement test and the listening comprehension, participants were assessed as being at one of six levels of English competence, from A1 (beginner) to C2 (very advanced). The mean placement scores (and number of participants) were: A1 = 155 (6); A2 = 198 (7); B1 = 557 (28); B2 = 708 (37); C1 = 847 (15); C2 = 940 (9).

Results

One English L1 speaker and 2 Hungarian L2 speakers of English who had lexical-decision accuracy rates below 70% across all trials were omitted from the analyses, leaving 76 English L1 speakers and 102 Hungarian L2 speakers of English. Lexical-decision latencies were measured from the onset of visual target presentation. Incorrect responses to word targets and correct responses two standard deviations from the mean latency (computed separately for each participant's target and baseline trials) were discarded (6% discarded for English L1 speakers; 7.5% discarded for Hungarian L2 speakers of English).

Analysis 1. L1 versus L2 English speakers

Mean lexical-decision latencies and accuracy levels are reported in Table 3. Priming, which was calculated as the difference between the latencies for the baseline and test conditions, is also shown in Table 3. Statistical analyses were based on mixed-effect modelling of the unaveraged response latency data (Baayen, Davidson, & Bates, 2008), with participants and items as random factors and condition (baseline, test), lexicality (word vs. nonword), stress (SW vs. WS), and English background (L1, L2) as fixed factors.

There was a main effect of English background, $F(1, 21300) = 1{,}300$, $p < .001$, indicating that L1 speakers, unsurprisingly, had overall faster lexical decision latencies (L1: 533 ms; L2: 587 ms). There was also a main effect of condition, $F(1, 21300) = 1{,}905.71$, $p < .001$, which demonstrates that responses were consistently faster on test trials than on baseline trials (see Table 3). However, because we are primarily interested in the effects of lexicality, stress, and language background on priming (i.e., the difference between baseline and test conditions), the exploration of the experimental hypotheses focuses on reliable interactions

Table 3. *Average lexical-decision latencies, accuracy, and levels of priming for L1 versus L2 English speakers*

	English L1 speakers		*English L2 speakers*		
Prime	*SW*	*WS*	*SW*	*WS*	*Overall means*
Word context	506 (98.2)	494 (98.7)	546 (99.0)	545 (98.8)	523 (98.7)
Baseline	563 (97.0)	560 (98.0)	629 (99.0)	624 (98.5)	594 (98.1)
Mean priming	57	66	83	79	71
Nonword context	516 (98.0)	507 (98.7)	556 (97.0)	555 (97.0)	534 (97.7)
Baseline	556 (98.2)	560 (98.4)	623 (97.5)	625 (97.9)	591 (98.0)
Mean priming	40	53	67	70	57
Overall means	535 (97.9)	530 (98.5)	589 (98.1)	587 (98.0)	560 (98.1)

Note: Mean lexical-decision latencies in ms; % accuracy in parentheses; levels of priming = baseline − test. In all cases, the priming effect was significant (i.e., latencies to the experimental trials were significantly faster than those to the baseline trials) at $p \leq .001$. S = strong. W = weak. L1 = native language. L2 = second language.

between condition and the other factors (lexicality, stress, language background).

An interaction between condition and lexicality, $F(1, 21300) = 21.18$, $p < .001$, suggested that priming was greater following word than following nonword contexts, but this did not interact with English background, $F(1, 21300) < 1$ (L1: 16 ms; L2: 13 ms; see Figure 1). As indicated by the interaction between condition and English background, $F(1, 21300) = 48.99$, $p < .001$, the L2 speakers showed greater priming than the L1 speakers (see Table 3 and Figure 1). Although this effect was somewhat unexpected, it could simply reflect an inflation of any differences at the slower latencies of the L2 speakers (see main effect of language background above). It could also suggest that the L2 speakers took greater advantage of auditory information when performing the visual lexical-decision task, actively using this source of information as a supplementary strategy in the face of non-native input.

