UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Use of Word Segmentation Cues in Adults: L1 Phonotactics versus L2 Transitional Probabilities

Permalink

https://escholarship.org/uc/item/4bv8z5fs

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

ISSN

1069-7977

Authors

Finn, Amy S.
Hudson Kam, Carla L.

Publication Date

2006

Peer reviewed

Use of Word Segmentation Cues in Adults: L1 Phonotactics versus L2 Transitional Probabilities

Amy S. Finn (amyfinn@berkeley.edu)
Department of Psychology, 3210 Tolman Hall
Berkeley, CA 94720 USA

Carla L. Hudson Kam (clhudson@berkeley.edu)
Department of Psychology, 3210 Tolman Hall
Berkeley, CA 94720 USA

Abstract

We investigate whether adult learners' knowledge of phonotactic restrictions on word forms from their first language (L1) impact their word segmentation abilities in a new language. Adult learners were exposed to a speech stream in which language specific and non-language specific cues for word segmentation were pitted against one another. English rules about possible phonetic combinations (phonotactics) and transitional probabilities of syllables conflicted such that predictive transitional probabilities generated words that were phonotactically impossible in English. A control with phonotactically viable items was also run. At test, participants choose between words defined by deterministic transitional probabilities and words that are phonotactically possible in English, but have much lower transitional probabilities. A baseline of their abilities to track transitional probabilities in the stimuli was also collected. Results suggest that although participants are able to track the transitional probabilities in these stimuli, they are not using them to segment and extract words. Control subjects, however, do use transitional probabilities to segment words. This pattern of results is resilient, holding up with substantial increases in exposure and even when segmentation is encouraged by explicitly giving participants one of the words in the stream prior to exposure.

Introduction

A great deal of research has documented that adult language learners do not tend to be as proficient as children. (e.g. Johnson & Newport, 1989; Mayberry & Lock, 2003). One possible source for adults' difficulties is interference – their knowledge of a previously learned language interferes with their acquisition of features in a new language (Odlin, 1989). Typically, work on transfer and interference has focused on how aspects of the L1's grammar affect the learning of the same kind of feature in a second language (L2). For example, such features can include word order or relative clause formation. However, interference need not be so direct; knowledge of one aspect of a language could interfere with a learner's ability to acquire a different aspect of the new language. The current work investigates this second type of interference. In particular, we investigate the effect of a learner's L1 phonology—specifically phonotactics—on word segmentation.

Finding the boundaries between words is an important early step in language acquisition, and a great deal of recent research in first language acquisition has been directed at understanding how it is that infants are able to segment words from running speech. It is now clear that infants are able to use at least 3 cues present in their input in order to extract words: transitional probabilities (Saffran, Aslin & Newport, 1996), rhythm or prosody (Jusczyk, Houston & Newsome, 1999), and phonotactics (Mattys, Jusczyk, Luce & Morgan, 1999).

In a very well-known study, Saffran and colleagues showed that infants are able to use transitional probabilities to determine word boundaries (Saffran, Aslin & Newport, 1996). In these studies, syllables are presented one after another with no pauses in between; but some syllables are paired together 100% of the time and others are paired only 33% of the time. Words are defined as sequences of syllables with high transitional probabilities between the syllables, so that three syllables with transitional probabilities of 1.0 are a word, whereas a sequence with one high and one low transitional probability (e.g. 1.0 and .33) is not a word. Learners are able to determine word boundaries based on this information from very little exposure to the stimuli.

There is also evidence that infants use language specific cues to segment words. One such cue is rhythm or prosody. Jusczyk, Houston & Newsome (1999) for example found that 7.5 month old infants can parse out words that follow the more typical stress pattern of the language, and will mis-parse a speech stream when it contains words with a non-canonical stress pattern. Thiessen & Saffran (2003) showed that 9 month olds will use English prosodic cues to parse an artificial language, showing that infants' usage of prosodic cues is very robust, even at this early age. Importantly, this cue to word segmentation is language specific. In order to use this cue, learners must have already parsed at least a few words in their language so as to learn the typical stress pattern. Transitional probabilities, however, can be used without having previously segmented any units.

Another language specific cue is phonotactics, the constraints a language places on the ordering of segments within and between the words. These can be constraints such as allowable word initial consonant clusters, allowable phonetic combinations in syllable internal places, and so on. For example –ng is not allowed syllable initially in English. Mattys and colleagues presented infants with artificial, bi-syllabic words that contained two types of consonant

clusters—within-word and between-word. The former had a high probability of occurring within the words and the latter a high probability of occurring at word boundaries. They found that infants are able to segment words based solely on this information (Mattys, Jusczyk, Luce & Morgan, 1999). Phonotactics, like transitional probabilities, are a type of co-occurrence information. However, it is co-occurrence given a particular position within a word. Therefore, like prosody, in order to use phonotactic information to segment words learners must already know some words.

Not surprisingly, it appears that infants might start out using statistical information for word segmentation, but after they have segmented a few words—and thereby had an opportunity to learn more about their specific language—they move on to using more language-specific strategies. In fact, Thiessen & Saffran (2003) showed that infants' cue choice changes with age: when cues conflict and either can be used to segment a speech stream, younger infants (6 months old) use transitional probabilities, and 9 month old infants use prosody, not transitional probabilities. Johnson & Jusczyk (2001) have a similar finding.

It is an open question, however, as to what strategies adults use when learning a new language. Are they like the older infants using the language specific strategies they already know or do they start fresh, using the statistics to build a new set of language specific strategies?

The Current Research

In our current work, we seek to test whether adults apply language specific strategies from their L1 in order to segment a new language. More specifically, studies presented here ask if adults' previously existing phonotactic knowledge interferes with their ability to segment and extract words in a new language.

This is an open question of much interest because interference, as an explanation for sensitive period effects, cannot be taken as a given. Mayberry and colleagues, for example, consistently show that late language learners with no previous exposure to a language consistently perform worse than late language learners who have an L1 (Mayberry & Lock, 2003). Nonetheless, it is still of interest to investigate the role of interference with an eye toward sensitive periods using a statistical learning paradigm because (1) many studies have highlighted interference as very important (see Oldlin, 1989 for review) and (2) the artificial language literature, in large part, does not consider the role of previous knowledge.

Experiment 1

In the first experiment we simply pitted two word segmentation cues—transitional probabilities and English phonotactic rules about word initial clusters—against one another. If adults use transitional probabilities and ignore phonotactic constraints from their L1, this would suggest that they *are* starting fresh with each new language. If instead, they parse the speech stream according to English phonotactic rules, it would indicate that their prior knowledge is interfering with the acquisition of a new language. Crucially, we also constructed control stimuli

where word initial clusters were complex, but legal according to English phonotactics. Since these stimuli do not conflict, we expect participants to parse the speech stream according to transitional probabilities.

Method

Participants All 40 participants were native speakers of English. Experimental participants were 7 men and 13 women, ranging in age from 19 to 21 years. Control participants were 6 men and 14 women ranging in age from 18 to 22 years. All individuals participated in partial fulfillment of class requirement.

Materials Experimental and control stimuli each consisted of 8 two-syllable words (CCVCV), each beginning with a consonant cluster. For the experimental stimuli, these CC onsets violate the word-initial phonotactic rules of English. In the control stimuli, CC onsets are licit.

Experimental stimuli were /tfo-bu/, /čpʌ-zi/, /bte-gʌ/, /kmo-du/, /ʊti-sa/, /fse-lo/, /psu-ne/, and /θmʌ-re/[1]. These words were presented in quasi-random order with no pauses between them. The stimuli were created so that a word never followed itself, and all words were equally likely to follow all other words, yielding transitional probabilities of 100% for word internal syllable transitions and 14.2% for word boundary syllable transitions.

