Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Memory and Language 53 (2005) 225–237

Journal of Memory and Language

www.elsevier.com/locate/jml

# Phonology impacts segmentation in online speech processing ☆

Luca Onnis [a,*], Padraic Monaghan [b], Korin Richmond [c], Nick Chater [d]

[a] Department of Psychology, Cornell University, 14850 Ithaca, NY, USA
[b] Department of Psychology, University of York, York YO10 5DD, UK
[c] Institute for Applied Cognitive Science and Department of Psychology, University of Warwick, Coventry CV4 7AL, UK
[d] Centre for Speech Technology Research, Division of Informatics, University of Edinburgh, 2 Buccleuch Place, Edinburgh EH8 9LW, UK

Received 19 March 2004; revision received 21 February 2005
Available online 19 April 2005

## Abstract

Peña, Bonatti, Nespor, and Mehler (2002) investigated an artificial language where the structure of words was determined by nonadjacent dependencies between syllables. They found that segmentation of continuous speech could proceed on the basis of these dependencies. However, Peña et al.'s artificial language contained a confound in terms of phonology, in that the dependent syllables began with plosives and the intervening syllables began with continuants. We consider three hypotheses concerning the role of phonology in speech segmentation in this task: (1) participants may recruit *probabilistic phonotactic* information from their native language to the artificial language learning task; (2) *phonetic properties* of the stimuli, such as the gaps that precede unvoiced plosives, can influences segmentation; and (3) *grouping by phonological similarity* between dependent syllables contributes to learning the dependency. In a series of experiments controlling the phonological and statistical structure of the language, we found that segmentation performance is influenced by the three factors in different degrees. Learning of nonadjacent dependencies did not occur when (3) is eliminated. We suggest that phonological processing provides a fundamental contribution to distributional analysis.


*Keywords:* Artificial language learning; Statistical learning; Segmentation; Phonology

## Introduction

Artificial language learning (ALL) provides a methodology for a highly controlled analysis of how learners can learn to extract structure from speech-like stimuli. By using small-scale artificial languages, the structure of which can be learned during the course of a brief experimental session, it is possible to put learning processes under the experimental microscope. Yet when ALL studies are conducted with adult participants, there is inevitably a substantial possible complication—that the adults' knowledge of their native language may influ-

☆ Part of this work was carried out while Luca Onnis and Padraic Monaghan were at the University of Warwick. L. Onnis, and N. Chater were supported by the European Commission Project HPRN-CT-1999-00065. P. Monaghan was supported by a Human Frontiers of Science Grant RGP0177/2001-B. K. Richmond was supported by EPSRC Grant No. GR/R94688/01. Thanks to Marjolein Mercx for data collection, and to Morten Christiansen, Rebecca Gómez, and Friederike Schlaghecken for helpful comments. We also thank James Morgan, Letitia Naigles, and an anonymous reviewer.
* Corresponding author. Fax: +1 607 255 8433.
*E-mail address:* lo35@cornell.edu (L. Onnis).

0749-596X/$ - see front matter 
doi:10.1016/j.jml.2005.02.011

ence their processing of the artificial language. In particular, the phonological structure of the ALL stimuli, and its relation to the phonological structure of the native language of the participants, provides a potentially rich source of information that learners may draw upon in performing experimental tasks (Brent & Cartwright, 1996; Friederici & Wessels, 1993; Mattys, Jusczyk, Luce, & Morgan, 1999).

In this paper, we investigate some of the phonological properties of ALL stimuli that may contribute to performance; and we suggest that taking account of these phonological properties may lead to a reinterpretation of some ALL findings. In particular, we focus on a series of segmentation experiments that assessed learners' ability to detect nonadjacent dependencies, i.e., dependencies between syllables that are not directly adjacent in connected speech. Currently, there are contradictory results in the literature. Peña, Bonatti, Nespor, and Mehler (2002) found that nonadjacent dependencies between syllables could be learned in an ALL task, and that this learning contributed to segmenting a pause-free stream of sounds into words. By contrast, Newport and Aslin (2004) obtained the opposite result—learners were not able to segment similar stimuli accurately. We attempt to reconcile these opposing data by looking at the contribution of phonological properties in Peña et al.'s experimental materials.

The results from Peña et al. (2002) were taken to support a separation between different types of computational processing in language learning. Their claims have contributed to the debate on the extent to which language acquisition is dependent on the statistical structure of the language environment, or on algebraic, rule-like computations (Hahn & Chater, 1998; Marcus, 1999; McClelland & Plaut, 1999). This question has been central to debates about language acquisition, and is ubiquitous at different levels of description of language structure. Peña et al. (2002) argued that statistical and algebraic computations could be reconciled: speech segmentation operates on the basis of statistical learning, whereas entirely separate algebraic computations are necessary for learning grammatical structure. In this paper, we present a series of experiments to show that the line of ALL evidence that they have pursued does not yet support this segregation of computational processes. We explore phonological confounds in the materials used by Peña et al. which reveal a complex but systematic interaction between phonological and structural information in ALL experiments.

### *Contributions of phonology to language learning*

Adults, young children, and infants readily find relations between adjacent items in sequences of stimuli, such as syllables (Saffran, Aslin, & Newport, 1996), tones (Saffran, Johnson, Aslin, & Newport, 1999), or visual items (Kirkham, Slemmer, & Johnson, 2002). In contrast, evidence for learning the relations among nonadjacent items is scarce, and seems to occur only under specific circumstances, such as when intervening material is highly variable (Gómez, 2002). Yet, computing adjacent information in a sentence like [the books on the shelf are dusty] would fail to detect the correct noun–verb agreement between *books* and *are*, instead producing [*The books on the shelf is dusty] with agreement between the adjacent noun *shelf* and the verb *is*. Nonadjacent dependencies are therefore an essential feature of language structure that must be available to the language user.

Peña et al. (2002) provided a set of intriguing ALL studies that seemed to suggest that nonadjacent dependencies can be learned, but that this learning can only be applied selectively. Specifically, they argued that knowledge of nonadjacent dependencies can be used for segmentation (which they take to be a statistical computation), although they cannot simultaneously be used for learning rules in the language (which they take to be an algebraic computation).

Here we focus on Peña et al.'s (2002) experiments on segmentation, and how far participants' segmentation performance provides evidence for the learning of nonadjacent dependencies. Seidenberg, MacDonald, and Saffran (2002) suggested that phonological properties of the stimuli might be a crucial confound.

Peña et al.'s participants were presented with continuous streams of syllables comprised of words of the form $A_iXB_i$, where there were three such $A_i\_B_i$ pairs, and $X$ was one of three syllables that randomly intervened between the $A_i\_B_i$ pair. The artificial language generated three sets of nine words altogether: the first set ($A_1XB_1$) was [pu-li-ki], [pu-ra-ki], [pu-fo-ki]; the second set ($A_2XB_2$) was [be-li-ga], [be-ra-ga], [be-fo-ga]; and the third set ($A_3XB_3$) was [ta-li-du], [ta-ra-du], [ta-fo-du]. In a subsequent forced-choice task, participants demonstrated a preference for words as they were construed (e.g., $A_1X_2B_1$, [pu-li-ki]) over part-words, i.e., sequences that spanned word boundaries (e.g., $X_2B_1A_3$, [li-ki-ta], or $B_3A_1X_2$, [du-pu-li]). Both word and part-word sequences had appeared in the training phase. In the absence of acoustic cues to word boundaries in the training stream, preference for words is presumed to be made on the basis of distributional information. Specifically, Saffran, Newport, and Aslin (1996) suggest that segmentation is determined by low transitional probabilities—word boundaries are presumed to be conjectured at points where the next syllable is particularly difficult to predict given the previous syllable.