There was no significant interaction between condition and stress, suggesting that priming was comparable for SW and WS primes, $F(1, 21300) = 2.63$, $p = .10$. In numerical terms, the priming difference was in fact in the opposite direction to that expected according to a metrical segmentation hypothesis (SW primes: 62 ms; WS primes: 67 ms). None of the other interactions with condition reached significance.

Lexical-decision accuracy was similar for the two groups: L1 = 93.7%; L2 = 93.5%, $F(1, 22782) < 1$. Given the range of L2 proficiencies, this result probably reflects the piloting that ensured that materials were well recognized by L2 speakers and the fact that L2 speakers took slightly longer to perform the task.

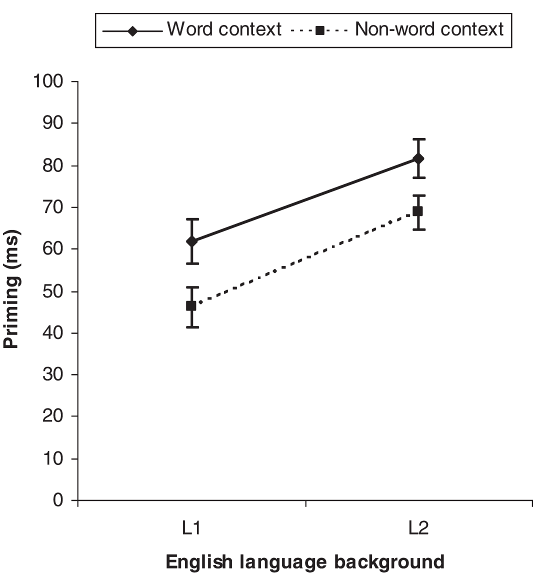

Figure 1. *Priming (ms) for all participants as a function of the lexical context preceding the prime (word vs. nonword) and participants' familiarity with English (native language, L1, vs. second language, L2). Error bars represent ± one standard error.*

Analysis 2. L2 speakers according to English competence

The above analysis makes clear that both native and non-native speakers use lexical knowledge as a cue to the placement of a subsequent word boundary. To explore the impact of L2 competence on the use of this segmentation strategy, we divided the L2 participants into groups according to their linguistic competence in English, as assessed by Dialang. Because there was a concentration of participants at proficiency levels B1 and B2, we merged the two lower levels, A1 and A2, to form the level A, and likewise merged the two upper levels, C1 and C2, to form the level C. Mean lexical-decision latencies, accuracy levels, and priming effects (difference between baseline latency and test latency) for the four groups are reported in Table 4.

Mixed-effect modelling with participants and items as random factors and condition (baseline, test), lexicality (word vs. nonword), stress (SW vs. WS), and L2 level (A, B1, B2, C) as fixed factors revealed an interaction between condition and lexicality, $F(1, 12172) = 9.78$, $p < .005$, which showed, as before, greater priming following a lexical than a nonlexical context. This lexical influence on priming was comparable across the four L2 levels, $F(1, 12172) < 1$ (see Figure 2): A = 12 ms; B1 = 10 ms; B2 = 15 ms; C = 12 ms. None of the other interactions with condition reached significance.

As a further check, we examined the relationship between participants' placement test scores and the size of lexical effect (the difference in mean priming following word context and nonword context). As shown in Figure 3, there was no significant correlation between these variables: Spearman's $rho = .085$, $p > .10$, supporting the finding suggested by the above analysis of variance (ANOVA) that there was no increase in lexical priming for more competent L2 speakers.

Accuracy was found to differ across the four L2 levels, $F(1, 13052) = 5.47$, $p < .001$, which pairwise comparisons showed was due to the lower accuracy rate of the A group, 95.3%, than of all other groups (B1 = 97.8%; B2 = 98.6%; C = 99.2%), none of which differed significantly.