Each word was generated with the text to speech program, SoftVoice (Katz, 2005). The synthesizer produced syllables with a monotonic F0 (fundamental frequency) of 83.62 Hz. All vowels are the same length regardless of placement next to particular consonants or other phonemes and there were no co-articulation effects. We used synthesized speech precisely because it allows better control of the above mentioned parameters. Any natural production of speech could provide segmentation cues through varying degrees of co-articulation, different vowel lengths, amplitudes, frequencies, etc. A sample stretch of input is presented here: '*kmodutfobuvteesaabtaeguhfsaelopsunaethmaraetfobubtaeg uhpsunaekmoduvteesaafsaelothmarae.*'

Control stimuli were constructed and concatenated using the same methods. The 8 words are /zwo-bu/, /kre-gʌ/, /plo-du/, /bli-sa/, /ʊre-lo/, /θrʌ-zi/, /twu-ne/, and /stʌ-re/.

In both conditions, participants were exposed for a total of 17 minutes and 59 seconds. Each word occurred 560 times during the presentation phase. After exposure participants were given a forced-choice test to see whether or not they had correctly identified the word boundaries as defined statistically, despite the presence of the violating consonant clusters. Test items were of two types: (1) Word vs. Non-word and (2) Word vs. Split-cluster word. Word test items were the statistically defined words to which participants had been exposed. Non-word test items consisted of the first syllable from one word paired with the second syllable from

---

[1] Some of these consonant clusters appear in natural languages, while some do not. We will return to this point in the discussion.

a completely different word, e.g. *kmo-rae*. Although participants had heard each of the two syllables in the Non-word, they had never heard the two syllables in succession.

Split-cluster test items consisted of an exposure word, minus the first consonant, and with another word's initial consonant at the end, giving them a CVCVC structure. We refer to these as split cluster words because we split the consonant cluster at the beginning of the word, making experimental stimuli viable in English. Note that the split-cluster test items have lower transitional probabilities than the words. For example, *mo-duth* does not violate English phonotactics, but the *th* sound only follows *du* 14.2% of the time in the exposure stimuli. By adding this sound to the end of the word, we are actually making the forced choice options easier for the participant: since 14.2% is a very low transitional probability '*duth*' should sound odd to them even though the word as a whole is more viable in English.

There were 8 of each type of test (Word vs. Non-word and Word vs. Split-cluster word) yielding 16 test items in total. Test items of the two types were interleaved quasi-randomly, and two different counterbalanced versions of the test were created. Test stimuli were generated with exactly the same procedure as exposure stimuli.

Procedure Participants were run individually in a quiet room. They wore headphones during presentation and testing to eliminate any outside noise. Participants were told simply to listen to the stimuli as best they could, not to tune it out and not to analyze or think too much about it. To encourage this, participants engaged in a coloring task during exposure.

Each participant was exposed to the stimuli for 17 minutes and 59 seconds after which time they completed the test. During test, participants were asked to choose which of two items were more likely to be words in the language they were exposed to. Forced choice items were presented over headphones, one after another with a 1 second pause in between. Participants were then given 3 seconds to circle their choice on paper; they were instructed to circle 1 if the first one sounded better and 2 if the second one sounded better. After the test, subjects were given a questionnaire asking about their demographic and language backgrounds.

Results

The dependant measures are participants' forced choice responses on Non-word and Split-cluster test items. Data are shown separately in Figure 1 for experimental (black bars) and control subjects (grey bars).

Experimental condition: A t-test on the responses from the Non-word questions showed that participants choose the word with deterministic transitional probabilities more often than chance $t(19) = 9.19$, $p < .001$ indicating that learners are able to track transitional probabilities in these stimuli.

A t-test on the Split-cluster test items showed that participants did not choose the correct word (defined by transitional probabilities of 100%) more often than chance $t(19) = -.175$, $p = .863$. This suggests that although learners have access to this statistical information, it does not seem to guide their segmentation or extraction of the words.

Control condition: A t-test on the responses from the Non-word questions showed that participants choose Words over Non-Words more often than chance $t(19) = 16.35$, $p < .001$. This indicates that they are also tracking transitional probabilities in these control stimuli. A t-test on the Split-cluster test items did however reveal a difference, showing that participants choose Words over Split-cluster words more often than chance $t(19) = 20.38$, $p < .001$. Therefore, when complex CC onsets are viable in English, there does not seem to be any interference for the word segmentation task.
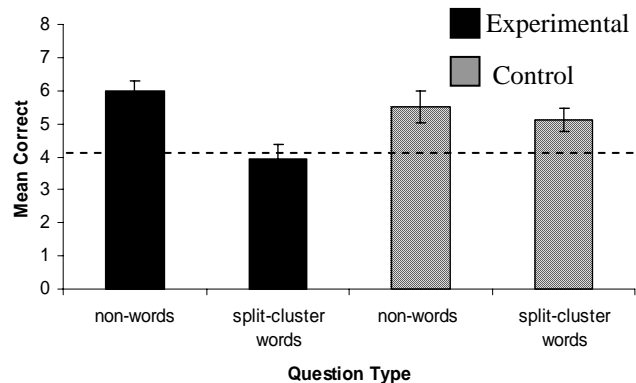

Figure 1: Experiment 1

Experiment 2

Results from Experiment 1 suggest that previous language experience can interfere with adult learners' abilities to use statistical information to extract words from the running speech. However, these learners were only exposed to speech stimuli for about 18 minutes. It could be that interference only plays a role early on but with increased exposure, this phenomenon fades. For this reason, in Experiment 2, we doubled exposure time.

Method

Participants were 20 native English speakers, 3 men and 17 women, ranging in age from 18 to 23 years. Individuals participated in partial fulfillment of class requirement.

Materials The stimuli for this experiment were the same as for the experimental condition in Experiment 1. Participants, however, listened twice as long for a total of 36 minutes. In so doing, they heard each word a total of 1,120 times. Test items were also the same as Experiment 1.

Procedure Save for the extension of exposure time, the procedure was exactly the same.

Results

A t-test on the responses to Non-word questions showed that participants choose the word with deterministic transitional probabilities more often than chance $t(19) = 2.38$, $p < .05$. However, a t-test on the Split-cluster test items showed that participants still did not choose the correct

word (defined by transitional probabilities of 100%) more often than chance $t(19) = .65$, $p = .522$. Therefore, doubling the exposure period from Experiment 1 did not alter the pattern of results.
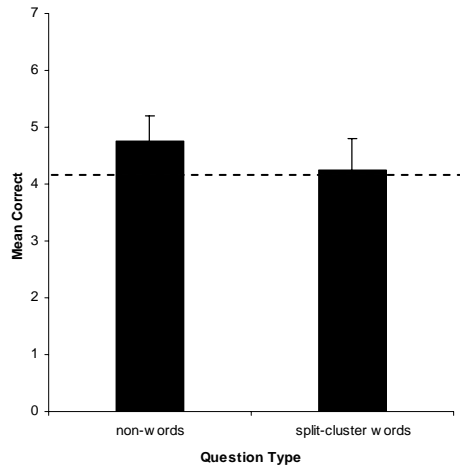

Figure 2: Experiment 2

Experiment 3

Results from Experiments 1 and 2 suggest that previous language experience may be interfering with learners' abilities to use statistical information in order to segment and extract words. This seems to be true even with substantial exposure time. Still, 36 minutes of exposure may not have been enough. For this reason, we decided to double the exposure time yet again, yielding 72 minutes of exposure total.

We chose to split the exposure period into two sessions occurring over two days. Research suggests that learning and memory is enhanced in adults and infants by sleep (Plihal & Born, 1997; Gomez, Bootzin & Nadel, in press; Stickgold & Walker, 2005). We reasoned that splitting the exposure and allowing for a night of sleep in between sessions might help the adult learners, and even assist them in properly segmenting the speech stream.

Method

Participants were 20 native English speakers, 4 men and 16 women, ranging in age from 18 to 33 years old**.** Individuals participated in partial fulfillment of class requirement.

Materials The stimuli and test items for this experiment were exactly the same as in Experiments 1 and 2, except that the total exposure time was 72 minutes—4 times that of Experiment 1, and double Experiment 2. During exposure, participants heard each word 2,240 times.

Procedure Save for the extension of exposure time, the procedure was exactly the same. On day one, participants got 36 minutes of exposure. On day two, participants received an additional 36 minutes of exposure and

completed the test. Sessions were always completed on consecutive days.

Results

A t-test on the responses to the Non-word questions showed, yet again, that participants choose the word with deterministic transitional probabilities more often than chance $t(19) = 9.45$, $p < .001$. A t-test on the Split-cluster test items also showed, yet again, that participants did not choose the correct word more often than chance $t(19) = 1.60$, $p = .126$. The pattern of results from experiments 1 and 2 was not altered even when exposure time is quadrupled and subjects are given an opportunity to sleep between exposure periods.
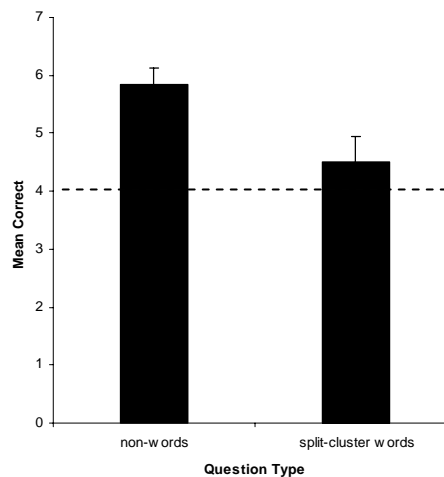

Figure 3: Experiment 3

Experiment 4

Taken together, the first three experiments strongly suggest that previous language experience interferes with learners' abilities to use statistical information to correctly segment and extract words from a new language. This finding, however, could be a product of this artificial exposure situation. In reality, language learners do not get exposed to completely unparsed stimuli for so long. That is, they generally will hear at least a few words in isolation. In fact, it has been argued that these initial words are what learners use to make initial generalizations about the prosody and phonotactics of their language—generalizations they later employ to further parse the speech stream (see Werker &Yeung, 2005). It has also been argued that having a few initial words can provide initial anchors in the speech stream, trimming it down and generally easing the task of segmentation (Bortfeld, Morgan, Golinkoff & Rathbun 2005).

For these reasons, in Experiment 4 we gave participants one of the words before exposure to see if this would improve their segmentation performance. We decided to give them *kmo-du* after an item analysis revealed that participants got this item correct 51% of the time in Experiments 1-3. (Some words were more likely to be

segmented correctly, and others were less likely.) Since their performance at test with this word is no different than chance, we thought this would be a good candidate to help them parse the speech stream.

Method

Participants were 20 native English speakers, 8 men and 12 women, ranging in age from 18 to 22 years old**.** Individuals participated in partial fulfillment of class requirement.

Materials The stimuli used in this experiment were exactly the same as stimuli used in Experiment 1 (lasting 17 minutes and 59 seconds, and consisting of 560 repetitions of each of the 8 words). Test items were also exactly the same. Both Non-word and Split-cluster word comparisons for *kmo-du* were collected but not included in the analysis.

Procedure The procedure for experiment 4 is exactly the same as for Experiment 1, except that participants were told before exposure that *kmo-du* is a word in the language they were about to listen to.

Results

A t-test on the responses to the Non-word questions showed that participants choose the word with deterministic transitional probabilities more often than chance $t(19) = 4.19$, $p < .001$. This is consistent with the previous experiments. A t-test on the Split-cluster test items was also consistent with previous experiments; participants still did not choose the correct word more often than chance $t(19) = 1.10$, $p = .286$. It appears that even when participants are given a full word before exposure, it does not help them to override their native language's phonotactic rule system. This is true even though participants remembered the word we gave them before exposure; all participants chose *kmo-du* with 100% accuracy.
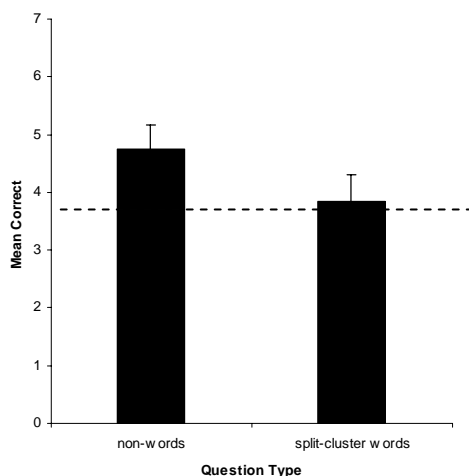

Figure 4: Experiment 4

General Discussion

In all the experiments above, it is clear from the Word vs. Non-word test items that the adult participants are tracking some kind of transitional probabilities in the new language; this is something that has been demonstrated in previous studies (e.g. Saffran, Newport & Aslin, 1996). However, the current data does not necessarily indicate what transitional probabilities they are tracking, participants could score above chance on the Word versus Non-word items by computing the transitions between syllables or between vowels in syllables, for example. Still, whatever they are tracking, it is clear that they are learning something about the statistics present in the input.

This brings us to the results for the Split-cluster words. For all experiments except the control, participants were at chance on these items. The interpretation of the control data is straight forward: participants are using transitional probabilities to guide their segmentation of the words. Unfortunately, chance performance for experimental stimuli is not as clear cut. If participants were *above* chance on these items, it would indicate that they are able to override their L1 knowledge of cues to word boundaries and use instead the statistics present in the input. They would be starting fresh. Alternatively, if their performance was significantly *below* chance, this would indicate complete interference of L1 phonotactics. The current data, however, are more intermediate, suggesting a moderate amount of interference. Although we are confident that some interference is going on, we cannot be sure as to precisely where they are segmenting the speech. What, if any, units are they extracting? Future studies aim to answer precisely this question. We plan to include a battery of more delicate forced choice options designed to probe the nature of the extracted units.

It is worth noting that while all participants were native speakers of English, some were also proficient in other languages. A close look at the grammars for all these languages revealed that only one language of which our sample of participants were proficient (French) allowed any of the consonant cluster onsets in our stimuli. We only have two speakers of French in our sample. This is too few to be able to draw conclusions about interference from a language other than English. We did however re-run analyses reported in the previous experiments without these individuals. Removing them did not alter the pattern of results.

In additional studies, we aim to explore if this early mis-segmentation bleeds into other aspects of language acquisition upstream, such as morphology. We are also eager to see which of the consonant clusters are more and less difficult to parse and whether this is related to how natural or common these clusters are in languages of the world. A related question is how these clusters are perceived. Is there some interference on a perceptual level or is this a question of leanabilty? Of central interest is what infants and children do with these exact stimuli. Is it necessary to be an infant to start fresh? Do you simply need to be within the critical period? These are all open questions.

Acknowledgments

References

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun. K. (2005). Mommy and me: familiar names help launch babies into speech-stream segmentation. *Psychological Science, 16(4),* 298-304.

Gómez, R. L., Bootzin, R. R., & Nadel, L. (in press). Naps promote abstraction in language learning infants. *Psychological Science.*

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, *44(4),* 706-716.

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(1),* 60-99.

Jusczyk, P.W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39(3-4),* 159-207

Katz, J. (2005). SoftVoice (Demo Program) [Computer Software]. Los Angeles, CA: SoftVoice, Inc.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology, 38(4),* 465-494.

Mayberry, R. I., & Lock, E. (2003). Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. *Brain and Language*, *87(3),* 369-384.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.

Plihal, W., Born, J. (1997). Effects of early and late nocturnal sleep on priming and spatial memory. *Psychophysiology, 36*, 571-582.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274(5294),* 1926-1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: the role of distributional cues. *Journal of Memory and Language, 35*, 606-621.

Stickgold, R., & Walker., M. (2005). Memory consolidation and reconsolidation: What is the role of sleep. *Trends in Neurosciences*, *28*, 408-415.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, *39(4),* 706-716.

Werker, J. F., & Yeung, H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences, 9(11),* 519-527