In Peña et al.'s experiments, adjacent transitional probabilities between any $A_i$ and $X$ and any $X$ and $B_i$ (within-word) were .33, while transitional probabilities between a $B_i$ syllable and an $A_i$ syllable (between-word) were .5 (words belonging to the same nonadjacent family did not follow one another). The nonadjacent dependencies between the $A_i$ and the $B_i$ were always 1, while nonadjacent dependencies across word boundaries were lower, $\Pr(A_i|X_{\text{previous}}) = .33$, $\Pr(X|B_{\text{previous}}) = .33$. Peña et al.'s experiments tested two alternative hypotheses: If participants are segmenting using the lowest transitional probabilities of adjacent items, as argued by Saffran, Newport, et al. (1996), they would prefer part-words, because the least predictable point in the strings is between the $A_i$ and $X$. Alternatively, if participants are sensitive to the $A_i\_B_i$ nonadjacent probabilities they would prefer words. Hence, in order to segment the speech stream correctly learners had to disregard adjacent probabilities and detect the nonadjacent ones. The results showed that nonadjacent dependencies of the syllables were learned and contributed towards segmentation. However, this result contrasts with findings from an experiment using a very similar artificial grammar, by Newport and Aslin (2004). They found no learning of nonadjacent syllables:[1] in one version of their study (Experiment 2: Language A) they used three nonadjacent syllable frames ([di_tae], [po_ga], and [te_bu]) and 2 intervening syllables ([ki] and [gu]), so that transitional probabilities between nonadjacent syllables were 1.0 and all other adjacent and nonadjacent transitional probabilities were .5 or lower. So far, these two different results have not been reconciled.

Seidenberg et al. (2002) pointed out that in each experiment in Peña et al.'s (2002) study, syllables in the same positions were used for all participants: all initial and final syllables began with a plosive consonant and all medial syllables began with a continuant. This phonological structure in Peña et al.'s study may have contributed to learning to segment speech for several reasons, considered below. However, if segmentation can be based on distributional information alone, as Peña et al. claim, then the phonological properties of their stimuli should not be crucial. If so, then learners should be able to segment an artificial language with the same nonadjacent dependencies as Peña et al.'s original experiment, but with no confounding phonological structure. After replicating Peña et al.'s segmentation experiment in Experiment 1, we test this in Experiment 2.

The first hypothesis about the role of phonology in speech segmentation in Peña et al.'s study is that knowledge of phonotactic constraints (whether absolute or probabilistic) derived from the participant's own language may be recruited to segment the experimental stimuli. It has been proposed that very young children may develop implicit knowledge of the distributional regularities of sounds in order to bootstrap the basic units of language. For instance, in the second half of their first year children begin to distinguish strings of sounds containing legal sequences in their language from illegal sequences (Jusczyk, 1999a, 1999b), and this might help them segment words correctly (e.g. [penguins.would] versus [*penguin.swould]). Likewise, in adult language processing, where speech segmentation has to be resolved online, McQueen (1998) has shown that Dutch listeners spot a word more easily when it is aligned phonotactically (e.g., *pil* in [pil.vrem] in Dutch) than when it is misaligned with a boundary (e.g., *pil* in [pilm.rem]).

Such phonotactic constraints can be absolute or probabilistic. Absolute constraints produce sentences that are illegal, as the ones mentioned above: for instance /zw/ and /vɹ/ never appear at the beginning of words in English, although they do in Dutch[2]. In contrast, probabilistic phonotactic constraints provide information about the likelihood of certain sounds occurring in certain positions within words, such as at word onset, word offset, or within the word. Kessler and Treiman (1997), in an extensive examination of English consonant–vowel–consonant (CVC) monomorphemes, found that not all consonant sounds were equally good word onsets in English. Mattys and Jusczyk (2001) found that infants listened to CVC words longer when the stimulus previously appeared in a sentential context with good phonotactic cues than when it appeared without such cues. They found that good cues to word boundaries were associated with high between-word probability as obtained from a corpus of child-directed speech. For example, in English, the between-word sequence /ng/ and [fh] are good cues, respectively, for onset and offset position ([bean gaffe hold…]), whereas [ŋg] and [ft] are bad cues respectively at onset and offset ([..fang gaffe tine…]). Mattys and Jusczyk also found that effective segmentation also resulted when good phonotactic cues occurred only at the onset or the offset of the target words in the utterances.

Several studies in infant and adult speech perception have documented the potential impact of speech cues in detecting word-like units. Johnson and Jusczyk (2001) found that in the presence of speech cues conflicting with distributional cues for word boundary assignment, the former were preferred by 8-month-olds in segmenting an artificial stream of sounds. The authors concluded that coarticulation and stress override distributional statistics, perhaps because this information is more readily available or perceptually more salient. Although there may be several types of speech cues involved in natural speech segmentation such as syllable lengthening (Quené, 1992), and metrical information (Norris,

[1] Newport and Aslin's results are not entirely negative, because they found successful nonadjacent learning of phonetic segments, such as consonants and vowels, e.g. [p_g_t] or [_a_u_e]. Here, we are interested in their findings of unsuccessful segmentation with nonadjacent syllables, because they contrast directly with Peña et al.'s.

[2] Though the phoneme is realized as /R/ in Dutch.

McQueen, Cutler, & Butterfield, 1997), these were eliminated by Peña et al. by creating a synthesized stream of concatenated syllables of the same duration, pitch, amplitude, and characterized by the absence of stress or other prosodic features. However, the presence of probabilistic phonotactic constraints, i.e., the skewed distribution of specific sounds in specific contexts, might not have been controlled thoroughly. It is therefore possible that probabilistic phonotactic information about the specific onsets of initial, medial, and final position syllables used by Peña et al. exerted an influence on the results, rather than participants learning the statistical or algebraic properties of the stimuli. In this case, an experiment with the same phonological structure as Peña et al.'s original experiment but with no nonadjacent structure ought to result in good segmentation performance. We test this in Experiment 3.

Another influence of phonology in ALL experiments may be due to the phonetic properties of speech stimuli, particularly when produced by speech synthesizer programs. Perruchet, Tyler, and Galland (in press) investigated the output of the MBROLA speech synthesizer in producing the continuous French speech in Peña et al.'s experiments, and found that unvoiced plosives were preceded by silent gaps unusually longer than those occurring in French natural speech. In MBROLA the silence gaps preceding unvoiced plosives are proportional to the length of the whole phoneme. The duration of phonemes in Peña et al. was considerably longer (116 ms) than the mean duration of unvoiced plosives in French (Perruchet et al. report a range of 60–120 ms with a median around 75 ms in speech samples of French). In consequence, there were silent onsets preceding the articulation of plosive sounds generated by MBROLA which were longer than in natural speech. Two out of three words in Peña et al. began with unvoiced plosives (/p/ and /t/), hence these words would be preceded by a gap in the speech stream, which Perruchet et al. (in press) suggested would contribute to segmentation before these consonants. To test this, we constructed speech stimuli with continuants in word initial position, and plosives in medial and final positions in the words (Experiment 4). If Perruchet et al. were correct then this would result in preference for part-words over words—part-words now beginning with plosives more often than words.

A third possible role of phonology in ALL experiments is that items in the speech stream may be assigned to the same word because they are grouped by phonological similarity. In Peña et al.'s materials, the first and the third syllable begin with a plosive, and are distinct from the continuant property of the intervening syllables. Thus, the role of the plosive in word onset position may only be effective when the final syllable also begins with a plosive, an issue which is addressed in Experiments 5 and 6.

Finally, note that it is possible that distributional information can be extracted and used in segmentation; but that the process of extracting this information operates in consort with phonological cues. Braine (1987) claimed that learning of grammatical structure could not be achieved unless there was phonological coherence among words of the same category. Similarly, Morgan and Newport (1981) showed that dependencies are more readily learned when learners are provided with phonological cues that link the stimuli between which the dependencies hold. This possibility is also explored.

## Experiment 1

In Experiment 1, we replicated Peña et al.'s study of segmentation based on nonadjacent dependencies within words in continuous speech, except that we used English speech stimuli and English participants.

### Method

#### Participants

Fourteen undergraduate and postgraduate students at the University of Warwick participated for £1. All participants spoke English as a first language and had normal hearing.

#### Materials and design

We used the same nine word types from Peña et al. to construct the training speech stream in Experiment 1. The set of nine words was composed of three groups ($A_i\_B_i$), where the first and the third syllable were paired, with an intervening syllable ($X$) selected from one of three syllables. The first set ($A_1XB_1$) was: [pu-li-ki], [pu-ra-ki], [pu-fo-ki]; the second set ($A_2XB_2$) was: [be-li-ga], [be-ra-ga], [be-fo-ga]; and the third set ($A_3XB_3$) was: [ta-li-du], [ta-ra-du], [ta-fo-du]. In IPA format, the $A_i\_B_i$ pairs were /pu__ki/,/bɛ__gæ/, and /tæ__dʌ/,and the intervening syllables were /ɹæ/, /fəʊ/, and /li/.

We used the Festival speech synthesizer (Black, Taylor, & Caley, 1990) using a voice based on British-English diphones at a pitch of 120 Hz, to generate a continuous speech stream lasting approximately 10 min. All syllables were of equal duration, and were produced at a rate of 4.5 syllables/second. The speech stream faded in for the first 5 s, and faded out for the last 5 s, so there was no abrupt start or end to the stream. Words were selected randomly, except that no $A_i\_B_i$ pair occurred twice in succession. The speech stream was constructed from 900 word tokens, in which each word occurred approximately 100 times. Examples of the speech stream for each Experiment are shown in Table 1. Adjacent transitional probabilities were as follows: within words, $\Pr(X|A_i)$ and $\Pr(B_i|X) = .33$; between adjacent words $\Pr(A_j|B_i) = .5$ (the greater predictability across word boundaries

Table 1
Training and test samples for each of the six experiments

| Experiment | Training sample | Test sample (word) | Test sample (part-word) |
|---|---|---|---|
| 1 | ..PURAKI-BELIGA-TAFODU-PULIKI-TARADU-BEFOGA.. | PURAKI | RAKIBE |
| 2 | As 1 but with syllables assigned differently to each participant, e.g., ..BEPURA-GATADU-LIKIFO-BEGARA-GAKIDU.. | BEPURA | PURAGA |
| 3 | ..PURAGA-BELIKI-TAFOGA-PULIDU-TARAKI-BEFODU.. | PURAGA | RAGABE |
| 4 | ..RAPUKI-LIBEGA-FOTADU-LIPUKI-RATADU-FOBEGA.. | RAPUKI | PUKILI |
| 5 | ..ZEPUVO-THITASHU-FOGISA-ZETAVO… | ZEPUVO | PUVOTHI |
| 6 | ..ZEPUSHU-THITASA-FOGIVO-ZETAVO… | ZEPUSHU | PUSHUTHI |

The first column lists the experiment, the second column lists a sample of the speech stream played during training. The third and fourth columns list a sample forced choice pair at test. Hyphens individuating word boundaries are added in this table for ease of reading, but no word boundary cues were present in the training.

Table 2
Adjacent and nonadjacent transitional probabilities between syllables in the speech stream of each Experiment

| Experiment | Adjacent transitional probabilities | | Nonadjacent transitional probabilities | |
|---|---|---|---|---|
| | Within-word $\Pr(X\|A_i)$ and $\Pr(B\|X)$ | Between-word $\Pr(A_j\|B_i)$ | Within-word $\Pr(B_i\|A_i)$ | Between-word $\Pr(A_i\|X_{\text{previous}})$ $\Pr(X\|B_{\text{previous}})$ |
| 1 | .33 | .5 | 1 | .33 |
| 2 | $\Pr(X\|A_i) = .5$ hi-freq, .25 lo-freq<br>$\Pr(B_i\|X) = .5$ hi-freq, .25 lo-freq | .5 | 1 | $\Pr(A_i\|X_{\text{previous}})$: .25 if $X$ in hi-freq<br>.375 if $X$ in lo-freq<br>$\Pr(X\|B_{\text{previous}})$: .33 |
| 3 | .33 | .33 | .33 | .33 |
| 4 | $\Pr(X\|A_i) = .5$ hi-freq, .25 lo-freq<br>$\Pr(B_i\|X) = .5$ hi-freq, .25 lo-freq | .5 | 1 | $\Pr(A_i\|X_{\text{previous}})$: .25 if $X$ in hi-freq<br>.375 if $X$ in lo-freq<br>$\Pr(X\|B_{\text{previous}})$: .33 |
| 5 | .33 | .5 | 1 | .33 |
| 6 | .33 | .33 | .33 | .33 |

arises because of the constraint that no $A_i\_B_i$ pair is immediately repeated). Nonadjacent transitional probabilities within words were $\Pr(B_i|A_i) = 1$, whereas between words they were $\Pr(A_i|X_{\text{previous}}) = .33$, $\Pr(X_j\ |B_{\text{previous}}) = .33$. Table 2 summarises the transitional probabilities between syllables for every Experiment.

For the test stimuli, part-words were formed from the last syllable of one word and two syllables from the following word ($B_iA_jX$), or from the last two syllables of one word and the first syllable from the following word ($XB_iA_j$). Participants were seated in a sound-proof room and were trained and tested separately. E-prime software was used to present training and test speech, which was played through centrally positioned loudspeakers.

*Procedure*

In the training phase, participants were instructed to listen to continuous speech and try and work out the words that it contains. They then listened to the training speech. In the test phase, participants were requested to respond which of two sounds was a word in the language they had listened to. They were then played a word and a part-word separated by 500 ms, and responded by pressing 1 on a computer keyboard if the first sound was a word, or 2 if the second sound was a word. After 2 s, the next word and part-word pair were played. In half of the test trials, the word occurred first. Seven participants heard a set of test trials with one set of words first, and the other seven participants heard the other set of words first.

*Results and discussion*

The results—illustrated in Fig. 1—replicated those of Peña et al. (2002). Participants preferred words over part-words, with a mean score of 28.2 (78%), standard deviation (*SD*) of 5.4, from a possible 36, where chance performance was 18. There was a significant preference for words over part-words: $t(13) = 7.084$, $p < .0001$. In addition, participants preferred words significantly more when they had to make a decision against part-words of the form $XB_iA_j$ (e.g. [li-ki-be] mean score 15.4 from a possible 18, $SD = 2.4$) as opposed to part-words of the form $B_iA_jX\_$(e.g. [ki-be-li], mean score 12.9 from 18, $SD = 3.6$), $t(13) = 3.194$, $p < .01$.

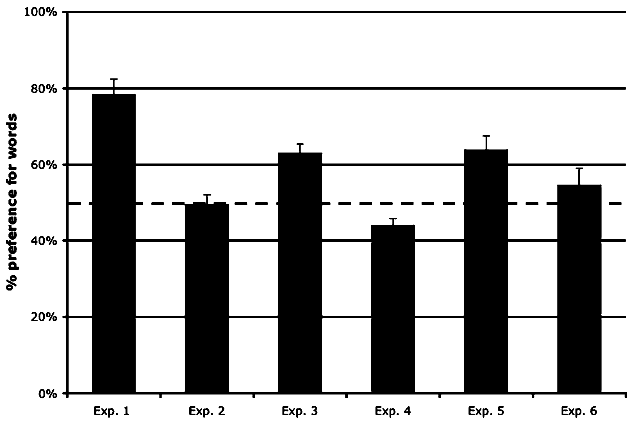


Fig. 1. Mean percentage preference for words over part-words in each segmentation task experiment. Experiment 1: original segmentation task replicating Pena et al. Experiment 2: segmentation task with randomized phonology. Experiment 3: segmentation task with no nonadjacent dependencies. Experiment 4: segmentation task with nonadjacent dependencies and continuant–plosive–plosive sound pattern. Experiment 5: learning from language with nonadjacent dependencies and with continuant–plosive–continuant phonological similarity. Experiment 6: performance with continuant–plosive–continuant phonological similarity but no nonadjacent dependencies. Error bars illustrate standard error of the mean. The dotted line represents chance level at 50%.

The replication of Peña et al.'s first experiment was a prerequisite to ensure direct comparison between the task being carried out on English and French participants. Even though the language and the synthesizer differed from those for the experiments on French, the same strong preferences for words over part-words were found in our study. Given the similarity between the distribution of plosives in English and French—plosives occur word-initially more than continuants—there remains the possibility that participants are guided in their responses by phonological properties of the language rather than by the statistical structure of the artificial language. Additional evidence for the impact of phoneme distribution comes from the significant preferences for words over $XB_iA_j$ part-words compared to words over $B_iA_jX$ part-words—the former beginning with a continuant while the latter beginning with a plosive. Decisions on forced choice pairs were harder when both word and part-word began with a plosive sound.

To test the possibility that word over part-word preferences were due to preferences for plosives in first position, we ran a control version of this study that broke the link between certain phonemes occurring in initial, medial, or final positions in Experiment 2. An additional source of preference for words over part-words was that words occurred approximately twice as frequently in the training speech corpus as part-words. We also controlled for this potential influence on the results in Experiment 2.

## Experiment 2

In Experiment 2, we controlled for preference of phonemes occurring in certain positions within words. We maintained the nonadjacent-dependency structure of the language from Experiment 1, but for each participant we randomly assigned each of the nine syllables from the first experiment to three $A_i\_B_i$ pairs and three $X$s. Each participant was therefore exposed to a training corpus that had the same $A_iXB_i$ structure as Experiment 1, but with phonemes assigned to different positions.

### *Method*

#### *Participants*

Fourteen students from the same population as in Experiment 1, but who had not participated in any other experiment reported here, participated for a £1 payment.

#### *Materials and design*

For each participant, we randomly assigned the 9 syllables from the first experiment to the $A_i$, $B_i$, and $X$ positions. Thus, each participant listened to speech with the same structure containing the nonadjacent dependencies, but with syllables assigned to different positions. For instance, the sequence $A_1XB_1$ was instantiated as [li-ki-pu] for one participant but as [be-ga-ra] for another. Once the syllables had been assigned to the positions within the words they remained in those positions for the duration of the experiment.

In addition, because part-words were half as frequent as words in the training phase in Experiment 1, we doubled the frequency of one of the words in each $A_i\_B_i$ family. Transitional probabilities were .5 between $A_i$ and the $X$ syllables in high-frequency words, and .25 in the low-frequency words; and .5 between $X$ and $B_i$ in high-frequency words, and .25 in the low-frequency $B_i$ syllables; and .33 between $B_i$ and $A_i$ syllables. The training speech was composed from concatenated words such that consecutive words were from different classes. There were approximately 150 instances of high-frequency words, and 75 of low-frequency words. The manipulation of the training stimuli equalizes the frequency of part words and words, in the test stimuli.[3]

Test items were composed of one of the lower-frequency $A_iXB_i$ words and either a $XB_iA_j$ or a $B_iA_jX$ part-word, where either $X$ and $B_i$ or $A_j$ and $X$ were from a high-frequency word. Both word and part-word sequences at test had then been heard with the same frequency during training. All 12 possible word and part-word pairs were used, and participants responded to 24 pairs, 12 of which had the word preceding the part-word, and 12 in which the part-word preceded the word.

*Procedure*

The training and testing procedure were identical to that for Experiment 1.

*Results and discussion*

The results are shown in Fig. 1. The mean response correct was 11.86 (49%), $SD = 2.3$, from a total of 24, which was not significantly different from chance, $t(13) = -.228$, $p = .824$.

The results for Experiment 2 contrast with those of Experiment 1. The key change that we made between Experiments 1 and Experiment 2 was to reassign syllables to different roles for each participant. The structure of the language was identical for both Experiment 1 and Experiment 2, however the strong preferences for words over part-words observed in Experiment 1 were completely absent from Experiment 2. That is, when the correspondence between plosives occurring word-initially and word-finally was removed there was no indication of learning the nonadjacent dependencies in the speech signal.[4] The results of Experiment 2 indicate that phonological structure has a profound effect on learning nonadjacent structure, when there is no sharing of phonological properties between first and third syllable then there is no evidence of segmentation.

Yet, is preference for phonemes in particular positions sufficient alone to result in preference for words over part-words? We tested this in Experiment 3, where the nonadjacent dependency structure was removed from the language, but the original positions of syllables from Experiment 1 were maintained.

## Experiment 3

In Experiment 3, we maintained the order of phonemes from Experiment 1, but broke the dependency between the first and the third syllable in each word. So, any first syllable was followed by any second syllable, which could be followed by any third syllable. This means that nonadjacent transitional probabilities between $A_i$ and $B_i$ syllables was reduced from 1.0 to .33.

*Method*

*Participants*

Fourteen students (who had not participated in any other experiment reported here) at the University of Warwick participated for £1.

*Materials and design*

The speech stream was constructed in the same way as for Experiment 1, except that the 9 syllables of Experiment 1 maintained their relative positions within words, but any combination of $A$, $X$, and $B$ could occur within a word. For instance, whereas in Experiment 1 the first syllable [pu] was always paired with the last syllable [ki], generating a nonadjacent frame [pu-$X$-ki], now it generated two more frames [pu-$X$-ga], and [pu-$X$-du]; likewise for the other syllables. Hence, the speech stream was comprised of 27 word types, and each word occurred approximately 33 times in the speech stream in randomized order with the constraint that no two adjacent words shared first, second, or third syllable. All transitional probabilities were now .33, including nonadjacent ones (see Table 2). Hence, there were no distributional cues for segmentation.

The test phase consisted of all 27 words, compared to part-words that were composed of either the last two syllables of the word followed by the first syllable of another word, or the last syllable of the word and the first two syllables of another word (e.g., the word $A_iXB_j$ was compared to the part-word $B_jA_iX$ or $XB_jA_i$). There were 13 comparisons between words and $XB_jA_i$ part-words, and 14 comparisons between words and $B_jA_iX$ part-words. Equal numbers could only have been achieved if a word had been repeated, or not all words had been used.

[3] A further difference between Experiments 1 and 2 is that Experiment 2 controls for frequency of words versus part-words. However, this difference is unlikely to account for the size of the difference in preferences between these experiments. Peña et al. (2002) report a frequency-controlled version of their Experiment 1 where preferences decline by only 2.8%. Hence, controlling for frequency between part-words and words at test seems to make little difference in these experiments.

[4] Peña et al. (2002) repeated their Experiment 1 by interchanging part-words for words during the training phase, and found a reduced, but significant, preference for words over part-words. However, if their control words were of the form plosive–plosive–continuant rather than continuant–plosive–plosive then their significant effect can still be attributed to a preference for words beginning with a plosive over part-words beginning with a continuant. We suggest that testing a single control is not sufficient for removing any preferences for phonemes in particular positions.

*Procedure*

The training and testing procedure was identical to that for Experiment 1.

*Results and discussion*

The results are shown in Fig. 1. Participants in this Experiment preferred words over part-words with a mean of 17.0 (63%), $SD = 2.4$, from a total of 27, which was significantly greater than chance, $t(13) = 5.416$, $p < .001$. When words were compared to part-words that began with a continuant ($A_iXB_i$ versus $XB_iA_j$), there was a significant preference for words (mean correct 9.8 out of 14, $SD = 1.8$), $t(13) = 5.643$, $p < .001$. There was no significant preference when words were compared to part-words beginning with a plosive ($A_iXB_i$ versus $B_iA_jX$), when mean correct was 7.2 out of 13, $SD = 2.1$, $t(13) = 1.261$, $p = .229$. Proportion correct scores for $A_iXB_i$ over $XB_iA_j$ part-words were greater than scores for $A_iXB_i$ over $B_iA_jX$ part-words, $t(13) = 2.456$, $p < .05$.

Though performance was significantly better than chance in Experiment 3, the overall preference for words over part-words was significantly lower than that for Experiment 1 (see Fig. 1), $t(26) = 3.313$, $p < .01$. This may be due to the language being more complicated in Experiment 3 than Experiment 1—there were 27 words compared to the 9 words of Experiment 1. Another alternative explanation is that the results of Experiment 1 indicate influences both of phonological preferences and learning of the nonadjacent dependencies. We return to this point in Experiment 5.

The results of Experiment 3 indicate that, even though there was no nonadjacent structure in the artificial language, participants still exhibited a preference for words over part-words, as defined by positions of phonemes. Experiments 1–3 suggest that preference for words over part-words is impacted by biases about word onsets that learners bring with them into the laboratory. It also seems that there is a bias against assigning word status to candidate strings that begin with a continuant sound, both in Experiment 1 and Experiment 3. In Experiment 4 below, we tested this bias by keeping the nonadjacent statistical relations of the original grammar as in Experiment 1, but having words beginning with continuant onsets and *all* part-words beginning with plosives. In line with the results of Perruchet et al. (in press), we predicted that learners would prefer part-words over words, even though this went against the nonadjacent-dependency structure of the language.

## Experiment 4

In Experiment 4, we maintained the same underlying $A_iXB_i$ structure as in Experiments 1 and 3, but used the syllables beginning with continuants as *A* syllables and the syllables beginning with plosive consonants as *X* syllables. This created words that began with continuant sounds and part-words that started with a plosive sound. If learners dispreferred continuant sounds as onsets they would prefer part-words (e.g. [be-li-pu] or [pu-ki-ra]) over words (e.g. [li-pu-ki]). In addition, this preference for part-words would indicate that phonological preferences overwhelm any effect of learning nonadjacent structure.

*Method*

*Participants*

Fourteen students (who had not participated in any other experiment reported here) at the University of Warwick participated for £1.

*Materials and design*

The grammar used was the same as for Experiment 1: $A_iXB_i$. The speech stream was composed of 9 words as in Experiment 1, but this time *A* syllables were instantiated as [li], [ra], and [fo], while *X* syllables were instantiated as [pu], [ta], [be]. The 9 words were [li-pu-ki], [li-ta-ki], [li-be-ki], [ra-pu-ga], [ra-ta-ga], [ra-be-ga]; and [fo-pu-ga], [fo-ta-ga], [fo-be-ga]. The training corpus was generated in the same way as for Experiment 1, except that frequency of words versus part-words was controlled as in Experiment 2, by doubling the frequency of one word in each of the three nonadjacent pairs.

The test phase was constructed in the same way as Experiment 1, and consisted of the lower-frequency words heard during training, compared to part-words that were composed of two syllables of one high-frequency word and one syllable of another high-frequency word. There were 24 test-pairs.

*Procedure*

The training and testing procedures were identical to those for Experiment 2.

### Results

The results are shown in Fig. 1. There was a significant preference for part-words (which began with a plosive) over words (which began with a continuant), with a mean of 10.6 (44%), $SD = 1.6$, from a total of 24, $t(13) =$ -3.33, $p < .01$. There was no significant preference for words (e.g. [li-pu-ki]) over $XB_jA_i$ part-words (e.g. [pu-ki-ra]), mean correct words chosen over part-words 5.6 out of 12, $SD = 1.5$, $t(13) = -1.104$, $p = .290$, but there was a significant preference for $B_jA_iX$ part-words (e.g. [be-li-pu]) over words, mean words chosen over part-words 5.0 out of 12, $SD = 1.0$, $t(13) = -3.606$, $p < .005$. These analyses confirm that participants rejected words in the language when they began with continuants, preferring instead to segment the speech at plosive onsets.

The results from Experiments 1 to 4 provide at the very least very weak evidence for learning of nonadjacent dependencies to drive segmentation, derived from the difference in scores between Experiments 1 and 3. However, Experiment 4 showed that nonadjacent-dependency learning could be over-ruled by preferences for words beginning with a plosive. The results of these four experiments are compatible with all three explanations of the role of phonology in segmentation. Syllable strings that begin with plosives might be preferred because of a bias for plosives at the onset of words in English and French, but also perhaps because of Perruchet et al.'s explanation in terms of the gap in speech prior to the expression of an unvoiced plosive. In addition, the third potential role of phonology—that syllables may be grouped by phonological similarity between dependent syllables—is still compatible with the results. Indeed, in Experiment 4, there was a stronger preference for part-words that had plosives in first and third position over words than part-words that had plosives in first and second position.

It is difficult to distinguish the first two accounts of the contribution of phonology to the learning task, but the third possibility can be tested in isolation. If the phonological similarity hypothesis holds, then segmentation should occur when $A_i\_B_i$ nonadjacent dependencies are instantiated as continuants and the intervening syllable is from a different category, such as a plosive. Experiment 5 tests this idea. Such an effect would indirectly account for why Newport and Aslin (2004) obtained no learning of nonadjacent syllables, because in their stimuli 2 out of 4 word frames were not phonologically similar ([pu_ra]and [lo_ki]), and all other syllables (word-initial, middle, and word-final) began with a plosive sound. This created an uninformative pattern *plosive–plosive–plosive* for 16 out of 20 words in the training set; and four words with a plosive-plosive-continuant pattern ([pi-di-ra], [pi-ku-ra], [pi-to-ra], [pi-pa-ra]) where phonological similarity is inconsistent with word boundaries.

## Experiment 5

### Method

#### Participants

Fourteen students (who had not participated in any other experiment reported here) at the University of Warwick participated for £1.

#### Materials and design

As in Experiment 1, the speech stream was composed of the 9 words respecting the $A_iXB_i$ grammar. However, syllables beginning with a continuant were assigned to the *A* and *B* positions, and syllables in the *X* position had plosive onsets. The three *A_B* pairs were [ze_vo], [thi_shu], [fo_sa], and the *X* syllables were [pu], [ta], [gi]. The continuants were chosen such that they were all fricatives, and pairs had different places of articulation. The training corpus was generated in the same way as for Experiment 1. In particular, as in Experiment 1, part-word and word frequencies in test are not equalized in the training stimuli. As noted in Footnote 3, this factor does not appear to make a substantial difference to the results.

The test phase was constructed in the same way as Experiment 1, and consisted of each word compared to $XB_iA_j$ and $B_iA_jX$ part-words. There were 36 test pairs.

#### Procedure

The training and testing procedures were identical to those for Experiment 1.

#### Results and discussion

The results are shown in Fig. 1. Participants preferred words over part-words with mean 22.6 times out of a maximum 36 (62.7%), $SD = 5.2$, which was significantly above chance, $t(13) = 3.318$, $p < .01$. Preference for words over $B_iA_jX$ part-words was significantly greater than chance, 12.3 from 18, $SD = 3.0$, $t(13) = 4.101$, $p < .001$, and preference for words over $XB_iA_j$ part-words was marginally significantly greater than chance, 10.3 from 18, $SD = 2.6$, $t(13) = 1.840$, $p = .089$. Performance for $B_iA_jX$ part-words was better than for $XB_iA_j$ part-words, $t(13) = 3.321$, $p < .01$, indicating that continuant–continuant–plosive patterns were preferred less than plosive–continuant–continuant part-words.

These results suggest that segmentation can occur on the basis of nonadjacent dependencies but only under certain circumstances where there is phonological similarity between the first and the third syllable. In Experiment 2, when there was no sharing of phonology between syllables, then nonadjacent dependencies were not accessed for the segmentation task. The results suggest that access to computing statistical nonadjacent dependencies requires that the dependencies are phonologically similar.

However, performance in Experiment 5 is significantly worse than Experiment 1, which combined word-initial plosives and nonadjacent structure, $t(26) = 1.662$, $p < .05$. This result suggests that there is some combination of preference for plosives in first position and sharing of first and third syllable that contributes to segmentation performance. However, it remains a possibility that the continuant–plosive–continuant phonological structure is sufficient on its own to drive preferences for words over part-words, and that the nonadjacent structure is irrelevant to performance on the task. To test this we ran Experiment 6 below, which has no nonadjacent structure but maintains the phonological property sharing between first and third syllable.

## Experiment 6

### Method

#### Participants

Fourteen students (who had not participated in any other experiment reported here) at the University of Warwick participated for £1.

#### Materials and design

Materials were created using the continuant-plosive-continuant pattern for syllable onsets in words, as used in Experiment 5. However, the dependency between particular first and third syllables was removed. Thus, in first position were the syllables [ze], [thi], [fo], in second position were [pu], [ta], [gi], and in third position [vo], [shu], [sa]. The training corpus was created in the same way as Experiment 3, with no syllable repeated in an adjacent word. Similarly, the test phase was constructed in the same way as Experiment 3, and consisted of all 27 words, compared to part-words that were composed of either the last two syllables of the word followed by the first syllable of another word, or a last syllable from another word and the first two syllables of the word being tested.

#### Procedure

The training and testing procedures were identical to those for Experiment 3.

### Results and discussion

The results are shown in Fig. 1. Participants preferred words to part-words with mean 14.7 out of 27 (54%), $SD = 4.5$. This was not significantly different from chance, $t(13) = 1.010$, $p = .331$. Preference for words was not observed when compared against $XB_iA_j$ part-words, mean 7.1 out of 13, $SD = 3.1$, $t(13) = .684$, $p = .506$. Nor was there a preference for words over $B_iA_jX$ part-words, mean 7.6 out of 14, $SD = 2.3$, $t(13) = 1.042$, $p = .316$. Unlike for Experiment 5, there was no significant difference for proportion correct on $XB_iA_j$ compared to $B_iA_jX$ part-words, $t(13) = -.031$, $p = .976$. Participants were not able to distinguish words from part-words when the phonological pattern involved continuants in first and third position but when there was no nonadjacent dependency structure. This pattern of results contrasts with the effect seen in Experiment 3, where a preference for plosives in first and third position was observed after training, even when there was no nonadjacent structure in the language.

## General discussion

In this paper, we have investigated the potential role of phonological processing in an ALL task. We found that, in the processing of nonadjacent dependencies for use in a segmentation task, the phonological properties of the dependent syllables within words were critical for learning to take place. Segmentation can take place on the basis of *adjacent* dependencies in sequences of syllables (Aslin, Saffran, & Newport, 1996), tones (Saffran et al., 1999), or visual items (Kirkham et al., 2002). Yet, finding nonadjacent dependency learning has proved elusive in several studies (e.g., Morgan & Newport, 1981; Newport & Aslin, 2004), and the data presented in this study helps to define the conditions under which such learning is possible.

Table 3 provides a summary of the six experimental designs. Experiment 1 replicated Peña et al.'s (2002) study in English: segmentation could proceed on the basis of nonadjacent dependencies. However, Experiment 2, which maintained the nonadjacent structure but altered the order of syllables from the first experiment, did not find evidence of learning, and indicated that nonadjacent dependencies cannot be used for segmentation under all circumstances. Indeed, Experiment 3 showed that segmentation could proceed on the basis of order of syllables only, though performance was not as good as when syllable order and nonadjacent structure was in place as in Experiment 1. Experiment 4 further showed that learning nonadjacent structure could be over-ruled by particular orders of syllables. These first four experiments provide evidence that plosives in initial position in words have a large contribution toward segmentation performance in Peña et al.'s experiments. These studies supported the claims that ALL performance was influenced either by phonotactic biases, or by latent prefer-

Table 3
Summary of the design of the experiments

| Experiment | Syllable positions | Nonadjacent structure | *p*-value |
|---|---|---|---|
| 1 | P–C–P | Y | <.0001 |
| 2 | Randomised | Y | ns |
| 3 | P–C–P | N | <.001 |
| 4 | C–P–P | Y | < .01* |
| 5 | C–P–C | Y | < .01 |
| 6 | C–P–C | N | ns |

The first column lists the experiment, the second column lists the experiment number in Peña et al.'s study. Syllable positions indicate the order of syllables within words in the Experiment, P–C–P: plosive–continuant–plosive; C–P–P: continuant–plosive–plosive; C–P–C: continuant–plosive–continuant. The Structure column indicates whether the language contained nonadjacent dependencies or not, and the effect indicates the statistical result. The asterisk (*) indicates a significant preference for part-words.

ence for plosives in first position due to their phonetic properties, or by a combination of the two.

To further assess the contribution of phonotactic biases in artificial speech stimuli, we conducted two measures of the distribution of phonemes in French and English. The first counted the percentage of words in a corpus that began with each consonant in the onset of the syllables in Peña et al.'s experiments. This was to measure the extent to which certain phonemes were more likely than others to begin words in French and in English. The second measure was the conditional probability that each phoneme was the onset of a word. This measure determines whether each phoneme in syllable onset position is an informative cue for the beginning of a word. Some phonemes may occur more frequently than others in all positions within the word, but this would not be reflected in the first measure that assesses the percentage of words with that phoneme as the onset.

For French, we used the LEXIQUE corpus (New, Pallier, Ferrand, & Matos, 2001) and we used the CELEX corpus for English (Baayen, Piepenbrock, & Gulikers, 1995). The results are shown in Table 4, for token frequencies of words, as well as type frequencies in parentheses. For the proportion of words beginning with each phoneme, more words in French begin with initial (/p/, /b/, /t/) and final (/k/, /g/, /d/) phonemes than with medial (/R/, /f/, /l/) phonemes (column 3), where initial, medial, and final refer to positions in the syllables that make up the artificial grammar used by Peña et al. (2002). In English more words begin with the initial phonemes than with medial or final phonemes (column 5). Using the second measure, in French the initial and final phonemes were more likely than medial phonemes to begin words (column 4). In English, the initial phonemes were more likely than medial and final phonemes to begin words.

We tested the consequence of forming a preference for words over part-words based only on the likelihood of the initial phoneme in word-initial position. If an initial phoneme occurs more often initially than a medial phoneme, then the word beginning with the initial phoneme is taken to be preferred over the part-word beginning with the medial phoneme. For instance, *puraki* would be preferred over *rakibe* as /p/ occurs more often initially than /R/ in French. From the 36 tests of word/part-word in the segmentation experiment in Peña et al.'s study, in the token frequency analysis in French 18 cases (50.0%) produced a preference for a word over a part-word. In English, 32 out of 36 words (88.9%) were preferred over part-words. For the second measure, for French again 18 words (50%) would be preferred over part-words, and 24 words (66.7%) in English would be preferred. For the type frequency analysis, in French the first measure resulted in preference for 20 words over part-words and 20 for the second measure. In English, there was a preference for 22 words over part-words for the first measure and 24 for the second measure.

A more conservative decision rule for selection of the preferred word is based on a Luce choice ratio (Luce, 1963) in which the probability of selecting the sequence beginning with say, /b/ rather than /R/ is

$$\Pr(/b/) = \frac{\text{freq}(/b/)}{\text{freq}(/b/) + \text{freq}(/R/)}.$$

For token frequencies, for the first measure in French this choice ratio results in 46.6% preference for words, and 60.1% in English. The Luce choice ratio for the second measure results in preference for 48.8% of words in French and 54.8% in English. For the type frequency analysis, in French there is a 52.5% preference for words in the first measure, and 51.9% for the

Table 4
Percentage of words beginning with each consonant for syllables in initial/medial/final word position in Peña et al.'s studies, and conditional probabilities of consonants beginning a word in French and English

| Position in $A_iXB_i$ words | Phoneme | French | | English | |
|---|---|---|---|---|---|
| | | % words beginning with phoneme | Word-onset probability | % words beginning with phoneme | Word-onset probability |
| Initial | /p/ | 8.2 (8.7) | .67 (.40) | 3.1 (8.4) | .44 (.34) |
| | /b/ | 1.7 (4.7) | .44 (.35) | 4.4 (6.8) | .65 (.44) |
| | /t/ | 3.5 (4.7) | .21 (.11) | 4.9 (4.6) | .20 (.09) |
| | | Total: 13.4 (16.9) | Overall: .44 (.32) | Total: 12.4 (19.8) | Overall: .43 (.29) |
| Medial | French /R/, English /ɹ/ | 2.9 (9.1) | .11 (.13) | 2.2 (6.7) | .20 (.21) |
| | /f/ | 2.7 (4.4) | .59 (.36) | 4.4 (5.2) | .64 (.36) |
| | /l/ | 9.1 (2.4) | .43 (.08) | 2.3 (3.7) | .19 (.12) |
| | | Total: 14.7 (15.9) | Overall: .37 (.19) | Total: 8.9 (15.6) | Overall: .34 (.23) |
| Final | /k/ | 7.2 (9.2) | .55 (.32) | 3.7 (9.4) | .35 (.30) |
| | /g/ | 0.9 (2.5) | .40 (.29) | 1.5 (3.0) | .53 (.35) |
| | /d/ | 13.0 (8.2) | .79 (.42) | 3.0 (6.7) | .19 (.16) |
| | | Total: 21.1 (19.9) | Overall: .58 (.34) | Total: 8.2 (19.1) | Overall: .36 (.27) |

Token frequency analysis, with type frequency analysis in parentheses.

second measure, and in English for the first measure there is a 54.3% preference for words and 50.5% for the second measure.

The corpus analyses provide mixed evidence of bias in terms of the distribution of the phonemes beginning syllables in Peña et al.'s stimuli. The preferences we found in the corpus analyses were very weak, resulting in no preference in French for words over part-words in the token frequency analysis, and a slight preference for words in the type frequency analysis. The biases for initial phonemes beginning words in English were stronger, but still much less than the observed preferences in the experimental results of Peña et al. and ours, and this suggests that additional contributions to preference result from other sources, such as phonological similarity between nonadjacent dependencies.

Experiment 5 indicated that nonadjacent dependencies could be learned if supported by a correspondence between dependent syllables in terms of phonological properties (Morgan & Newport, 1981), even though there were no plosives in first position—words in this experiment began and ended with continuants in the onset of the syllable. This result also ruled out the hypothesis by Perruchet et al. (in press) that the speech synthesizer alone was responsible for inducing segmentation at word-boundaries. Even if the Festival speech-synthesizer, as well as the French MBROLA synthesizer, produced unvoiced plosives at the beginning of words preceded by a silence before the onset, the results of Experiment 5 show that successful segmentation can occur with continuants as word onsets, outweighing this synthesizer bias.

Finally, Experiment 6 showed that phonological similarity is not sufficient on its own to drive segmentation, as performance was at chance level when there was no nonadjacent structure but only continuant–plosive–continuant phonological structure. Table 5 summarises the design of each Experiment and the resultant effects.

The phonotactic bias and the phonological similarity bias appear to interact in an additive way to segmentation performance (see Table 5). Experiment 1 with nonadjacent structure, plosive onset, and phonological similarity between dependent syllables resulted in the highest preference for words over part-words (28% above chance). Experiment 2, with plosive onset but no nonadjacent structure, and Experiment 5, with phonologically similar nonadjacent dependencies but with continuant onsets, resulted in segmentation performance at equivalent levels. Each was approximately half the level above chance that was found for Experiment 1. Absence of either plosive onset or phonological similarity was sufficient for performance to return to chance levels (Experiment 6), or even below if phonological preferences were violated, as in Experiment 4.

We performed an ANOVA on the combined results of Experiments 1, 3, 5, and 6, with presence/absence of nonadjacent structure and presence/absence of plosive as word-onset as factors. There was a main effect of structure, $F(1,52) = 9.915$, $p < .005$, with presence of structure resulting in better performance. There was also a main effect of presence of plosive as first sound, $F(1,52) = 10.365$, $p < .005$, with better performance when plosives were initial. There was no significant interaction between structure and initial plosive, $F < 1$.

The influence of phonological properties on learning language structure is not entirely surprising, given that there is a strong correspondence between phonological properties and phonotactic and grammatical structure in natural language. There is coherence among grammatical categories in terms of phonological properties (Kelly, 1992), which may well be a crucial contributor to the learnability of such grammatical structure (Braine, 1987). Monaghan, Chater, and Christiansen (in press) found that category learning in an ALL was significantly improved when words in the same category shared phonological information, for instance. Similarly, Brooks, Braine, Catalano, Brody, and Sudhalter (1993) found that learning of a gender-like classification in an ALL was only possible when words shared phonological properties.

The results from Peña et al.'s (2002) study indicate that segmentation can take place on the basis of nonadjacent transitional probabilities, but we have shown this only occurs when there is phonological similarity between the dependent syllables within words. When the contribution of phonology is removed no learning takes place, as in Newport and Aslin's (2004) studies, where a similar artificial grammar was employed. Peña et al.'s

Table 5
The additive contribution of phonological and nonadjacent dependency structure in Experiments 1–6

| Experiment | Statistical structure present? | Words have plosive onset? | Phonological pattern brackets statistical structure? | Percentage performance above chance (%) |
|---|---|---|---|---|
| 1 | Yes | Yes | Yes | 28 |
| 2 | Yes | No | No | 0 |
| 3 | No | Yes | No | 13 |
| 4 | Yes | No | No | −4 |
| 5 | Yes | No | Yes | 14 |
| 6 | No | No | No | 4 (ns) |

(2002) conclusions regarding separable processing for segmentation and for learning to generalize the structure of the language are therefore premature, given that segmentation only occurs under certain conditions precipitated by phonological properties of the speech. The conditions under which generalization may also occur require additional testing. The results of these experiments indicate that phonological factors in ALL experiments need careful experimental control, given the sensitivity of learners to both the phonological and distributional structure of artificial language learning materials.

## References


Aslin, R. N., Saffran, J. R., & Newport, E. L. (1996). Computation of conditional probability statistics by 8-month old infants. *Psychological Science, 9*, 321–324.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database (CD-ROM)*. Philadelphia, PA: University of Pennsylvania.

Black, A. W., Taylor, P., & Caley, R. (1990). The Festival Speech Synthesis System. Available from http://www.cstr.ed.ac.uk/projects/festival.html, Centre for Speech Technology Research (CSTR), University of Edinburgh, Edinburgh, UK.

Braine, M. D. S. (1987). What is learned in acquiring word classes: A step toward an acquisition theory. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 65–87). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brooks, P. B., Braine, M. D. S., Catalano, L., Brody, R. E., & Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language, 32*, 79–95.

Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition, 61*, 93–125.

Friederici, A. D., & Wessels, J. M. I. (1993). Phonotactic knowledge and its use in infant speech perception. *Perception & Psychophysics, 54*, 287–295.

Gómez, R. (2002). Variability and detection of invariant structure. *Psychological Science, 13*, 431–436.

Hahn, U., & Chater, N. (1998). Similarity and rules: distinct. exhaustive? empirically distinguishable? *Cognition, 65*, 197–230.

Jusczyk, P. (1999a). How infants begin to extract words from speech. *Trends in Cognitive Sciences, 3*, 323–327.

Jusczyk, P. W. (1999b). How infants begin to extract words from speech. *Trends in Cognitive Sciences, 3*, 323–328.

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*, 1–20.

Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99*, 349–364.

Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language, 37*, 295–311.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence of a domain general learning mechanism. *Cognition, 83*, B35–B42.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*. New York: Wiley.

Marcus, G. F. (1999). Do infants learn grammar with algebra or statistics? *Science, 284*, 436–437.

Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition, 78*, 91–121.

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*, 465–494.

McClelland, J. L., & Plaut, D. C. (1999). Does generalization in infant learning implicate abstract algebra-like rules? *Trends in Cognitive Sciences, 3*(5), 165–201.

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language, 39*, 21–46.

Monaghan, P., Chater, N., & Christiansen, M. H. (in press). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition*.

Morgan, J. L., & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior, 20*, 67–85.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique, 101*, 447–462.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of nonadjacent dependencies. *Cognitive Psychology, 48*, 127–162.

Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology, 34*, 191–243.

Peña, M., Bonatti, L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*, 604–607.

Perruchet, P., Tyler, M. D., Galland, N., & Peereman R. (in press). Learning nonadjacent dependencies: No need for algebraic-like computations. *Journal of Experimental Psychology: General*.

Quené, H. (1992). Durational cues for word segmentation in Dutch. *Journal of Phonetics, 20*(3), 331–350.

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

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*, 27–52.

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.

Seidenberg, M. S., MacDonald, M. C., & Saffran, J. R. (2002). Does grammar start where statistics stop? *Science, 298*, 553–554.