Table 4. *Average lexical-decision latencies, accuracy, and levels of priming for English L2 speakers according to English competence*

	English L2 speakers by assessment level								
	A		*B1*		*B2*		*C*		
Prime	*SW*	*WS*	*SW*	*WS*	*SW*	*WS*	*SW*	*WS*	*Overall means*
Word context	637	626	526	530	534	527	539	545	558
	(97.6)	(96.6)	(98.7)	(99.1)	(99.3)	(98.8)	(99.5)	(99.7)	(98.7)
Baseline	726	696	613	619	612	610	623	616	639
	(92.0)	(94.0)	(96.0)	(97.0)	(98.0)	(98.0)	(99.0)	(98.0)	(96.5)
Mean priming	89	70	87	89	78	83	84	71	81
Nonword context	651	644	536	537	547	537	542	553	568
	(95.0)	(96.2)	(100)	(98.4)	(99.0)	(98.8)	(99.0)	(99.5)	(98.2)
Baseline	713	716	612	616	606	609	615	612	637
	(94.2)	(96.2)	(95.8)	(97.4)	(98.6)	(98.0)	(99.5)	(98.7)	(97.3)
Mean priming	62	72	76	79	59	72	73	59	69
Overall means	682	671	572	576	575	571	580	582	593
	(94.7)	(95.8)	(97.6)	(98.0)	(98.7)	(98.4)	(99.3)	(99.0)	(97.7)

Note: Mean lexical-decision latencies in ms; % accuracy in parentheses; levels of priming = baseline − test. In all cases, the priming effect was significant (i.e., latencies to the experimental trials were significantly faster than those to the baseline trials) at $p \leq .001$. S = strong. W = weak. L1 = native language. L2 = second language.

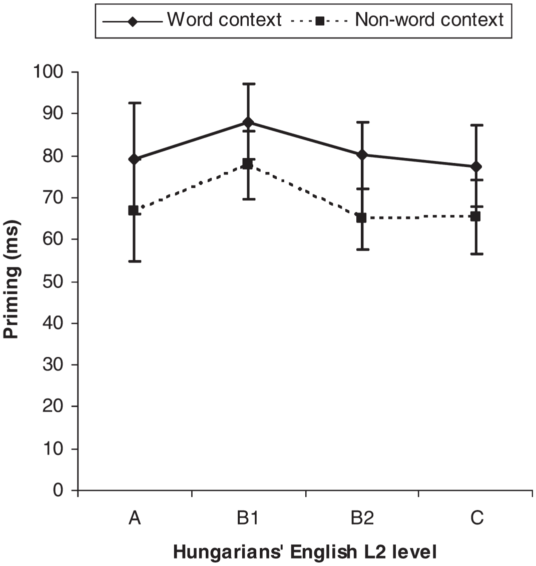

Figure 2. *Priming (ms) for all Hungarian participants as a function of the lexical context preceding the prime (word vs. nonword) and the Hungarian participants' levels of English proficiency, from A (low) to C (high). Error bars represent ± one standard error.*

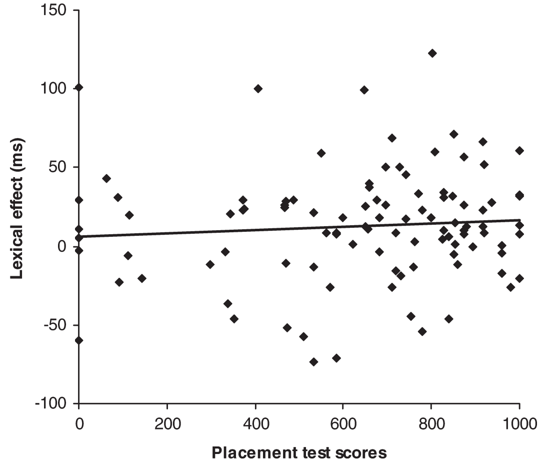

Figure 3. *Correlation between Hungarian participants' placement test scores and the size of their lexical segmentation effect (i.e., the difference in mean priming following word context and nonword context).*

