Speech Segmentation by Native and Non-Native Speakers: The Use of Lexical, Syntactic, and Stress-Pattern Cues

Lisa D. Sanders*
Helen J. Neville
University of Oregon
Eugene

Marty G. Woldorff
Duke University
Chapel Hill, NC

Varying degrees of plasticity in different subsystems of language have been demonstrated by studies showing that some aspects of language are processed similarly by native speakers and late-learners whereas other aspects are processed differently by the two groups. The study of speech segmentation provides a means by which the ability to process different types of linguistic information can be measured within the same task, because lexical, syntactic, and stress-pattern information can all indicate where one word ends and the next begins in continuous speech. In this study, native Japanese and native Spanish late-learners of English (as well as near-monolingual Japanese and Spanish speakers) were asked to determine whether specific sounds fell at the beginning or in the middle of words in English sentences. Similar to native English speakers, late-learners employed lexical information to perform the segmentation task. However, non-native speakers did not use syntactic information to the same extent as native English speakers. Although both groups of late-learners of English used stress pattern as a segmentation cue, the extent to which this cue was relied upon depended on the stress-pattern characteristics of their native language. These findings support the hypothesis that learning a second language later in life has differential effects on subsystems within language.

KEY WORDS: bilingual, speech segmentation, lexical, syntax, stress pattern

Many studies have shown that differences in language experience (e.g., learning an L1 or L2, age of acquisition) affect how language is processed. However, past studies have typically focused on a single aspect of language. In contrast, this study investigates multiple aspects of language, measuring the influence of language experience on the use of lexico-semantic, syntactic, and prosodic information. Previous research has suggested that some subsystems of language remain flexible throughout life such that individuals learning an L1, an L2 early in life, or an L2 later in life achieve similar levels of processing. Other subsystems are greatly affected by differences in language experience. Measuring the effects of language experience on different subsystems of language in the same individuals performing the same task makes it possible to determine the relative plasticity of these subsystems.

* Currently affiliated with the University of Maryland, College Park

Linguistic Subsystems

The results of behavioral and electrophysiological studies suggest that non-native speakers process the words and meanings of a second language in much the same way as native speakers. For example, deaf native signers who learn English later in life can detect semantic anomalies in their L2, and the event-related potentials (ERPs) elicited by these anomalies are similar to those of native speakers (Neville, Mills, & Lawson, 1992). Similar results have been reported for native Chinese late-learners of English (Weber-Fox & Neville, 1996) and native Japanese late-learners of German (Hahne & Friederici, 2001).

In contrast, considerable evidence suggests that differences in language experience have marked effects on the ability to process syntactic information in a native-like manner. Age of acquisition, characteristics of languages known at the time of acquisition, and manner in which a language is learned all affect L2 syntactic processing (De Groot & Kroll, 1997; Johnson & Newport, 1989, 1991; McDonald, 2000; Scovel, 1988). In fact, differences in language experience that have little effect on semantic processing can significantly affect syntactic processing (Weber-Fox & Neville, 1996; Hahne & Friederici, 2001).

Phonological and prosodic processing may be particularly susceptible to differences in how and when a language is acquired. Many studies of pronunciation indicate that non-native speakers retain a foreign accent long after learning their L2 (e.g., Flege, Yeni-Komshian, & Liu, 1999; Guion, Flege, Liu, & Yeni-Komshian, 2000). Other studies have shown that acquiring lexical tone is very difficult for native speakers of nontonal languages (Leather, 1987) and that native speakers of tone or pitch-accent languages (including Japanese) do not use metrical stress as do native speakers of accentual languages (including Spanish and English) when processing English as a second language (Archibald, 1997).

Speech Segmentation

To process speech, listeners must break continuous streams of sound into potentially meaningful units. Lexico-semantic, syntactic, and prosodic information have all been shown to play a role in the segmentation process (e.g., Cutler & Butterfield, 1992; Cutler, Mehler, Norris, & Segui, 1983, 1986; Sanders & Neville, 2000). Because these different types of linguistic information can be used as segmentation cues, the investigation of speech segmentation offers an opportunity to index different subsystems of language using the same task.

No previous studies reported in the literature have measured the effects of differences in language experience on the use of lexico-semantic and syntactic segmentation cues. However, studies on the use of prosodic segmentation cues by bilingual speakers have been conducted. For example, syllable structure has been shown to be important for segmenting syllable-timed languages such as French, Italian, and Spanish (Cutler, Mehler, Norris, & Segui, 1983, 1986). Native English speakers who began learning French very early in life do not show syllable segmentation effects (Cutler, Mehler, Norris, & Segui, 1989, 1992). Similarly, a rhythmic cue that has been shown to be important for segmenting Japanese is not used by native English speakers listening to either English or Japanese (Cutler & Otake, 1994; Otake, Hatano, Cutler, & Mehler, 1993). These results have led to the proposal that the rhythmic characteristics of a language (e.g., syllable-timed, stress-timed) are important in determining how native speakers of that language will segment continuous speech in any language. Additionally, these results have been used to conclude that non-native speakers do not acquire rhythmic segmentation cues other than those used for their L1.

Hypotheses

In a previous study, native English speakers were shown to use lexico-semantic information to help them perform a segmentation task (Sanders & Neville, 2000). In the present study, the same segmentation task was given to late-learners of English. Because native and non-native speakers have been shown to detect and process semantic anomalies in similar manners, we hypothesized that non-native speakers would also be able to use lexico-semantic information to segment speech.

Native English speakers have also been shown to use syntactic information to segment speech (Sanders & Neville, 2000). However, previous research suggests that late-learners of a language may not process syntactic information in a native-like manner. Therefore, we hypothesized that the non-native speakers in the present study would not be able to use syntactic information as a segmentation cue to the same extent as native speakers.

English is a stress-timed language, with a typical pattern of strong stress on the first syllable of a word followed by weaker stress on the remaining syllables (Cutler & Carter, 1987). Therefore, listeners can assume that strongly stressed syllables are likely to be word initial and unstressed syllables are likely to be word medial. Indeed, stress pattern has been shown to be an important segmentation cue for native English speakers (Cutler & Butterfield, 1992; Sanders & Neville, 2000). The ability of non-native speakers to use stress pattern as a segmentation cue may be affected by the characteristics of L1 in addition to whether a language is learned as an L1 or an L2.

To test this hypothesis, we included groups with different L1 backgrounds (Japanese and Spanish) in this study. If non-native speakers fail to use rhythmic segmentation cues other than the rhythmic cue relevant for their L1, native speakers of Japanese (mora-timed) and Spanish (syllable-timed) would not be expected to use stress pattern as a segmentation cue when listening to English. Alternatively, native Japanese and native Spanish speakers might differ in their abilities to use stress pattern as a segmentation cue in English. Japanese does not use loudness and duration as indications of lexical stress (Hyman, 1977), whereas Spanish does use loudness and duration as indications of stress but has a typical stress pattern different from that for English. If learning a new segmentation cue, but not a new pattern, is difficult for second-language learners, then only Spanish speakers would be expected to use English stress pattern as a segmentation cue. If using a different stress pattern in a native language interferes with the ability to learn a new stress pattern, only Japanese speakers would be expected to use English stress pattern as a segmentation cue.

In addition, it was important to measure the performance of native Japanese and Spanish speakers who did not know English. By doing so, it was possible to determine if late-learners' use of English segmentation cues could best be explained by their having learned new cues or by their having transferred cues from their native language. If native Japanese and native Spanish late-learners of English acquire new segmentation cues, then monolingual Japanese and Spanish speakers should not be able to use those cues. However, if late-learners of English apply segmentation cues from their native languages rather than learning new ones, then non-English speakers should be able to use these cues as well.

Method

Participants

Four groups of participants were included in this experiment: native Japanese late-learners of English (JE), native Spanish late-learners of English (SE), near-monolingual Japanese speakers (J), and monolingual Spanish speakers (S). A description of each group is provided in Table 1. The performance of these four groups was compared to that of monolingual English speakers (E) previously reported in Sanders and Neville (2000).

Because English is regularly taught as part of the curriculum in Japanese schools, it was not possible to find adult monolingual Japanese speakers. Therefore, native Japanese speakers with little English experience were

Table 1. Participants.

Group	*N*	*M* age	Gender (# women)	Language experience
E	16	20;9	11	native English speakers (Sanders & Neville, 2000) less than two years of foreign language study never lived in a non-English-speaking country
JE	16	24;1	11	native Japanese speakers began learning English after the age of 12 (*M* = 12;7) moved to the U.S. after the age of 15 (*M* = 18;5) lived in the U.S. for an average of 5.8 years (minimum = 2 years) at time of study: students at University of Oregon who used English predominantly
SE	14	26;0	9	native Spanish speakers began learning English after the age of 12 (*M* = 17;4) moved to the U.S. after the age of 16 (*M* = 19;10) lived in the U.S. for an average of 6.3 years (minimum = 2 years) at time of study: university students in San Antonio who used English predominantly
J	16	22;4	10	native Japanese speakers began learning some English after the age of 12 (*M* = 12;6) never lived in an English-speaking country at time of study: university students at Tokyo University who used Japanese predominantly
S	9	28;1	3	native Spanish speakers never studied written or spoken English lived in the U.S. less than 2 years (*M* = 7 months) at time of study: had completed at least 1 year of university-level classes in Spanish and used Spanish predominantly

Note. Mean age is in years;months.

selected for this study. They reported that their study of English was primarily focused on written, rather than spoken English. None of these subjects had been taught by a native English speaker, had friends or family who were native English speakers, or had lived in an English-speaking country. However, all of the subjects reported having access to English movies, books, and songs.

Stimuli

The sentences used for this study are characterized in detail in Sanders and Neville (2000) and will be described briefly here. The 900 sentences were designed to vary the amount of lexical, syntactic, and stress-pattern information available to the listener. Starting with normal English sentences (*semantic* sentences, N = 300), lexical information was reduced by replacing all of the open-class words with nonwords to create sentences that had normal English syntax and prosody but little meaning (*syntactic* sentences, N = 300). Syntactic information was reduced by replacing all of the remaining words and morphemes in the syntactic sentences with nonwords to create sentences with normal English prosody but little meaning and little syntactic information (*acoustic* sentences, N = 300). Representative examples of each sentence type are given in Table 2. IPA transcriptions of these example sentences are shown in Appendix A.

Most of the sentences contained a task-relevant "target" phoneme or phoneme combination (21 different targets). Stress pattern was varied by including words that contained these targets in different positions and in syllables of different stress. The targets were selected such that they were, with equal probability, (a) the first sound in a word with a stressed first syllable (Strong-stress, Initial-position); (b) the first sound in a stressed syllable in the middle of a word (Strong-stress, Medial-position); (c) the first sound in a word that was unstressed on the first syllable (Weak-stress, Initial-position); or (d) the first sound in an unstressed syllable in the middle of a word (Weak-stress, Medial-position). Therefore, both strong-stress initial-position and weak-stress medial-position targets occurred in words with a normal English stress pattern (strong stress on the initial syllable, e.g., *bot*tles and *tim*ber). Both strong-stress medial-position and weak-stress initial-position targets occurred in words with a less common English stress pattern (unstressed on the initial syllable, e.g., to*bac*co and bal*loon*).

An equal number of each type of target word (N = 60) was used. Words containing targets were never among the first or last three words of the sentence. Sixty sentences of each type that did not contain targets also were included. Examples of words with and without target phonemes are shown in Table 2. All of the sentences and examples of the targets produced in isolation were recorded by the same female native English speaker.

As reported previously (Sanders & Neville, 2000), care was taken to determine that the different sentence types were matched on as many physical characteristics as possible. Specifically, sentences were matched on speech rate, duration, position of target, and pitch contour. Word-initial and word-medial syllables that contained targets were matched on loudness, duration, and fundamental frequency.

Table 2. Semantic, syntactic, and acoustic sentences.

Condition	Sentence	Example
SI	Semantic	In order to recycle *b*ottles you have to separate them.
	Syntactic	In order to lefatal *b*okkers you have to thagamate them.
	Acoustic	Ah ilgen di lefatal *b*okkerth ha maz di thagamate fon.
SM	Semantic	If the only thing in it were to*b*acco it wouldn't cause so much harm
	Syntactic	If the ilmy shord in it were do*b*atty it wouldn't gaff so much hilm.
	Acoustic	Os fa ilmy shord el ok hon do*b*atty ag hapsel gaff sha nes hilm.
WI	Semantic	The child stopped crying when a *b*alloon was given to her.
	Syntactic	The ferp trepped plawing when a *b*arreal was kaffen to her.
	Acoustic	Sa ferp trepp plawel ron i *b*arreal hof kaffem gi wem.
WM	Semantic	I saved money since lowgrade tim*b*er worked for this project.
	Syntactic	I cheft rono since miltrok del*b*er meld for this plassig.
	Acoustic	O cheft rono zalf miltrok del*b*er meld sith foch plassig.
TA	Semantic	Try looking under the afghan for the toy you lost.
	Syntactic	Qui medding under the ithdon for the kay you moft.
	Acoustic	Qui medden amkel fa ithdon sal cha kay wa moft.

Note: SI = Strong stress, Initial position; SM = Strong stress, Medial position; WI = Weak stress, Initial position; WM = Weak stress, Medial position; TA = Target Absent.

Procedure

JE were tested in two 1.5-hour sessions in a sound-attenuating room at the University of Oregon. J were tested in groups of three in single 2.5-hour sessions at Tokyo University. SE and S were tested alone in single 2.5-hour sessions in a sound-attenuating room at the University of Texas Health Sciences Center in San Antonio. JE and SE were given language questionnaires and task instructions in English. J were given questionnaires and instructions that had been translated from English into Japanese by a native Japanese speaker. S were given questionnaires and instructions translated into Spanish by a team of native English and native Spanish speakers.

All participants completed a brief language questionnaire and 60 practice trials before the 900 test trials. All sounds were presented binaurally over headphones at approximately 60 dB above normal hearing threshold. All visual information was presented on a computer monitor 55 inches away from the subject.

During each trial, participants first heard the sound of the target for that sentence presented in isolation and saw a letter or letters representing that target on the screen. Subjects were instructed to listen for the target sound in the sentence that followed 1100 ms later. Participants were asked to press one button if they heard a target at the beginning of a word or nonword, to press a different button if they heard the target in the middle of a word or nonword, and not to press any button if they did not hear a target. The next trial began 1500 ms after the end of a sentence regardless of subject response. The letter or letters that represented the target remained on the screen for the entire trial. The 900 sentences were presented in random order, with the exception that no two versions (*semantic*, *syntactic*, and *acoustic*) of the same sentence were presented with fewer than 80 other sentences in between. During the test trials, participants were offered a break after every 20 sentences.

Analyses

A possible concern in interpreting the results was that the phonemes presented in isolation may have sounded more similar to targets in word-initial or word-medial positions or in stressed or unstressed syllables. For that reason, only trials on which subjects indicated that they had detected the target phoneme (by pressing either of the buttons) were included in analyses. Thus, if any subject heard the sounds presented in isolation and in the sentences as different phonemes, they would have been expected not to respond (as they were asked to do when the target did not occur in the sentence). Localization accuracy was measured by dividing the number of trials on which subjects successfully detected a target and determined whether it was word-initial or word-medial by the number of trials on which subjects correctly detected the target.

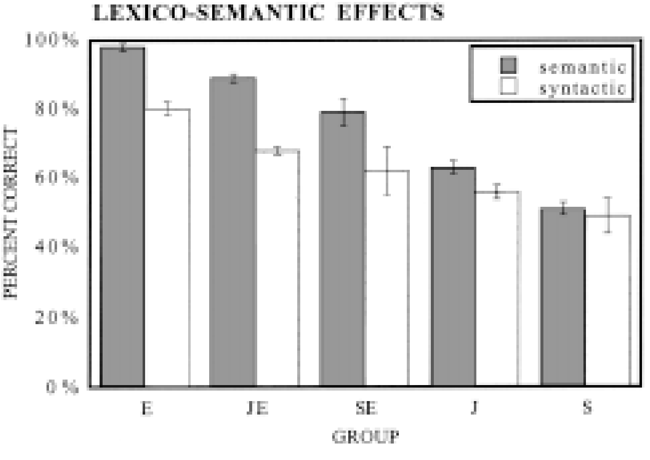
Mixed design 5 (group) × 3 (sentence type) × 2 (stress) × 2 (target position) ANOVAs were performed. Included were a group of native English speakers (E) from a previous study (Sanders & Neville, 2000). Because significant interactions ($p < .01$) for group and each of the other factors were found, each non-native group was entered into an ANOVA with E separately. Additionally, a 3 (sentence type) × 2 (stress) × 2 (target position) repeated-measures ANOVA was performed for each group alone. Planned comparisons between sentence types (*semantic* and *syntactic*, *syntactic* and *acoustic*) and stress patterns (normal and infrequent) were performed for each group separately and for each group with E. Bonferroni corrections for multiple comparisons were applied.

Results

Lexical Effects

Each group's accuracy on the *semantic* and *syntactic* sentences is shown in Figure 1. As reported previously (Sanders & Neville, 2000), E showed better performance for targets within *semantic* (M = 98%) than within *syntactic* (M = 80%) sentences. It was hypothesized that both JE and SE would also be able to use lexical information to perform the task. For JE, sentence type did affect performance [$F(2, 30) = 433$, $p < .01$], with higher accuracy on the *semantic* (M = 89%) than on the *syntactic* (M = 68%) sentences [$t(15) = 21.4$, $p < .01$; $d = 11.1$]. Similar effects of sentence type on accuracy were found for SE [$F(2, 13) = 22.41$, $p < .01$]. Again, performance was better for *semantic* (M = 79%) than for

Figure 1. Percent correct for the *semantic* and *syntactic* sentences.



syntactic (M = 62%) sentences [$t(13) = 4.61$, $p < .01$; d = 2.6]. When JE and E or SE and E were included in an ANOVA with two levels of sentence type (*semantic* and *syntactic*), there were no significant interactions between group and sentence type. This indicates that the increases in accuracy associated with the presence of lexical information did not differ for native and late-learners of English.

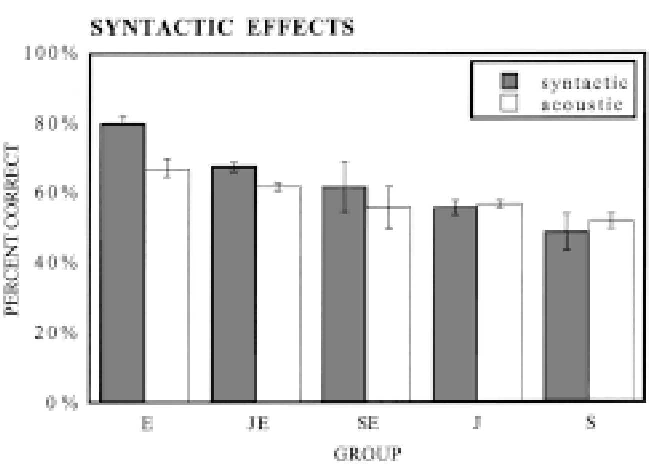
Because neither J nor S reported being able to understand conversational English, they were not predicted to show a benefit of lexical information on performance. This prediction held true for S, who showed no effect of sentence type on accuracy. However, J did show a sentence-type effect on localization accuracy [$F(2, 30)$ = 7.08, $p < .01$] such that performance was better for the *semantic* (M = 63%) than the *syntactic* (M = 56%) sentences [$t(15) = 3.67$, $p < .01$; d = 1.9]. As can be seen in Figure 1, the effect of lexical information on performance was not as large for this group as for E, JE, or SE. The difference in the size of this effect for E and J was indicated by a group-by-sentence interaction [$F(1, 30)$ = 23.41, $p < .01$] when only *semantic* and *syntactic* sentences were included in an ANOVA.

Syntactic Effects

As shown in Figure 2, E performed more accurately with the *syntactic* (M = 80%) than with the *acoustic* (M = 67%) sentences. Neither group of late-learners was hypothesized to use syntactic information to the same extent as native speakers. However, JE were actually slightly more accurate with targets in *syntactic* (M = 68%) than in *acoustic* (M = 62%) sentences [$t(15) = 5.81$, $p < .01$; d = 3.0]. When JE and E were entered into an ANOVA with only *syntactic* and *acoustic* sentence conditions, a group-by-sentence-type interaction [$F(1, 30)$ = 32.06, $p < .01$] reflected the finding that the difference for the two sentence types was larger for E than for JE. For SE, performance on the *syntactic* sentences (M = 62%) was not significantly better than performance on the *acoustic* sentences (M = 56%).

J and S were not expected to be able to take advantage of available syntactic information to perform the task. As predicted, there were no differences in performance on the *syntactic* and *acoustic* sentences for these groups.

Figure 2. Percent correct for the *syntactic* and *acoustic* sentences.



Stress-Pattern Effects

For E, there was a stress-by-position interaction on phoneme localization. When the data were grouped to compare the normal English stress pattern (strong-initial and weak-medial) to an infrequent English stress pattern (weak-initial and strong-medial), it was found that this group was more accurate with the normal pattern (M norm = 87%, M infreq = 76%), as shown in Figure 3. Performance on each condition is provided separately for each group in Appendix B.

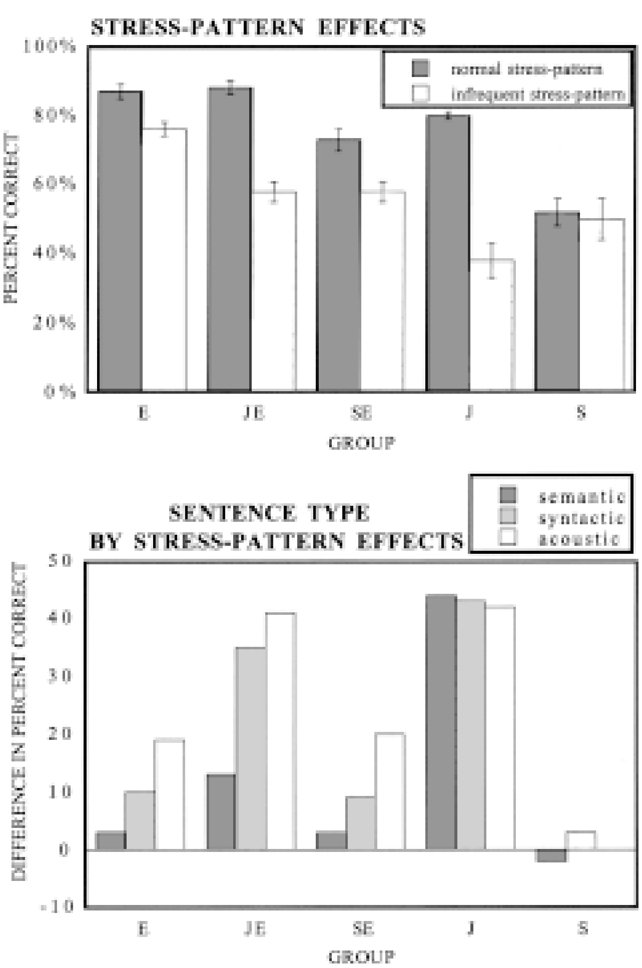
JE also showed a stress-by-position interaction on accuracy [$F(1, 15)$ = 308, $p < .01$] such that performance was better for the normal English stress pattern (M = 88%) than for the infrequent English stress pattern (M = 58%) [$F(1, 15)$ = 16.4, $p < .01$; d = 2.1]. Moreover, when E and JE were compared in the same ANOVA, a group-by-stress-pattern interaction [$F(1, 30)$ = 42.10, $p < .01$] indicated that JE showed a *larger* stress-pattern effect than E.

Because JE and SE had different types of experience with lexical stress before they began learning English, it was predicted that the effects of stress pattern in the present task would differ across these groups. However, like E and JE, SE evidenced a stress-by-position interaction on performance [$F(1, 13)$ = 17.33, $p < .01$]. Accuracy was higher for the normal English stress pattern (M = 73%) than for the infrequent English stress pattern (M = 58%) [$F(1, 13)$ = 17.04, $p < .01$; d = 2.3]. However, there was no group-by-stress-pattern interaction for E and SE.

Some of the stimuli used would receive the same lexical stress if the normal English pattern or the normal Spanish pattern were applied, but this was not true of other items. To determine if the stress-pattern effects found for SE could best be explained by the application of typical Spanish stress patterns or by the acquisition and application of typical English stress patterns, the stimuli were divided into two sets. The first set included items for which the typical stress patterns of the two languages predicted the same lexical stress (N = 38). The second set included items for which the typical stress patterns of the two languages predicted different lexical stress ($N = 202$). SE showed better performance with normal ($M = 70\%$) than infrequent ($M = 57\%$) stress for the subset of items for which both languages predicted the same lexical stress [$F(1, 13) = 8.77$, $p < .01$; $d = 1.6$]. They also showed better performance with normal English stress pattern ($M = 75\%$) than infrequent English stress pattern ($M = 58\%$) for the subset of items for which the two languages predicted different lexical stress [$F(1, 13) = 12.18$, $p < .01$; $d = 1.9$], consistent with the hypothesis that SE were applying the typical English stress pattern to all items.

To determine if the greater use of stress pattern by JE as compared to E could best be explained by native language experience, experience with English, or some combination of the two, it was important to test native Japanese speakers with little English experience. J also evidenced a particularly large stress-by-position interaction [$F(1, 15) = 199.51$, $p < .01$] on phoneme localization, performing better with the normal English stress pattern ($M = 80\%$) than with the infrequent English stress pattern ($M = 38\%$) [$F(1, 15) = 29.70$, $p < .01$; d = 2.8]. When J and E were included in the same ANOVA, there was a large group-by-stress-pattern interaction [$F(1, 30) = 80.61$, $p < .01$] such that J showed a larger effect of stress pattern than did E (Figure 3). Furthermore, a group-by-stress-pattern interaction was found for J and JE [$F(1, 30) = 22.08$, $p < .01$] such that J showed a larger effect of stress pattern than JE.

Figure 3. Percent correct in words with normal English stress pattern and infrequent English stress pattern collapsed across sentence type (top panel). Difference in percent correct (normal – infrequent) by sentence type and group (bottom panel).



Unlike the other three groups, S did not show a stress-by-position interaction or any significant differences in the comparisons of the normal English stress pattern to the infrequent English stress pattern (overall or for either subset of stimuli).

Item analyses were also performed on the stress-pattern effects for each group of subjects. The results of these analyses were the same as the by-subject ANOVAs, indicating that no small subset of the stimuli was driving the effects reported above.

Stress-Pattern-by-Sentence-Type Effects

As previously reported, E relied more heavily on stress pattern when other cues—specifically lexical and syntactic information—were absent. They showed a stress-pattern-by-sentence-type interaction such that the effect of stress pattern was larger for *acoustic* sentences than for *syntactic* sentences, and larger for *syntactic* sentences than for *semantic* sentences, as shown in the bottom panel of Figure 3.

JE were able to use English stress pattern to segment speech. However, it was not clear if, like E, they would rely on this segmentation cue to a greater extent when other segmentation cues were absent. In fact, stress pattern did interact with sentence type for JE [$F(2, 30) = 18.41$, $p < .01$] such that the stress-pattern effect was larger for the *syntactic* and *acoustic* sentences than for the *semantic* sentences.

SE also showed a stress-pattern-by-sentence-type interaction [$F(2, 13) = 14.65$, $p < .01$]. For this group, the effect of stress pattern was only significant for the *syntactic* [$F(1, 13) = 18.71$, $p < .01$; $d = 2.4$] and *acoustic* [$F(1, 13) = 27.57$, $p < .01$; $d = 2.7$] sentences and was larger for the *acoustic* sentences.

Because neither J nor S were able to use lexical or syntactic information to the same extent as native English speakers, these groups would not be expected to rely on stress pattern to different extents for the different sentence types. This hypothesis was supported. Although J showed large stress-pattern effects, there was no stress-pattern-by-sentence-type interaction for this group. S did not show an effect of stress pattern across sentence types, nor was there a stress-pattern-by-sentence-type interaction for this group.

Summary

E, JE, SE, and J all performed better on the *semantic* sentences than on the *syntactic* sentences. E showed a larger effect of lexical information than J (Figure 1). E and JE performed better on the *syntactic* sentences than the *acoustic* sentences. However, this effect was larger for E (Figure 2).

E, JE, SE, and J showed a stress-by-position interaction such that their performance was more accurate for targets in words with the normal English stress pattern (stress on the first syllable) than those with the infrequent English stress pattern (unstressed on the first syllable). The effect of stress pattern was larger for JE and J than for E (Figure 3, top panel).

For E, JE, and SE, the stress-pattern effect interacted with sentence type. For E, the stress-pattern effect was larger in *acoustic* than *syntactic* sentences and larger in *syntactic* than *semantic* sentences. For JE, there was no significant difference in the size of the stress-pattern effect for the *syntactic* and *acoustic* sentences, but both of these sentence types showed a larger effect than did the *semantic* sentences. For SE, stress pattern was significant only for the *syntactic* and *acoustic* sentences. No stress-pattern-by-sentence-type interaction was found for either J or S (Figure 3, bottom panel).[1]

Discussion

Semantic and Lexical Processing

JE and SE did not acquire English until after the age of 12. However, both of these groups clearly benefited from the presence of lexical information—to an extent similar to that among native speakers. Although no single test or task can be used to definitively conclude that there are no differences in the ways in which native and non-native speakers process lexical or semantic information, the fact that both groups of late-learners were able to use the lexical information supports the hypothesis that the lexico-semantic system remains relatively plastic beyond the age of 12.

In contrast, S did not benefit from the presence of lexical information, indicating that they either did not know the English words used in the study or were not able to use the information to perform the task. Either way, this finding suggests that differences in lexical information in the *semantic* and *syntactic* sentences, and not differences in acoustic or prosodic information, resulted in better performance with the *semantic* sentences for those who knew English. J did benefit from the presence of lexical information, but not to the same extent as the groups that knew English well. It is possible that learning written English in school and exposure to English-language television, movies, and songs contributed to this group's ability to recognize at least some English words and to use this information to determine word onsets and detect target phonemes.

Syntactic Processing

No group of non-native speakers used syntactic information to the same extent as native speakers. This is consistent with a large literature showing that late acquisition of a language is associated with an inability to process syntactic information in a native-like manner. JE demonstrated that they could use syntactic information to help determine word onsets and detect targets, but not to the extent as could E. The other three groups of non-native speakers did not show any significant differences for the *syntactic* and *acoustic* sentences, indicating that language experience was the important factor in being able to make use of the differential information provided by these sentence types. These findings indicate that syntactic processing abilities are affected by learning another language first and/or by delaying second-language acquisition until after the age of 12, and suggest that syntactic systems do not remain as plastic as lexical/semantic systems.

Phonological and Prosodic Processing

From a previous study (Sanders & Neville, 2000), it was clear that native English speakers were more accurate at determining where word onsets occurred in continuous speech when targets were located in words with the normal English stress pattern. Surprisingly, both JE and J also were more accurate at detecting targets in words with the normal English stress pattern. It is clear that J knew at least some English because they performed better with the *semantic* sentences than with other sentence types. However, it was not expected that their limited English experience would be sufficient to help them learn and make use of English stress patterns.

Three plausible explanations for the finding that native Japanese speakers showed better performance with normal English stress pattern include the following: (1) Non-native speakers can learn new segmentation

[1] We conducted an additional experiment with all four groups in which participants performed a phoneme-detection task without indicating target location. The phoneme-detection task was employed to provide a more online measure of segmentation. Reaction times on that task (M = 805 ms) were much shorter than those for the present study (M = 1444 ms), although other results were similar across experiments. Results of the phoneme-detection study can be obtained by contacting the authors.

cues in general. (2) Something specific about stress as a segmentation cue makes it easy to learn. (3) There are segmentation cues that exist in Japanese and English that co-occur with English stress. The results of at least one other study of bilingual speech segmentation suggest that non-native speakers may be able to learn segmentation cues specific to their second language (Goetry & Kolinsky, 2000). Interestingly, this research was also concerned with the acquisition of stress pattern as a segmentation cue. Perhaps loudness and duration are such acoustically salient attributes that non-native speakers, even with very little experience with stress, are able to process stress information. However, even if stress information itself were salient, non-native speakers would still have to learn a new stress *pattern* to use stress as a segmentation cue.

Alternatively, it may be the case that the native Japanese participants (JE and J) were able to use some segmentation cue that is effective in both Japanese and English. It is difficult to rule out this possibility, because very little is known about segmentation cues for Japanese. As previously noted, there is some evidence that native Japanese speakers use morae to segment Japanese, but the existence of such a metrical segmentation cue does not preclude the existence of other useful segmentation information. The data from the native Japanese speakers in this study could either indicate that both groups were applying a Japanese segmentation cue that happens to co-occur with stress in English or that both groups had enough exposure to English to learn a new segmentation cue.

The findings concerning native Spanish speakers in the present study may help differentiate between these alternatives. S did not display any stress-pattern effects for phoneme localization, whereas SE did use stress pattern in the segmentation task and did so to an extent similar to that of native speakers. Furthermore, SE showed stress-pattern effects even on a subset of stimuli for which typical Spanish stress and typical English stress differ, suggesting that they were not applying Spanish stress-pattern rules to segment the speech in this experiment. As the native Spanish speakers did not apply Spanish-appropriate segmentation cues to English, it might be hypothesized that the native Japanese speakers did not apply Japanese-appropriate segmentation cues to English. This would mean that the most likely explanation of the stress-pattern findings for native Japanese speakers involves learning a new segmentation cue.

In either case, the fact that SE were able to use normal English stress pattern to segment speech in itself indicates that non-native speakers are able to learn new segmentation cues. In general, this suggests that the phonological or prosodic subsystems of language underlying the ability to learn stress pattern as a segmentation cue remain plastic beyond the age of 12.

Multiple Segmentation Cues

As reported previously, monolingual English speakers can use multiple segmentation cues flexibly. With the present set of stimuli, E relied on remaining segmentation cues to a greater extent when other types of segmentation information were not available. Because the non-native speakers in this study were less able to use some types of available segmentation information (e.g., syntactic), it was hypothesized that they may have relied on the segmentation cues that they could use to a greater extent. This hypothesis was supported by the localization accuracy of both native Japanese and native Spanish speakers. For both groups of late-learners of English (JE and SE), the stress-pattern effect was larger with *syntactic* and *acoustic* sentences than it was with *semantic* sentences, suggesting that these groups relied on stress pattern as a segmentation cue to a greater extent when lexical and semantic information was absent. As expected, the same was not true when syntactic information was absent, because the non-native speakers did not use syntactic information as a segmentation cue.

J and S showed no stress-pattern-by-sentence-type interactions. Again, these data support the hypothesis that listeners rely on any segmentation cues that are both available and usable. Because the non-English speakers were not using lexical or syntactic segmentation cues, they were not expected to rely more heavily on stress-pattern cues when lexical and syntactic information was absent.

Proficiency

Instead of administering general tests of language proficiency, specific tasks which indexed the use of lexical, syntactic, and stress-pattern segmentation cues were employed in this study. This methodology and the results reinforce two important points about the nature of language proficiency. First, the finding that the same non-native speakers evidenced native-like performance in some conditions but not others indicates that proficiency is not monolithic; speakers can and do attain different levels of proficiency for different subsystems within language. Thus, language proficiency cannot be adequately or accurately measured by tasks that index ability for only one aspect of language or by tasks that attempt to index language ability in general. Second, it is likely that even *within* a given subsystem of language individuals will master some aspects of language processing and yet fail to master others. In fact, it has been reported that non-native speakers process different syntactic structures to different extents (Weber-Fox & Neville, 1996). Therefore, measuring proficiency in individual subsystems of language may not even be possible. Instead of attempting to "measure proficiency," a more useful approach to the study of bilingual language processing might involve the careful inclusion of appropriate control groups (e.g., native speakers), thus allowing for a more meaningful and interpretable comparison of performance on several specific language tasks.

Sensitive Periods

In the present study, we compared native and non-native speakers with different language backgrounds (Japanese, Spanish) and different amounts of experience (extensive, almost none). Because we did not measure the effects of the same language experience at different stages of development, our results do not speak to the existence or structure of a sensitive period for language acquisition (Bruer, 2001). However, the finding that different subsystems of language show different degrees of plasticity may have important implications, both for interpreting previous findings related to sensitive periods in language development and for how this issue should be addressed in the future.

Within perceptual systems, there is evidence that specific subsystems have different sensitive periods. For example, within the visual system, altered experience has the most marked effects on subsystems such as ocular dominance columns and orientation selectivity at different time-periods in development; other subsystems retain the ability to change throughout life (Harwerth, Smith, Duncan, Crawford, & von Noorden, 1986; Kaas, Krubitzer, Chino, Langston, Polley, & Blair, 1990; Maurer & Lewis, 1998; Tychsen, 2001). The same may be true for language systems (Neville & Bavelier, 2001; Neville & Bruer, 2001). The results from the present study, which show that the same type of language experience can have differential effects on different subsystems of language, support the hypothesis that some linguistic subsystems may be more developmentally constrained than others. Rather than attempting to define a single sensitive period for language acquisition, it will be important to determine if and when specific subsystems within language have sensitive periods.

Conclusion

Overall, the results indicate that subsystems within language are differentially affected by both the characteristics of L1 and whether a language is learned as an L1 or as an L2 later in life. Consistent with previous research, the pattern of results suggests that lexical and semantic subsystems retain the ability to change to a greater degree than do syntactic subsystems. In contrast with previous studies of bilingual phonology, the results suggest that at least one aspect of prosody—lexical stress—can be learned after the age of 12. The findings also indicate that segmentation cues can be used flexibly by both native and non-native speakers, such that cues that are both available in the speech stream and usable by the listener are employed to a greater extent when other segmentation cues are either absent or not accessible to the listener. The finding that some segmentation cues are easily acquired by late-learners of a second language whereas others are not both provides more information about how language is processed in general and suggests ways in which adult language learners may more easily learn to process speech in a second language.

Acknowledgments

This research was presented in part as a poster entitled "Speech segmentation by bilingual speakers" at the 1998 annual meeting of the Cognitive Neuroscience Society (San Francisco). The research was supported by NIH, NIDCD grant DC000128. The authors wish to thank Wayne O'Neil and Yokio Otsu for help in collecting data in Japan and Donna Coch for comments on the manuscript.

References

Archibald, J. (1997) The acquisition of English stress by speakers of nonaccentual languages: Lexical storage versus computation of stress. *Linguistics, 35,* 167–181.

Bruer, J. T. (2001). A critical and sensitive period primer. In D. Bailer, Jr., J. T. Bruer, F. J Symons, & J. W. Lichtman (Eds.), *Critical thinking about critical periods* (pp. 3–26). Baltimore, MD: Paul H. Brookes.

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., & Segui, J. (1983). A language-specific comprehension strategy. *Nature, 304,* 159–160.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllables' differing role in the segmentation of French and English. *Journal of Memory and Language, 25,* 385–400.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. *Nature, 340,* 229–230.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology, 24,* 381–410.

Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language, 32,* 358–378.

De Groot, A. M. B., & Kroll, J. (1997). *Tutorials in bilingualism: Psycholinguistic perspectives.* Mahwah, NJ: Lawrence Erlbaum Associates.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41,* 78–104.

Goetry, V., & Kolinsky, R. (2000). The role of rhythmic cues for speech segmentation in monolingual and bilingual listeners. *Psychologica Belgica, 40,* 115–152.

Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics, 21,* 205–228.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition, 4,* 123–141.

Harwerth, R., Smith, E., Duncan, G., Crawford, M., & von Noorden, G. (1986). Multiple sensitive periods in the development of the primate visual system. *Science, 232,* 235–238.

Hyman, L. M. (1977). *Studies in stress and accent* (Southern California Occasional Papers in Linguistics No. 4). Los Angeles: University of Southern California Department of Linguistics.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21,* 60–99.

Johnson, J., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition, 39,* 215–258.

Kaas, J., Krubitzer, L., Chino, Y., Langston, A., Polley, E., & Blair, N. (1990). Reorganization of retinotopic cortical maps in adult mammals after lesions of the retina. *Science, 248,* 229–231.

Leather, J. (1987). F0 pattern inference in the perceptual acquisition of second language tone. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 59–81). Dordrecht: Foris.

Maurer, D., & Lewis, T. L. (1998). Overt orienting toward peripheral stimuli: Normal development and underlying mechanisms. In J. E. Richards (Ed.), *Cognitive neuroscience of attention: A developmental perspective* (pp. 51–102). Hillsdale, NJ: Erlbaum.

McDonald, J. L. (2000). Grammaticality judgements in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics, 21,* 395–423.

Neville, H. J., & Bavelier, D. (2001). Variability of developmental plasticity. In J. L. McClelland & R. S. Siegler (Eds.), *Mechanisms of cognitive development: Behavioral and neural perspectives* (pp. 271–287). Mahwah, NJ: Lawrence Erlbaum Associates.

Neville, H. J., & Bruer, J. T. (2001). Language processing: How experience affects brain organization. In D. Bailer, Jr., J. T. Bruer, F. J. Symons, & J. W. Lichtman (Eds.), *Critical thinking about critical periods* (pp. 151–172). Baltimore, MD: Paul H. Brookes.

Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex, 2,* 244–258.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language, 32,* 358–378.

Sanders, L. D., & Neville, H. J. (2000). Lexical, syntactic, and stress-pattern cues for speech segmentation. *Journal of Speech, Language, and Hearing Research, 43,* 1301–1321.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech.* New York: Newbury House.

Tychsen, L. (2001). Critical periods for development of visual acuity, depth perception, and eye tracking. In D. Bailer, Jr., J. T. Bruer, F. J. Symons, & J. W. Lichtman (Eds.), *Critical thinking about critical periods* (pp. 67–80). Baltimore, MD: Paul H. Brookes.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience, 8,* 231–256.

Received August 16, 2001

Accepted January 28, 2002

DOI: 10.1044/1092-4388(2002/041)

Contact author: Lisa D. Sanders, PhD, Linguistics Department, University of Maryland, 1401 Marie Mount Hall, College Park, MD 20742-7505.
E-mail: lsanders@wam.umd.edu

Appendix A. IPA Transcriptions of Example Sentences Shown in Table 2

Condition	Sentence	Example
SI	Semantic	ɪn ɔɹɾɝ tu ɹisaikl bɑɾlz ju hæf tə sɛpəreɪt ðɪm
	Syntactic	ɪn ɔɹɾɝ tu ləfæɾl bɔkəɹz ju hæf tə θægəmeɪt ðɪm
	Acoustic	ɑ ɪlgən di ləfæɾl bɔkəɹθ hɑ mæz di θægəmeɪt fʌn
SM	Semantic	ɪf ðʌ oʊnli θɪŋ ɪn ɪt wɝ təbæko ɪt wʊdnt kɑz so mʌtʃ hɑɹm
	Syntactic	ɪf ðʌ ɪlmi ʃɔɹd ɪn ɪt wɝ dobæɾi ɪt wʊdnt gæf so mʌtʃ hɪlm
	Acoustic	ɑs fʌ ɪlmi ʃɔrd ɛl ɔk hɔn dobæɾi æg hæpsl gæf ʃɑ nɛs hɪlm
WI	Semantic	ðʌ tʃɑild stɑpt kɹaɪŋ wɛn ə bəlun wʌz gɪvən tə hɝ
	Syntactic	ðʌ fɝp trɛpt plɔɪŋ wɛn ə bəɹil wʌz kæfən tə hɝ
	Acoustic	sʌ fɝp trɛp ploʊl rɑn i bʌril hɑf kæfəm gɪ wɛm
WM	Semantic	aɪ seɪvd məni sɪns loʊgreɪd tɪmbɝ wɝkt fɝ ðɪs pɹɑdʒəkt
	Syntactic	aɪ tʃeft ɹɑno sɪns mɪltrɑk dɛlbɝ mɛld fɝ ðɪs plæsəg
	Acoustic	o tʃɛft ɹɑno zælf mɪltrɑk dɛlbɝ mɛld sɪə fɑtʃ plæsəg
TA	Semantic	traɪ lʊkiŋ ʌndɝ ðʌ æfgæn fɝ ðʌ tɔɪ ju lɑst
	Syntactic	kwɪ mɛɾiŋ ʌndɝ ðʌ ɪədən fɝ ðʌ kaɪ ju mɑft
	Acoustic	kwɪ mɛɾəŋ æmkl fɑ ɪədən sæl tʃʌ kaɪ wʌ mɑft

Appendix B. Mean Accuracy by Stress-Pattern Condition

		Sentence		
Group	Stress Pattern	Semantic	Syntactic	Acoustic
JE	Strong-Initial	96%	87%	80%
	Strong-Medial	91%	65%	50%
	Weak-Initial	74%	36%	34%
	Weak-Medial	95%	83%	85%
SE	Strong-Initial	83%	71%	64%
	Strong-Medial	70%	50%	42%
	Weak-Initial	77%	58%	50%
	Weak-Medial	83%	69%	69%
J	Strong-Initial	85%	82%	79%
	Strong-Medial	27%	23%	27%
	Weak-Initial	57%	47%	45%
	Weak-Medial	82%	73%	77%
S	Strong-Initial	55%	55%	50%
	Strong-Medial	42%	47%	50%
	Weak-Initial	58%	47%	55%
	Weak-Medials	50%	47%	53%