Discussion

The fact that English L1 listeners use lexical knowledge for segmentation of intelligible speech and ignore stress cues is a clear replication of our previous findings (Mattys et al., 2005). What is remarkable is that all L2 speakers also used lexicality as a segmentation cue for English, regardless of proficiency level. That is, the recognition of a word in the auditory stream allowed them to impose a segmentation structure on that stream, even if their knowledge of English was relatively minimal. This segmentation by lexical subtraction strategy, therefore, does not require a substantial native-like lexicon to operate, in line with findings for six-month-old infants (Bortfeld et al., 2005). This is also consistent with evidence that lexical activation is an automatic process in both monolinguals (e.g., Onifer & Swinney, 1981) and bilinguals (Hamers & Lambert, 1972).

Despite the dominance of segmentation by lexical subtraction even among our least proficient non-native speakers, this strategy can obviously not be initiated in the absence of any L2 experience at all, and we assume that at some elementary level of proficiency, the lexical effect would disappear. In our L2 group, although there was no interaction between L2 level and lexicality, it is the case that the A group was overall numerically slower than the other three groups (mean latencies: A = 677 ms; B1 = 574 ms; B2 = 573 ms; C = 581 ms), as well as less accurate (accuracy rates, as given above: A = 95.3%; B1 = 97.8%; B2 = 98.6%; C = 99.2%). Thus, the priming effects found for the A-group participants may be somewhat inflated due to their overall slowness, allowing more time for priming effects to appear.

To explore this possibility, we removed the 3 A-group L2 participants with the longest overall latencies (mean: 1,003 ms) in order to bring the mean response latency of the A group down to a level comparable with that of the other L2 groups (mean of new 10-participant A-group: 578 ms). The lexical effect for these 10 participants almost completely disappeared (2 ms). It should also be noted that the 3 slowest A-group participants were overall more accurate than the others: 98.4% compared with 94.2%. Thus, at the lowest level of L2 competence, there may be a split in terms of response strategy: Some participants take substantially longer to respond, are more accurate, and—like the higher competence groups—show greater priming following word contexts than following nonword contexts. Elimination of these careful responders left little sign of a lexical effect in the A group of L2 participants, suggesting a limit to the utilization of the segmentation by lexical subtraction strategy not evident—due to the heterogeneity of the A group—in the overall analysis.

As far as stress is concerned, the literature on language development indicates that very young learners of English utilize a metrical segmentation strategy initially (e.g., Jusczyk, Houston, & Newsome, 1999), while, at the same time, exploiting lexical knowledge as soon as it is available (Bortfeld et al., 2005). Stress cues were, however, not exploited by Hungarian listeners, even those with a low English proficiency level. This is despite the complete consistency of word-initial stress placement in Hungarian, which should make it a wholly reliable cue to a preceding word boundary. Thus, as for English L1 perception, the weighting of lexical knowledge cues for L2 segmentation appears to be greater than that of sublexical cues, at least with regard to word stress. This is in line with the hierarchical approach to speech segmentation proposed by Mattys et al. (2005).

However, it could be argued that Hungarians failed to exploit stress because the acoustic cues to stress differ somewhat between English and Hungarian, in particular with regard to the marked vowel reduction in unstressed syllables in English. Likewise, our suggestion that Hungarian speakers may use stress for L1 segmentation is based on logical inferences from a wealth of studies carried out on other languages, mostly English, but also fixed stress-initial Finnish; it has not been tested directly. While there is no reason to believe that Hungarian speakers would simply ignore a cue with so much probabilistic potential, it is also important to remember that the reliability of a particular cue ought to be considered relative to that of others. While there are numerous studies on the use of stress relative to phonotactic cues, acoustic–phonetic cues, and lexical knowledge in English, there are virtually none in Hungarian. Although unlikely, reliance on stress for segmenting Hungarian might be mitigated by other highly reliable cues. A detailed investigation of the L1 segmentation strategies used by Hungarian studies is therefore needed. However, whatever native Hungarians' L1 segmentation strategies, the dominance of segmentation by lexical subtraction as an L2 strategy is clearly indicated in our present results.

Even the low-proficiency L2 speakers tested here had several years of intermittent exposure to English speech, which may have put them beyond the point at which metrical segmentation is used and certainly would have allowed accumulation of sufficient lexical knowledge for use in segmentation. Longitudinal studies of L2 learners could reveal whether native-appropriate segmentation strategies dominate in the very early stages. However, it must be stressed that the lowest proficiency English speakers here were well below a functional level of conversational competence, and yet some, though not all, automatically used known words as stepping stones in the speech stream. Thus, the concept of segmentation by lexical subtraction appears to be a powerful and possibly universal word-finding mechanism at the core of both language development and second-language acquisition.

Original manuscript received 10 December 2008
Accepted revision received 16 April 2009
First published online 22 July 2009

REFERENCES

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subject and items. *Journal of Memory and Language*, *59*, 390–412.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX Lexical Database (Release 2). Philadelphia, PA: Linguistic Data Consortium.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me. *Psychological Science*, *16*, 298–304.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, *31*, 218–236.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 133–142.

Cutler, A., Mehler, J., Norris, D. G., & Seguí, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, *24*, 384–410.

Dahan, D., & Brent, M. R. (1999). On the discovery of novel wordlike units from utterances: An artificial-language study with implications for native-language acquisition. *Journal of Experimental Psychology: General*, *128*, 165–185.

Fónagy, I. (1958). A hangsúlyról [On stress]. In P. Fábián (Ed.), *Nyelvtudományi Értekezések* (Vol. 18). Budapest, Hungary: Akadémiai Kiadó.

Hamers, J. F., & Lambert, W. E. (1972). Bilingual interdependencies in auditory perception. *Journal of Verbal Learning and Verbal Behavior*, *11*, 303–310.

Honbolygó, F., Csépe, V., & Ragó, A. (2004). Suprasegmental speech cues are automatically processed by the human brain: A mismatch negativity study. *Neuroscience Letters*, *363*, 84–88.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*, 159–207.

Liss, J. M., Spitzer, S., Caviness, J. N., Adler, C., & Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *Journal of the Acoustical Society of America*, *104*, 2457–2466.

Mády, K., Bombien, L., & Reichel, U. D. (2008). Is Hungarian losing the vowel quantity distinction? In R. Sock, S. Fuchs, & Y. Laprie (Eds.), *Proceedings of the 8th International Seminar on Speech Production, Strasbourg* (pp. 445–448). Strasbourg, France: INRIA.

Mattys, S. L. (2004). Stress versus coarticulation: Toward an integrated approach to explicit speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 397–408.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*, *134*, 477–500.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition*, *9*, 225–236.

Sanders, L. D., Neville, H. J., & Woldorff, M. G. (2002). Speech segmentation of native and non-native speakers: The use of lexical, syntactic and stress pattern cues. *Journal of Speech, Language and Hearing Research*, *45*, 519–530.

Siptár, P., & Törkenczy, M. (2000). *The phonology of Hungarian*. Oxford, UK: Oxford University Press.

Smith, M. R., Cutler, A., Butterfield, S., & Nimmo-Smith, I. (1989). The perception of rhythm and word boundaries in noise-masked speech. *Journal of Speech and Hearing Research*, *32*, 912–920.

Taalas, P. (1999). DIALANG: An Internet-based system for diagnostic language assessment. *Proceedings of WebNet World Conference on the WWW and Internet 1999* (p. 1766). Chesapeake, VA: AACE.

Vroomen, J., Tuomainen, J., & de Gelder, B. (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language*, *38*, 133–149.

Weber, A. (2000). The role of phonotactics in the segmentation of native and non-native continuous speech. In A. Cutler, J. McQueen, & R. Zondervan (Eds.), *Proceedings of the Workshop on Spoken Access Processes* (pp. 143–146). Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics.