

RESEARCH ON SPOKEN LANGUAGE PROCESSING

Progress Report No. 28 (2007)

Indiana University

**Implementing and Testing Theories of Linguistic Constituency I:
English Syllable Structure¹**

Vsevolod Kapatsinski

*Speech Research Laboratory
Department of Psychological and Brain Sciences
Indiana University
Bloomington, Indiana 47405*

¹ I would like to thank the NIH for financial support through Training Grant DC-00012 and Research Grant DC-00111 to David Pisoni. Many thanks to Luis Hernandez for his help in creating the experimental program and to Adam Buchwald for helpful comments on an earlier draft of this paper.

Implementing and Testing Theories of Linguistic Constituency I: English Syllable Structure

Abstract. This paper proposes and tests an experimental method to evaluate models of linguistic constituency, including: 1) Connections within constituents are stronger than connections spanning constituent boundaries, 2) A constituent is more likely to be parsed out of the signal than a non-constituent (i.e., constituents are processing units), and 3) Both constituents and non-constituents are units, with constituents simply having higher frequency than non-constituents. The method, XOR learning, is designed to distinguish between associability of a whole and associability of its parts. Subjects learn to associate the whole with a different response than the response both of its parts have been associated with. We apply the method to the onset-rime organization of English CVC syllables with a lax vowel, showing that native English speakers can learn rime-affix associations but not body-affix associations. This difference in associability between bodies and rimes is observed in the absence of associability differences between onsets and codas, the only parts that bodies and rimes do not share. Competing theories of linguistic constituency are implemented in a Hebbian framework where parts of the syllable extracted from the signal become associated with affixes they co-occur with. Assuming automatic phonemic categorization, experimental results are explained only by models that assume that rimes and bodies differ in the level of activation they have during training. Applications of the experimental method and its variants to linguistic constituency in other domains are discussed.

Introduction

Theories of Constituency

This paper introduces a method to distinguish between different theories of linguistic constituency. Thus, the question we would like to address is: what are constituents? What does it mean to say that in an English syllable consisting of an onset, a nucleus, and a coda, the nucleus forms a constituent with the coda and not with the onset?

The traditional answer to this question in linguistic theory has been that the rime (nucleus+coda) is allocated a node in the tree structure while the body (onset+nucleus) is not (e.g., Fudge, 1987; Selkirk, 1982). A tree structure is a type of a network and, like in any network, it consists of nodes connected by links. By definition, then, a node is something that can be connected to/associated with something else. Thus, in the traditional view of linguistic constituency, constituents can be associated with other units, i.e., **constituents are associable, while non-constituents are not**. Thus, under this view, if the rime is a constituent while the body is not, rime-affix associations should be learnable while body-affix associations should not be.

In order to associate a unit X with another unit Y, the two units must be extracted from the signal. Thus, things that are associable must be extracted from the signal. In other words, something that is associable must be a **processing unit**. Under the traditional view of constituency, then, **constituents are processing units** (cf. Cutler et al., 2001; Mehler, 1981). That is, at the very least, if the rime is a constituent and the body is not, the rime should be more likely to be extracted from the acoustic signal than the body is.

An alternative to the tree-structural view of constituency is the dependency-based view, applied to syllabic constituency by Vennemann (1988) and Anderson and colleagues (e.g., Anderson & Ewen, 1987). Under this view, neither constituents nor non-constituents are allocated nodes. Rather, **connections between parts of a constituent are stronger than connections that cross constituent boundaries**.

Under this view, to say that the rime is a constituent while the body is not means to say that the nucleus is connected to the coda more strongly than to the onset. The dependency-based view does not straightforwardly predict a difference in associability between constituents and non-constituents. Any such difference would be an epiphenomenon, deriving from differences in associability between parts that the constituent and the non-constituent do not share. Thus, in the case of the body, the rime would be expected to be more associable than the body if and only if the coda is more associable than the onset.

Finally, processing units may differ in how associable they are, depending on factors like frequency and the cumulative strength of associations they already have (e.g., Kamin, 1969; Moder, 1992). That is, nodes may differ in associability. The associative learning literature indicates that **frequent stimuli are harder to associate than infrequent stimuli** (the phenomenon known as pre-exposure, or desensitization effects, see Hall, 2003, for review). In the linguistic literature, Bybee and Brewer (1980) and Moder (1992) have argued that frequent words have weaker connections to similar words than infrequent ones. Thus, if a rime is more frequent than a body, the rime may be expected to be less associable than the body. Given this potential influence on associability, rimes and bodies may be equally likely to be parsed out of the signal (i.e., constituents and non-constituents may be equally **salient**) and still differ in associability. Equal salience is proposed within full-listing models in which **all possible segment strings (up to a particular length) are parsed out of the signal** (Skousen, 1989; see also Bod 1998 for an analogous approach to syntax in which all possible subtrees are parsed out).

In this paper, we will compare associability of rimes to associability of bodies and then implement the various theories of syllable structure in a common framework to see which can account for the experimental data.

XOR Learning

As discussed in the previous section, the tree-structural view of constituency predicts that rimes should be more associable than bodies in English while the dependency-based view claims that any such differences should be attributable to differences in associability between onsets and codas. Thus, to distinguish between the two alternatives, it would be helpful to have a way to train subjects on body-affix or rime-affix associations while controlling segment-affix associability. Such a way is provided by XOR learning.

A classic **XOR** (exclusive-or) distribution is one in which stimuli containing either A or B are classified as being of type X, while stimuli containing both A and B and stimuli containing neither A nor B are classified as being of type Y ($A, \text{not } B \rightarrow X$; $\text{not } A, B \rightarrow X$; $A, B \rightarrow Y$; $\text{not } A, \text{not } B \rightarrow Y$). This distribution has been most famously used by Minsky and Papert (1969) to argue against two-layer connectionist networks as models of cognition. Minsky and Papert showed that such purely distributed models cannot represent the XOR relation because if A is associated with X and B is associated with X, AB necessarily is, since there is no representation for the complex unit AB. Algorithms for learning in three-layer networks were later introduced in part to handle XOR. In such networks, some node(s) in the hidden layer, which intervenes between the input and output, are activated when AB is presented and not when A or B is presented in isolation, allowing the network to respond to AB in a different way than it responds to A or B.

In the present experiment, we used a modified version of the XOR distribution in which A is associated with X, B is associated with X, while C is associated with Y as is D but AB is associated with Y while CD is associated with X, as shown in figure 2. This distribution was chosen to avoid asking participants to form a category defined solely in negative terms (e.g., syllables in which the coda is not B and the nucleus is not A). In the distribution in Figure 2, all the dependencies involve specific segments or bigrams present in the stimuli.

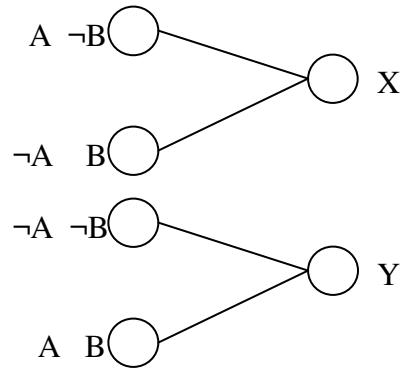


Figure 1. A classical XOR task where one response (X) is required when presented with either A or B but not both. If neither A nor B is presented or both A and B are presented, a different response, Y, is required. Here A might be an onset, and B might be a nucleus or A might be a nucleus with B being a coda.

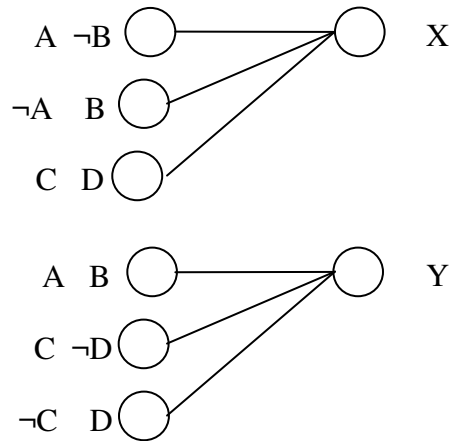


Figure 2. The task used in the present experiment. One response is associated with A in the presence of something other than B, with B in the presence of something other than A, and with CD. The other response is associated with C in the presence of something other than D, with D in the presence of something other than C, and with AB. Here A and C may be onsets with B and D being nuclei and AB and CD being bodies or A and C may be nuclei, B and D codas, and AB and CD rimes. AD and CB are never presented.

In the experiment presented here, X and Y are affixes. They can either precede or follow the stem. A, B, C, and D are individual segments. For some subjects, AB and CD are bodies, while for others they are rimes. We are interested, then, in how easily subjects learn AB-Y and CD-X associations when AB and CD are rimes compared to the case in which AB and CD are bodies.

The experiment is divided into two stages. In the first stage, AB and CD are never presented and subjects thus learn that A is paired with X, as is B, and that C and D are paired with Y. Thus, subjects learn segment-affix associations. By looking at subjects' accuracy with novel syllables containing familiar onsets or codas after the first stage, we can compare associability of onsets to associability of codas.

Since bodies and rimes share nuclei, all subjects learn nucleus-affix associations during stage I. Therefore, the results of stage I allow us to assess between-subject differences in learning rate. Thus, we can ensure that subjects who are assigned to learn rime-affix associations are not simply better learners than those that are assigned to learn bodies.

In the second stage, AB and CD tokens are introduced and subjects learn that AB is (surprisingly) associated with Y while CD is paired with X. Subjects are then asked to predict the affixes of unfamiliar syllables containing the now familiar AB and CD. The results of Stage II training indicate how easy it is to learn rime-affix vs. body-affix associations. Since we know how easy it is to learn coda affix vs. onset-affix associations from the results of Stage I, we can determine whether differences in associability between bodies and rimes can be explained by differences in associability between onsets and codas as predicted by dependency-based models of syllable structure.

Evidence for Syllabic Constituency

In order to investigate the relationship between constituency and associability, we need a case where constituency is uncontroversial. Such a case is provided by English CVC syllables with lax vowels. There are a number of reasons to believe that such syllables have an onset-rime structure (/C/ /V C/) and not body-coda structure (/C V/ /C/). That is, the vowel goes with the following consonant and not the preceding one. Furthermore, there is some evidence for the existence of the parts of syllabic constituents, the segments.

Strong evidence has been provided for the involvement of syllable structure in visual word recognition. In priming experiments reported by Ferrand et al. (1996) for French, Carreiras and Perea (2002) and Alvarez et al. (2004) for Spanish, and Ashby and Rayner (2004) for English the target word started with either a CVC or a CV syllable. The prime was presented visually and had either the form CV**** or CVC***. It was presented so fast that the subjects did not consciously notice its identity. All segments of the prime were present in the target. For instance, the primes may be *pa***** and *pas**** and the targets may be *pasivo* and *pastor* (Carreiras and Perea, 2002).²

If all that mattered for phonological priming were the number of segments or letters shared between the prime and the target or the duration of the shared part, we would expect CVC primes to produce more priming than CV primes for both types of targets. However, both studies showed a reliable interaction: while CVC primes produced more priming than CV primes for CVCCVC targets (*pas**** primed *pastor* more than *pa***** did), CV primes produced more priming than CVC primes for CVCVCV targets (*pa***** primed *pasivo* more than *pas**** did). These results follow directly from the syllable structure of the targets: *pas* shares a syllable with *pastor* but not with *pasivo*, while *pa* shares a syllable with *pasivo* but not with *pastor*. These studies provide convincing evidence that subjects are sensitive to the syllable structure of the word, although ambiguity remains regarding whether the sharing of syllables or syllabic constituents is at issue. While *pas* shares an onset and a rime with *pastor*, it only shares an onset with *pasivo*. Nonetheless, these data constrain possible models of syllable structure in that there must be nodes for either syllables or syllabic constituents, i.e., it is not sufficient to have only segment units.

Evidence for the onset-rime structure is provided by the fact that categorical co-occurrence restrictions involve VC's and not CV's in English (e.g., Fudge, 1987; Selkirk, 1982). In particular, lax vowels in English require a consonant to follow them while they do not require a preceding consonant.

² Alvarez et al. (2004) have addressed the concern that the effect is orthographic, rather than phonological in nature: the syllable priming effect did not diminish when the spelling of the first syllable was not shared by the prime and the target (*vi.rel-vi.rus* vs. *bi.rel-vi.rus*).

This also means that in a syllable with a lax vowel, the body is not a possible word while the rime is, which contributes to the constituency difference between the body and the rime in such syllables.

In addition, Kessler and Treiman (1997) and Lee (2006) have shown that there are statistical reasons for grouping vowels with codas in English: given the vowel, the coda is somewhat predictable while the vowel is not predictable given the onset. Treiman et al. (2000) found that wordlikeness judgments for nonsense CVC's are affected by the probabilistic constraints on vowel-coda co-occurrence.

Treiman (1983, 1986) and Derwing (1987) found that when English speakers are asked to use the beginning of the first (C)CVC(C) word and the end of the second (C)CVC(C) word to form a new word, they use the onset of the first word and the rime of the second and not the body of the first and the coda of the second. Treiman et al. (2000) found that the tendency is a little stronger with high-frequency rimes than with low-frequency rimes (although above 90% in both cases).

Treiman and Danis (1988) showed that when subjects are asked to recall a long list of words, they tend to make errors that are novel recombinations of previously presented onsets and rimes and not bodies and codas. In addition, Nelson and Nelson (1970), Vitz and Winkler (1973), Derwing and Nearey (1986), Bendrien (1992), and Yoon and Derwing (2001) found that CVC words sharing rimes are perceived to be more similar than words sharing bodies by English speakers in sound similarity judgment tasks. However, Geudens et al. (2005) have called the relevance of these results for testing syllabic constituency into question by showing that Dutch speakers judge syllables sharing rimes to be more similar than syllables sharing bodies yet show no preference for recombining onsets and rimes as opposed to bodies and codas in a serial recall task.

Lee (2006) has shown that the statistics favor the body in Korean, which explains why Korean speakers, unlike English speakers, tend to produce body-coda recombinations in serial recall when presented with syllables in which the nucleus co-occurs with the coda as strongly as it co-occurs with the onset. Furthermore, if English speakers are asked to memorize atypical syllables in which the nucleus co-occurs with the onset more than with the coda, they too tend to produce body-coda recombinations. Finally, Korean speakers presented with syllables in which the nucleus co-occurs with the coda more than with the onset, which are not typical in Korean, tend to produce onset-rime recombinations. These results constrain models of syllabic constituency in that the difference between bodies and rimes in English cannot be due to the fact that the beginning of the rime follows the beginning of the body within the syllable but rather must be due to the statistics of between-segment co-occurrence. In this paper, we will only use syllables in which statistics of co-occurrence favor the onset-rime division.

A number of studies have provided support for the psychological reality of segments. Vitz and Winkler (1973) found that sound similarity judgments for a pair of words had a correlation of 0.9 with number of mismatched segments. Kapatsinski (2006) observed that the mean likelihood of interrupting a word before replacing it in spontaneous speech production is very strongly correlated with log number of segments in the word ($r^2=.991$). Stemberger (1983) and Jaeger (2005), among others, find that most speech errors involve substitutions of single segments. Boothroyd and Nitttrouer (1988), Nearey (1990, 2003), Benki (2003), and Felty (2007) found that accuracy of identification of nonsense syllables in noise is highly accurately approximated by a linear combination of average identification accuracies of the component segments. Hockema (2006) found that word segmentation based on segment transitions would be highly successful in English. Finally, some evidence for units smaller than syllabic constituents is provided by the fact that not all rimes are equally acceptable in English, e.g., */aʊp/.

Given the existence of evidence for both the onset-rime division of the types of syllables used in the present study and the status of segments as processing units, we can ask whether the rime is a processing unit as well or if differences in between-segment connection strength are sufficient to account for the structure of English syllables. Before getting into the main part of the paper, it is important to note that the present study concerns the nature of syllabic constituency in English syllables with lax vowels

that are, furthermore, morpheme-initial. There are a number of factors that make the rime a better constituent than the body in such cases which may not distinguish other types of syllabic constituents from other types of non-constituents. For instance, bodies that end in a lax vowel is not a possible word, while bodies that end in tense vowels are possible words. In addition, Davis (1989) has argued, based on speech error evidence, that the word-initial onset is particularly poorly integrated into the rest of the word. Thus, the present paper intends to show that in a case in which there is an extremely clear constituency difference between the body and the rime, the rime is much more associable than the body. However, this may not be the case with other types of rimes and other types of bodies. The nature of a proposed constituency difference is an empirical question that needs to be investigated separately in each case. There is no a priori reason to believe that the same model should be used to model constituency in such disparate domains as syntax and phonology or even in different contexts within phonology. This paper provides one way of selecting an appropriate model for a particular type of linguistic constituent.

The Experiment

Methods

The Paradigm. In our experiment, we wanted to dissociate constituent associability from associability of the component segments. Thus, we needed the subjects to learn that a whole is associated with a different response than either of its parts. In other, words, if AB is associated with Y, then A is associated with X and B is associated with X.

Native English speakers were randomly assigned to the four experimental groups shown in Table 1. As can be seen from the table all subjects were exposed to co-occurrences between the vowels /æ/ and /ʌ/ and the affixes /mɪn/ and /num/. However, /mɪn/ and /num/ came after the stem (were suffixes) for groups II and IV but they came before the stem (were prefixes) for groups I and III. In addition to being exposed to vowel-affix correlations, subjects in groups I and II were exposed to rime-affix and coda-affix correlations while subjects in groups III and IV were exposed to body-affix and onset-affix correlations.

Group	Associate	Part relations	Whole relations
I	Rimes & prefixes	num- CæC num-CVʃ mɪn-CʌC mɪn- CVg	mɪn- Cæʃ num-Cʌg
II	Rimes & suffixes	CæC-num CVʃ-num CʌC-mɪn CVg-mɪn	Cæʃ-mɪn Cʌg-num
III	Bodies & prefixes	num-CæC num-ʃVC mɪn- CʌC mɪn- gVC	mɪn- ʃæC num-gʌC
IV	Bodies & suffixes	CæC-num ʃVC-num CʌC-mɪn gVC-mɪn	ʃæC-mɪn gʌC-num

Table 1. The experimental design: what do subjects have to learn?

Two things to note about the design are that each subject has to learn two XOR distributions as in Figure 2, and that the temporal relation between the affix and the stem (prefixation vs. suffixation) is manipulated independently of whether the to-be-associated part of the stem is a constituent.

If the subjects were exposed to a single XOR distribution (e.g., num-CæC, num-CVʃ, but mɪn-Cæʃ) they would not need to infer anything about associations of the parts (/æ/ and /ʃ/ in the example) since all regularities can be defined in terms that only involve /æʃ/: the presence of /æʃ/ indicates the presence of /num/ while its absence indicates the presence of /mɪn/.

The experimental design shown in Table 1 allows us to control for locality effects and possible differences in associability between suffixes and prefixes. Thus, if prefixes are more associable than suffixes, associations involving prefixes should be easier to learn than those involving suffixes, regardless of whether the rime of the stem or the body of the stem is involved. Alternatively, association between adjacent parts of the speech stream may be easier to learn than an association between non-adjacent parts. Such a result has been obtained in statistical learning of segment and syllable co-occurrences (Newport & Aslin, 2004; Bonatti et al., 2005). If this is the case, body-prefix and rime-suffix associations should be more learnable than body-suffix and rime-prefix associations.

The Sequence of Training and Testing Stages. Participants were first trained on vowel-affix and consonant-affix co-occurrence relations. During the 4 minute training session, they listened to stems containing the relevant consonants and vowels (but not to ones containing the relevant rimes or bodies) paired with affixes. They were instructed not to press any buttons. The stimuli were arranged so that every stem-affix combination (“**word**”) was followed by another word whose stem differed from the stem of the first combination by one segment (either a consonant or a vowel), e.g., /bæʃ-mɪn/ followed by /bɪʃ-num/ or /bɪg-num/ followed by /bug-num/. The words in these minimal pairs differed in the affix they took on half of the trials. Thus, the difference between the words in a minimal pair was irrelevant for affix choice in half of the pairs (as in /bɪg-num/ followed by /bug-num/). This balancing ensured that the training does not place stimuli sharing the rime next to each other more often than it places stimuli sharing the body next to each other.

This training session was followed by a testing session in which subjects were presented with stems they have already heard but they heard Gaussian noise in place of the affix. The noise had the same amplitude contour and duration as the average of the two affixes (/mɪn/ and /num/). The subjects were instructed to guess which affix has been replaced with noise. Once they made their guess, the correct stem+affix combination was pronounced. In the subsequent generalization block, the subjects heard novel stems paired with noise. No feedback was given.

This initial stage of the experiment, “**training on parts**”, leads to several critical comparisons. We can compare the associability of onsets and codas, allowing us to assess whether any differences found between bodies and rimes can be explained by differences between onsets and codas. Further, to ensure that subjects assigned to learn rime-affix dependencies are not just (by chance) better learners than those assigned to learn body-affix dependencies, we can compare individual differences in learning ability across subject groups by looking at how well the subjects are learning vowel associations. Third, we can compare the associability of prefixes and suffixes and assess the existence of locality effects, since the onset is adjacent to the prefix and the coda is adjacent to the suffix.

Finally, we can determine base rates of generalization from a set of rimes or bodies to a novel rime/body that consists of segments that have been presented and associated with the same response but have never been presented together. The last piece of information has dual significance: 1) we may expect that there will be less generalization to an unfamiliar rime than to an unfamiliar body if subjects are spontaneously storing rime-affix but not body-affix pairings during this stage, and 2) reaction to stimuli containing the crucial wholes (/ʃæ/ and /gʌ/ or /æʃ/ and /ʌg/) prior to training on those wholes provides us with a baseline level of accuracy to which the accuracy level following training on wholes can be compared.

After the generalization block, subjects go through the training-feedback-generalization cycle again. The only difference is that now they receive training on both parts and wholes. Thus subjects in groups I and II learn rime-affix, vowel-affix and coda-affix associations while those in groups III and IV learn body-affix, vowel-affix and onset-affix associations.

Table 2 summarizes how much total training and testing subjects get for each type of potential constituent. Since previous research (Bonatti et al., 2005; Creel et al., 2006) suggests that vowels are less associable than consonants, subjects received more vowel training than consonant training.

Constituent	Training Trials		Feedback Trials		Total		Generalization Trials	
	Stage I	Stage II	Stage I	Stage II	Stage I	Stage II	Stage I	Stage II
C	32	20	20	10	52	30	16	8
V	32	24	20	10	52	34	28	22
Body/rime	0	46	0	10	0	56	12	22
Total	64	92	40	30	104	122	56	52

Table 2. How much practice do they get?

Participants. The participants were introductory psychology students who received course credit in exchange for their participation. They were not rewarded for accuracy or speed. All participants reported being native English speakers with no history of speech, language, or hearing impairment. There were 17 participants in each of the four subject groups (rime-prefix, rime-suffix, body-prefix, and body-suffix).

Procedures. The stimuli were presented over headphones at a comfortable listening level. Subjects were seated at a testing station consisting of a computer on a desk surrounded by cubicle walls. Instructions appeared on the screen between the stages of the experiment. The instructions are presented in the appendix. The subjects could take as much time as they needed to read the instructions. Subjects were randomly assigned to one of the four subject groups and one of the four computers in the room.

Materials. Stems and affixes were recorded separately and concatenated using Praat. There are a few noteworthy things about the stimuli: 1) the segments that begin suffixes are the same segments that end prefixes (nasals), 2) prefixes and suffixes are acoustically identical and not prosodically integrated with the stem, 3) bodies are less frequent than rimes, while onsets and codas are similar in frequency, although onsets are somewhat more frequent than codas, and 4) stimuli used for testing do not include syllables used for training and include consonants and consonant clusters that are not used in training.

Controlling for Frequency. Convergent estimates of the frequencies of various rimes, bodies, onsets, vowels, and codas were obtained from the MRC Psycholinguistic Database (Coltheart et al., 1981; http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm), Kessler and Treiman (1997, <http://brettkessler.com/SyllStructDistPhon/>), and the Hoosier Mental Lexicon (Nusbaum et al., 1984; <http://128.252.27.56/Neighborhood/SearchHome.asp>). In order to minimize physical differences between onsets and codas, certain consonants were excluded from consideration: 1) voiceless stops due to the presence of aspiration in onsets of stressed syllables but not in codas, 2) nasals since they have much more nasalizing influence on the preceding vowel than on the following one, 3) /r/ because of fusion with the preceding vowel, 4) /l/ because of vast differences in pronunciation in onset and coda positions, 5) /w/ and /j/ because of restriction to word-initial position, 6) /d/ and /z/ because of possible morphological interpretation in coda position. Affricates were also eliminated because they may be more internally complex.

Table 3 presents the data across databases. Both of the consonants selected for the experiment are relatively balanced in how frequently they are used in the onset vs. in the coda. If anything, the consonants are slightly more frequent in the onset. The databases display a remarkably high agreement regarding the distributions of the consonants.

Database Consonant	MRC			Kessler and Treiman (1997)			HML		
	On	Cd	%On	On	Cd	%On	On	Cd	%On
b	216	90	71	154	62	71	150	75	67
g	102	97	51	88	67	57	80	77	51
v	61	85	42	45	54	45	39	65	38
f	144	97	60	92	68	58	109	70	61
s	183	169	73	126	116	52	121	113	52
ʃ	89	70	56	65	44	60	70	54	52

Table 3. Type frequencies of the consonants in codas vs. onsets of monomorphemic CVC words (consonants eventually used are in bold)

Table 4 shows the frequency distributions for rimes and bodies involving the consonants derived from the Hoosier Mental Lexicon. Since the databases display such a high agreement rate, only HML results are shown. As can be seen from the table, there is evidence that the chosen vowels are linked to the coda more strongly than they are linked to the onset. The rimes in the study are more frequent than the bodies. Thus, formal and usage-based criteria for constituency converge for the stimuli used in the present study.

	gV	Vg	%rime	ʃV	Vʃ	%rime
æ	9	20	69	8	21	82
ʌ	5	17	77	5	12	71

Table 4. Type frequencies of rimes and bodies of monomorphemic CVC words for the chosen consonants and vowels in the Hoosier Mental Lexicon (chosen rimes and bodies in bold)

Prefixes and Suffixes. The syllables /mɪn/ and /num/ were used as both the prefix and the suffix. The syllables were chosen so that none of the rules had an obvious phonetic motivation. In addition, we ensured that the set of consonants adjoining the stem was the same for prefixes and suffixes. These consonants (nasals) were chosen to be relatively perceptible, and relatively unlikely to interfere with the perception of the adjacent stem consonant. In addition, we made sure that they do not cause perceptual resyllabification.

The Generalization Stimuli. The generalization stimuli were of four types: 1) stimuli that contained consonant clusters in the unattended position (the position not involved in the generalizations always contained one consonant during training), e.g., /plæʃ/ for subjects learning rime-affix associations or /ʃælp/ for those learning body-affix associations, 2) stimuli that contained /l/ as the consonant in the unattended position (the consonant did not occur during training), e.g., /læʃ/ for rime-trained subjects or /ʃælp/ for body-trained ones, and 3) stimuli in which the two consonants of the stem were identical, e.g., /zæz/, /gʌg/, /ʃɪʃ/. No differences between generalization stimulus types were found, hence the results reported later are averaged across generalization stimulus types.

Recording the Stimuli. All syllables involved in the study were produced by a single male native American-English speaker who was unaware of the purpose of the study. In addition to the stimuli used in the study, the speaker also produced a large number of distractors that did not involve the target rimes and bodies. The affixes were produced only once by the speaker. The speaker did not know that these syllables had any special status. Each syllable was produced in isolation in response to a visual prompt appearing on a monitor for a fixed amount of time.

Results

Bodies vs. Rimes

After Training on Wholes. As shown in Figure 3 and Table 5, subjects who were exposed to rime-affix correlations pressed the appropriate button in response to novel stimuli containing a relevant rime in about 70% of the cases, which is significantly above chance (50%) according to a one-sample t-test ($t=5.955, df=33, p<.0005$). That is, the subjects succeeded in learning to respond with /mɪn/ when presented with a syllable ending in /æʃ/ and to respond with /num/ when presented with a syllable ending with /ʌg/. By contrast, subjects who were exposed to body-affix correlations did not learn those correlations, responding at chance levels. That is, when presented with a syllable that began with /ʃæ/ or /gʌ/ the subjects responded with either /mɪn/ or /num/ with equal probability.

Analysis of variance with constituency, affix location, and correct response (mɪn vs. num) as independent variables and accuracy as a dependent variable showed the rime vs. body difference to be statistically significant ($F(1,66)=123.431, p<0.0005$) with a large effect size (Cohen’s $d=1.17$) while there is no significant difference in associability between prefixes and suffixes ($F(1,66)=.431, p=.517$), no significant effect of correct response and consequently rime identity (mɪn vs. num, æʃ/ʃæ vs. gʌ/ʌg ($F(1,66)=1.026, p=.315$), and no significant interactions. Planned by-subjects and by-items t-tests confirmed the significance of the effect of constituency (by subjects: $t=5.401, df=66, p<.0001$, by items: $t=13.445, df=42, p<.0001$).

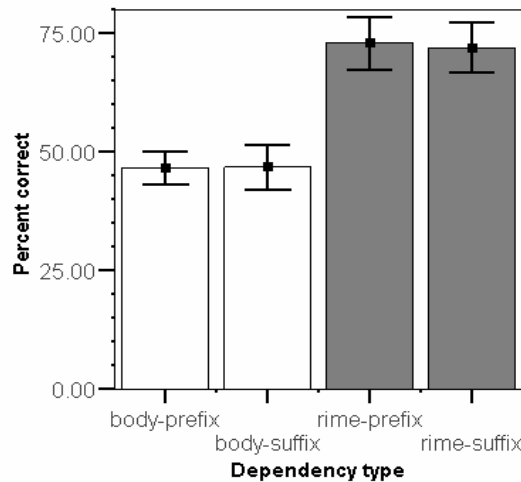


Figure 3. Rime associations are easier to learn than body associations (bars show means, error bars show by-subject standard errors)³

³ By-item standard errors are smaller (+/-1.4%).

	rime-prefix	rime-suffix	body-prefix	body-suffix
%correct	73	72	47	47
Average across affixes	73		47	

Table 5. Subjects' generalization accuracy (% correct) on rime-affix vs. body-affix correlations after training.

These results suggest that rimes are more associable than bodies. At this point we can reject any explanation of these results that is based on differences in associability between prefixes and suffixes as well as an explanation based on locality effects. The rime is more associable than the body regardless of whether the subjects are learning prefix associations or suffix associations and thus also regardless of whether the affix is adjacent to the stem-internal segment sequence that it co-occurs with. In addition, lack of significant interactions with correct response indicates that individual bodies and rimes functioned similarly.

Prior to Training on Wholes. We need to determine whether the results obtained might be due to differences in the likelihood of generalizing consonant and vowel associations learned during stage I to novel rimes and bodies. This section provides the results from generalization trials following training on parts (prior to training on wholes).

Prior to training on stimuli that contain the relevant bodies and rimes, the subjects are trained to acquire associations of the consonants and vowels that compose those rimes and bodies. Since the whole is associated with a different response than both of its parts in XOR learning, training on parts alone should make stimuli containing the whole become associated with responses appropriate for stimuli that contain only one of the parts. If this happens, accuracy on wholes should be below chance prior to training on wholes. Table 6 shows the results. The rime-suffix condition is significantly different from the other conditions (the interaction of constituent and affix is significant: $F(1,64)=5.51, p=.019$).

%correct	rime-prefix	rime-suffix	body-prefix	body-suffix
Generalization 1	29	43	28	26

Table 6. Subjects' generalization accuracy (% correct) on rime-affix vs. body-affix correlations prior to training

Given this result, subjects who are about to learn rime-suffix dependencies need to improve by fewer percentage points than other subjects, including those learning body-affix dependencies, to reach the same level of performance following Stage II. In the remainder of this section we address this concern by showing that a subject's accuracy level following Stage I is in fact a very poor predictor of the same subject's accuracy level following Stage II indicating that differences in accuracy between groups following Stage I do not give rise to differences in accuracy following Stage II.

Table 7 shows bivariate Pearson correlations between each subject's performance on vowels, consonants and wholes during the various stages of the experiment. Within a training stage, accuracy on wholes has a strong negative correlation with accuracy on segments.⁴ However, accuracy following Stage I (C1, V1, or W1) does not correlate with accuracy on wholes following Stage II (W2). This indicates that differences in the accuracy level prior to training on wholes cannot account for differences in accuracy

⁴ This is expected because whenever a whole is perceived, its parts are perceived as well. Thus, when one hears something like /Cæf-mɪn/ one re-associates /æ/ and /f/ with /mɪn/.

level following training on wholes. Thus, despite the fact that subjects are more accurate on rime-suffix associations than on other whole-affix associations prior to training, this cannot account for the difference between how well subjects perform on rime-suffix vs. body-prefix dependencies following Stage II training.

		W1	C1	V1	C2	V2
W1	r		-.623	-.428		
	sig.		.000	.000		
W2	r	.111	-.016	-.011	-.526	-.504
	sig.	.366	.897	.930	.000	.000

Table 7. Relations between each subjects’ performance on various stimuli in different stages of the experiment (W = wholes, i.e., bodies or rimes, V = vowels, C = consonants, stage 1 precedes training on wholes, stage 2 follows).

Familiar vs. Novel Syllables. One possible interpretation of the better performance of subjects exposed to rime-affix dependencies compared to subjects exposed to body-affix dependencies is that all subjects are actually learning syllable-affix dependencies but that such pairings are easier to learn when syllables paired with the same affix are similar. Previous research has found that English speakers judge syllables sharing rimes to be more similar than syllables sharing bodies (Nelson & Nelson, 1970; Vitz & Winkler, 1973; Derwing & Nearey, 1986; Bendrien, 1992; Yoon & Derwing, 2001). Therefore, subjects could perform better when exposed to rime-affix dependencies than when exposed to body-affix dependencies if they are learning syllable-affix pairings in either case and there is no difference in associability between bodies and rimes (cf. Geudens et al., 2005). However, if subjects are learning syllable-affix pairings, they should perform better when presented with familiar syllables than when presented with novel syllables.

Table 8 compares subjects’ performance on novel syllables (not presented during training) and familiar syllables. Separate ANOVA’s were conducted for rime-affix and body-affix subject groups. There is no main effect of syllable familiarity ($F(1,66)=.002, p=.98$). Examination of Table 8 shows that the only subjects for whom there is a significant difference between familiar and novel syllables in the expected duration are subjects acquiring body-prefix associations. For rimes, the effect is in the other direction ($F(1,33)=5.433, p=.02$): subjects perform slightly better on novel syllables. Therefore, we can reject the hypothesis that our subjects are learning syllable-affix associations instead of rime-affix associations.

Constituent	Rime				Body			
	Prefix		Suffix		Prefix		Suffix	
Stimuli are	Familiar	Novel	Familiar	Novel	Familiar	Novel	Familiar	Novel
%correct	71	74	66	71	60	49	49	49

Table 8. Testing with familiar vs. novel syllables

Consonants and Vowels

In this section we compare associability of onsets to associability of codas. This is necessary to show that the rime/body difference is not due to an onset/coda difference. We also examine whether the

subjects are really forming rime-affix associations and not just re-associating consonants and/or vowels with an inappropriate response by examining consonant and vowel associations after training on wholes.

Prior to Training on Wholes. Table 9 shows how accurate subjects were on consonant-affix and vowel-affix associations prior to training on wholes. The results were analyzed using an ANOVA with constituent, affix, segment type (consonant vs. vowel), and correct response as independent variables. There is a significant difference between consonants and vowels ($F(1,66)=48.108, p<.0005$): accuracy is lower for vowel-affix associations than for consonant-affix associations. In addition, there is a significant affix type by segment type interaction ($F(1,64)=16.908, p<.0005$) and a marginally significant segment type by correct response interaction ($F(1,64)=3.702, p=.054$). The difference between consonants and vowels is larger when suffix occurrence is being predicted than when prefix occurrence is, as shown in Table 9.

Association Type	Local		Non-local		Local		Non-local	
	Coda	Vowel	Coda	Vowel	Onset	Vowel	Onset	Vowel
%correct	61	56	71	58	73	55	67	60

Table 9. Accuracy on consonant-affix and vowel-affix relations prior to training on wholes

Table 9 shows that consonants were more associable than vowels. This result is especially striking given the fact that the subjects received more instruction with vowels than they did with consonants: while there were 30 training trials involving consonant-affix relations, there were 44 involving vowel-affix relations (Table 2). The lower associability of vowels compared to consonants is in line with findings of Bonatti et al. (2005), who found that non-adjacent statistical dependencies between vowels are harder to learn than non-adjacent dependencies between consonants but not with Newport and Aslin (2004), who failed to find a difference.

In addition, there is no significant effect of syllabic position on consonant associability. It is not the case that codas are significantly more or less associable than onsets. Finally, subject groups do not differ significantly on how well they acquire vowel associations. This result indicates that the differences that are observed between subjects acquiring rime associations and those acquiring body associations are not simply due to how good the subjects assigned to those groups are at associative learning of the type tested in this experiment.

After Training on Wholes. The reason to examine how well subjects perform on segments after being trained on wholes is that it could be the case that training on wholes simply causes subjects to re-associate the segments with the response that is appropriate for the whole. In this section, we show that this hypothesis is ruled out because accuracy on segments does not fall below chance, indicating that segments do not get associated with the response that is appropriate for the rime.

After body-affix or rime-affix associations are introduced, generalization accuracy on vowels and consonants displays main effects of constituent ($F(1,66)=10.421, \text{rime vs. body}, p=.001$), and segment type ($F(1,66)=23.851, \text{consonant vs. vowel}, p<.0005$).

Table 10 shows how accurate subjects were on consonant and vowel associations after having been introduced to body or rime associations as a function of which constituent was introduced, where the affix was located relative to the root, what the segment was, and how much training was given. The table shows a main effect of constituent: subjects are somewhat more accurate on segment associations when they are exposed to body, rather than rime associations. This could be due to a competition between the rime and the segments it contains. Since the body never gets associated with a response, it does not

compete with the segments it contains. In addition, the main effect of segment type is present: subjects are more accurate on consonants than on vowels.

Association	Local		Non-local		Local		Non-local	
	Coda	Vowel	Coda	Vowel	Onset	Vowel	Onset	Vowel
Associated	57	44	52	44	61	53	53	48
%correct	48		46		49		55	

Table 10. Accuracy on consonant-affix and vowel-affix relations after body/rime training: effects of affix location, segment type, and constituent/consonant location.

Comparing Learning Rates

Table 11 displays accuracy as a function of amount of training on a unit type. The number of training trials is the number of training trials preceding Generalization 1 for consonants and vowels and Generalization 2 for rimes and bodies. The baseline is chance for consonants and vowels and accuracy prior to training on wholes for rimes and bodies. Table 11 shows that the speed of acquiring the association (change in accuracy by number of training trials) is much smaller for vowels than it is for consonants and much larger for rimes than for either consonants or vowels.

	Rime	Body	Consonant	Vowel
Training only	0.80	0.50	0.50	0.19
Training & feedback	0.66	0.41	0.32	0.12

Table 11. Learning rates for rimes, bodies, consonants and vowels (per trials of exposure). The top row of numbers use the number of training trials in the denominator while the bottom row uses the sum of training and feedback trials.

The Model

The Framework

Representing Syllable Structure. In the present framework, the syllable is represented as a matrix of resting activation levels, as shown in Table 12. Each cell in the table corresponds to a node or a link. The cells corresponding to nodes are the ones for which the row label and the column label are identical. The columns correspond to the nodes to which the links point while the rows are the nodes from which the links originate. Cells corresponding to links of strength equal to zero are left empty. Note that if a node has a resting activation level of zero, the strengths of links heading to and radiating from the node are assumed to have a strength of zero. Thus nodes that have a strength of zero simply don't exist (to the rest of the network). Since the body does not exist in the structure shown in Table 12, cells corresponding to links pointing to and from the body are merged and shaded, indicating that they must be equal to zero.

Table 12 shows a version of the traditional structure of the syllable in which the rime is a node while the body is not, and each part is connected only to the whole that immediately dominates it. Thus, for instance, the nucleus is not connected to the syllable because it is dominated by the rime. However, one should note that this representation is more specific than the traditional tree diagram in that we specify whether a given connection is excitatory or inhibitory as well as exact strengths of both nodes and connections. Furthermore, all links in the present framework are directed. Thus, for instance, in Table 12,

wholes inhibit parts while parts excite wholes, and the amount of inhibition is smaller than the amount of excitation.

	Onset	Nucleus	Coda	Body	Rime	Syllable
Onset	$R_O = 1.0$					$R_{O>S} = 0.3$
Nucleus		$R_N = 0.7$			$R_{N>R} = 0.3$	
Coda			$R_C = 1.0$		$R_{C>R} = 0.3$	
Body						
Rime		$R_{R>N} = -0.2$	$R_{R>C} = -0.2$		$R_R = 0.5$	$R_{R>S} = 0.3$
Syllable	$R_{S>O} = -0.2$				$R_{S>R} = -0.2$	$R_S = 1.0$

Table 12. An example of a syllable structure.

The syllable structure in Table 12 can also be represented as in Figure 4.

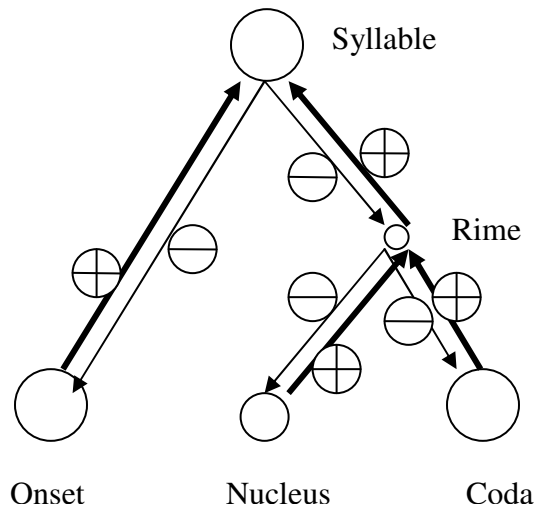


Figure 4. A graphical representation of the syllable structure in Table 12: The coda, the onset and the syllable are the most salient nodes, followed by nucleus, followed by rime. There is no body node. There are excitatory bottom-up connections and weaker inhibitory top-down connections.

The resting activation values displayed in the syllable matrix are used to determine how much activation will be received by memory traces representing each part of a syllable presented to the model. The amount of activation received by a chunk of a particular type is increased when 1) the resting activation level of the node representing the type is high, 2) there are many strong excitatory links pointing to the node, and 3) the inhibitory links pointing to the node are weak and few in number. In addition, there is a free parameter in the model, which determines whether a node loses (some) activation that spreads from it into the radiating links. If this parameter is positive, the amount of activation stored in the node after a syllable containing its referent is presented is reduced when the cumulative strength of links radiating from the node, i.e., the sum of absolute values of their activation levels, is high.

The resting activation levels are used to derive the proportion of input activation that is going to be stored in representations corresponding to each part of a syllable (and the syllable as a whole) after the syllable has been presented.

For each node, the formula in (1) is applied where X is the current node, $\neg X_1$ through $\neg X_{\text{num_earlier}}$ are the nodes that represent chunks that end before X ends in the syllable or that end at the same time and are part of X, and k is the free parameter that determines how much activation that spreads out of the node is lost to the node. When the end of the syllable is reached, the formula in (2) is applied to each node where remaining nodes are the ones that come after X in the syllable or of which X is a part. The updating sequence and corresponding *num_earlier* and *remaining* numbers are shown in (3). Note that multiplying the final result by R_X ensures that nodes that don't exist do not end up getting any activation and thus are not associable.⁵ This multiplication represents the likelihood of parsing the chunk out of the signal during testing. If the likelihood is zero, then how often the chunk has co-occurred with a given affix during training is not relevant.

$$(1) \quad A_X = R_X - \frac{k * \sum_{i=1}^5 |R_{X>\neg X_i}|}{\sum_{i=1}^5 |R_{X>\neg X_i}| + R_X} + \sum_{i=0}^{\text{num_earlier}} \frac{A_{\neg X_i} * R_{\neg X_i>X}}{\sum_{j=1}^5 |R_{\neg X_i>\neg X_j}| + |R_{\neg X_i>X}|}$$

$$(2) \quad A_X = R_X * (A_X + \sum_{i=0}^{\text{remaining}} \frac{A_{\neg X_i} * R_{\neg X_i>X}}{\sum_{j=1}^5 |R_{\neg X_i>\neg X_j}| + |R_{\neg X_i>X}|})$$

- (3) Formulas are applied in the following order:
onset → nucleus → body → coda → rime → syllable

where

num_earlier =	0	1	2	3	4	5
remaining =	5	4	3	2	1	0

Training. The model is presented with a list of syllables paired with affixes. For each onset, nucleus, coda, body, rime and syllable encountered, the model calculates the number of times it is encountered with each of the affixes, e.g., it might remember that the onset /g/ co-occurred 18 times with /num/ and 0 times with /min/ during stage 1.

After going through Stage I training, the model saves the co-occurrence statistics that it has gathered. Numbers of chunk-affix co-occurrences gathered during Stage II are simply added to the numbers gathered in Stage I. Thus, if the onset /g/ co-occurs 24 times with /min/ and 10 times with /num/ during Stage II, the model will know that overall /g/ co-occurred with /num/ 28 times and with /min/ 24 times.

Each time the model encounters a chunk-affix pair, it will increment the counter for that chunk-affix pair by the activation value A for that chunk. Thus, in our example, if A_{onset} is 0.9, the onset /g/'s /num/ **score** will be $18 * 0.9 = 16.2$ following Stage I. Thus, activation level of a chunk depends on where it is in the syllable, and nodes with low activation levels are less associable than nodes with high activation levels where activation level is a product of the resting activation level of the node, how much activation

⁵ When R_X is zero, the summed strength of all links radiating from it is set to 1 to avoid dividing by zero. The choice of 1 is arbitrary since the only number divided by the strength of links radiating from a non-existent node is zero thus the result of division is always zero.

and inhibition it receives from other nodes, and, if $k > 0$, how much activation it sends away to excite or inhibit other nodes.

There is a free parameter in the model that allows the model to assign different weights to each stage of training. This is done by dividing all scores from a given stage by the parameter after training on the immediately following stage. The parameter has no effect on performance on the test immediately following the training stage for which it is set. Thus, if the parameter is set to 2 during Stage I, all co-occurrence numbers from stage I will be divided by 2 before being added to co-occurrence numbers from Stage II. Thus, in our example above, on Test II, which follows Stage II, the model will think that /g/ co-occurred with /num/ 19 times, even though in reality it co-occurred with it 28 times because half of the co-occurrences that happened during Stage I will be forgotten or discounted. Thus, /g/'s /num/ score following stage II will be $18*0.9*0.5+10*0.9=17.1$.

Thus, the overall score for a given chunk-affix pair is given by the formula in (4) where S is the score for a given chunk-affix pair, D_i is the decay parameter associated with stage I, which is always equal to 1 for the training stage immediately preceding the test, and $P(\text{chunk}, \text{affix})_i$ is the number of times the chunk co-occurred with the affix during stage i .

$$(4) S_{(\text{chunk}, \text{affix})} = \sum_{i=1}^{\#_of_preceding_stages} (P(\text{chunk}, \text{affix})_i * A_{\text{Chunk}} * D_i)$$

Testing. During testing the model is presented with novel syllables that are not paired with an affix. For each syllable, the model extracts all the chunks that are present in it. It then recalls the scores representing how often each of the chunks present in the syllable has co-occurred with each of the suffixes during training and how salient these pairings have been overall, depending on the activation level of the chunk and the distribution of chunk-affix co-occurrences across stages of training. To predict which affix should go with the presented syllable, we sum the min-scores of all the chunks in the syllable and subtract the sum of num-scores for each chunk in the syllable from the result. If the result is positive, /num/ is the predicted affix, while /mIn/ is predicted if the result is negative.

Thus, the prediction value for a given syllable is given in equation (5) where each syllable consists of six chunks: onset, nucleus, coda, body, rime, and syllable. One can think of the prediction value as the model's confidence in its choice or the likelihood that the model would go along with its choice when random noise is injected into its performance.

$$(5) \text{Pr}_{\text{Syllable}} = \sum_{i=1}^6 (S_{(\text{chunk}_i, \text{num})} - S_{(\text{chunk}_i, \text{mIn})})$$

Comparing Models

Any model of syllable structure should explain several basic results. Namely, it should predict that: rimes are more associable than bodies (Table 5), rimes (and possibly bodies) are more associable than segments in terms of speed of acquiring association (Table 11), it is possible to associate a rime with suffix X while its parts are not associated with any suffix, being at chance (Table 5 and Table 10), the coda is somewhat less associable than onset but the difference in associability between coda and onset is much smaller than that between rime and body (Table 5 vs. Table 9), training on parts of a rime does not generalize to the rime as much as training on parts of the body generalizes to the body (Table 6), "body associations" do not show full generalization to novel syllables while rime associations do (Table 8), the onset-rime structure is not universal, hence cannot be explained by the fact that the rime follows the body

temporally (Lee, 2006), and either syllable nodes or rime nodes are necessary to account for the fact that priming with two shared segments is more effective than priming with three shared segments when the two segments form a syllable in the target word (Fernald et al., 1996; Carreiras & Perea, 2002).

The fact that there is cross-linguistic variability in syllable structure (Yoon & Derwing, 2001) and that frequency of between-segment co-occurrence is a crucial determinant of whether the nucleus is more closely tied to the coda or to the onset within a single language (Lee, 2006) indicate that the rime/body difference is not about which comes first temporally or its segmental shape (CV vs. VC). Therefore, we will not consider explanations for the asymmetry that propose that connections from the preceding to the following are stronger than connections going in the reverse direction. If our model is to capture variation in syllabic constituency, explanations must focus on consequences of constituency, not precedence.

This leaves the following space of possibilities: 1) parts of constituents may have a lower resting activation level than parts of non-constituents, 2) part→whole connections may be stronger within constituents than across constituent boundaries, 3) part-part connections may be stronger within constituents than across constituent boundaries, and 4) constituents may have higher resting activation level than non-constituents. However, option 1 is ruled out by the finding that a difference in associability between wholes does not entail a difference in associability between parts: while rimes are more associable than bodies regardless of whether local or non-local associations are at issue, codas are less associable than onsets only for local associations.

For each of the three remaining possibilities, the following questions should be addressed: a) are there excitatory part→whole connections, b) are there whole→part connections, c) are the whole→part connections inhibitory or excitatory, and d) is activation that spreads from a node into the radiating links lost to the source node.

The following possibilities will be left outside the scope of this paper: 1) there is more inhibition coming into a non-constituent than into a constituent, 2) there are more links radiating from a non-constituent than from a constituent, and 3) inhibitory connections leading from a constituent to its parts are stronger than those leading from a non-constituent to its parts. The reason these models will not be considered is that the author can think of no plausible psychological interpretations for them.

We will start out with a maximally simple model in which the only nodes that exist are segment nodes. Then we will make additions to this model to create asymmetric structures and examine the resulting differences in performance, i.e., prediction value, relative to the original model. Every pair of models compared will differ only in how the rime or connectivity within the rime are handled, so we can compare the models by looking at their performance on tests given to subjects trained on rime-affix associations. In this way, we will determine which modifications increase associability of the constituent and change the amount of generalization to novel syllables and the extent to which they rely on increases in part associability to increase associability of the whole.

The Distributed Baseline

We start out with a model that is unable to capture the rime/body asymmetry with the intent of examining ways of elaborating it to produce the observed results. For the purposes of comparing the models, we have eliminated the asymmetry in amount of training between vowels and consonants during stage I so that vowel and consonant scores after stage I are identical if vowels and consonants are equally associable under the model considered.

The syllable structure is shown in Figure 5. In this model, the only nodes are onset, nucleus, and coda. There are no nodes corresponding to larger structures. There are no links within the syllable.

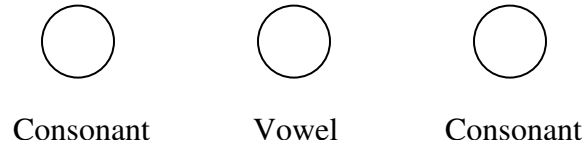


Figure 5. A network representation of a CVC syllable under the distributed baseline model.

Table 12 shows the results of training the model using the same stimuli presented to human subjects. The row labeled ‘C’ shows the model’s accuracy on stimuli that share a consonant with the training stimuli, the consonant being associated with one of the affixes. The row labeled ‘V’ shows accuracy on stimuli that contain a vowel that has been paired with a particular affix during training.

Stimulus type	Stage I		Stage II
	Familiar syllables	Novel syllables	Novel syllables
C	18	18	5
V	18	18	7
CV or VC		-36	-12

Table 12. The simple distributed model’s performance on testing trials following Stage I and II of training. Positive scores indicate that the association is correct while negative scores indicate that it is incorrect. The greater the absolute value of a score, the stronger the association. This is a version of the model with no discounting of Stage I training.

While the model does not achieve accurate performance on wholes without assuming that the training from Stage I is discounted, the score for a whole changes faster than the score for any individual part. How fast then does this purely distributed model predict the learning of whole-affix associations to be relative to the speed of learning part-affix associations?

The speed of learning associations of the whole (**V** for ‘velocity’) is given by accuracy on the wholes after stage II minus accuracy on the wholes after stage I divided by the number of whole-affix pairings during stage II training, or

$$(6) V_w = (\text{Pr}_{w_2} - \text{Pr}_{w_1}) / P_{w_2}$$

The speed of learning associations of a consonant is given by accuracy on the part after stage I divided by the number of consonant-affix pairings during Stage I. Similarly, accuracy on vowels is accuracy on vowels after Stage I divided by number of vowel-affix pairings during Stage I.

$$(7) V_c = \text{Pr}_{c_i} / P_{c_i}$$

$$(8) V_v = \text{Pr}_{v_i} / P_{v_i}$$

In the purely distributed model the accuracy score for the whole is always equal to the sum of accuracy scores for the parts multiplied by -1 . Thus,

$$(9) \text{Pr}_{w_i} = -(\text{Pr}_{c_i} + \text{Pr}_{v_i})$$

Every time a whole is perceived both of the parts are perceived as well. Thus, whenever a whole-affix pairing is presented in stage II, the parts of the whole are paired with the affix with which they are

not associated during stage I. Thus to derive the accuracy on a part after stage II we need to add the number of times that part was paired with the same affix as in Stage I and subtract the number of times it was paired with a different affix, the latter being equal to the number of times the whole was presented. The equation for the consonants is shown in (10).

$$(10) \Pr_{C_2} = \frac{\Pr_{C_1}}{d} + V_C * P_{C_2} - V_C * P_{W_2}$$

Now we can express the speed of acquiring whole associations as a function of the speed of acquiring part associations and number of times parts and wholes are presented during each stage of the experiment.⁶

$$(11) \quad V_W = (\Pr_{W_2} - \Pr_{W_1}) / P_{W_2}$$

$$\Pr_{W_1} = -(V_C * P_{C_1} + V_V * P_{V_1})$$

$$\Pr_{W_2} = -(V_C (\frac{P_{C_1}}{d} + P_{C_2} - P_{W_2}) + V_V (\frac{P_{V_1}}{d} + P_{V_2} - P_{W_2}))$$

Therefore,

$$(12) V_W = (V_C (P_{C_1} - \frac{P_{C_1}}{d} - P_{C_2} + P_{W_2}) + V_V (P_{V_1} - \frac{P_{V_1}}{d} - P_{V_2} + P_{W_2})) / P_{W_2}$$

If $d=1$, as in our current model,

$$(13) V_W = (V_C (P_{W_2} - P_{C_2}) + V_V (P_{W_2} - P_{V_2})) / P_{W_2}$$

In the present experiment,

$$P_{C_2} = 10$$

$$(14) P_{V_2} = 12$$

$$P_{W_2} = 23$$

Therefore,

$$(15) V_W = (13V_C + 11V_V) / 23 = 0.57V_C + 0.48V_V$$

Thus, under this simplest model, rimes and bodies are not predicted to be more associable in our experiment than both consonants and vowels unless

$$(16) \quad 0.43V_C < 0.48V_V$$

$$0.52V_V < 0.57V_C$$

⁶ Pr = prediction, p = frequency of occurrence, v = learning rate.

That is,

$$(17) 0.9V_C < V_V < 1.1V_C$$

This prediction contrasts with the experimental findings where rimes are more associable than single segments while V_C is far greater than V_V . The model is also unable to account for subjects' ability to learn to associate a rime with response X while its parts are not associated with response X. Finally, the model does not produce incomplete generalization, which we have observed with body-affix dependencies in the experiment.

We will now examine possible reasons for observing incomplete generalization, ways of making constituents more associable than their parts, and making constituents more associable than non-constituents of the same size.

Incomplete Generalization

There are two ways to produce incomplete generalization of constituent-affix associations: storage of partially overlapping constituents and storage of a larger unit, the syllable. Thus, we observe incomplete generalization of body-affix associations in the model shown in Figure 6 and in the one shown in Figure 7.

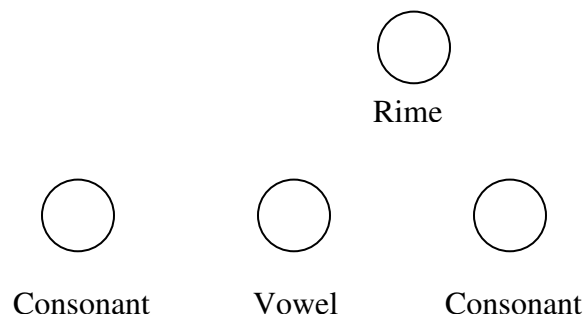


Figure 6. A model that produces incomplete generalization of body-affix associations due to the existence of a rime node.

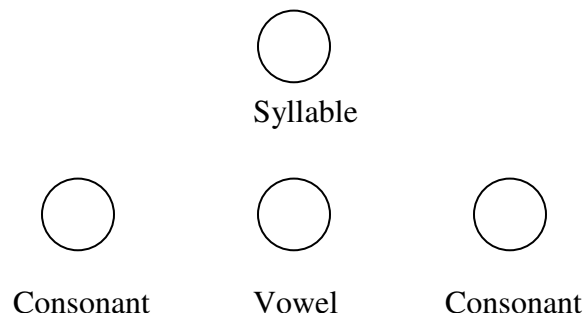


Figure 7. A model that produces incomplete generalization of body-affix associations due to the existence of a syllable node

Table 13 shows that scores on feedback trials, which consist of familiar syllables, increase relative to the distributed baseline while scores on generalization trials, which consist of novel syllables, stay the same when either rime nodes or syllable nodes are introduced.

	Distributed Baseline	Store Rimes	Store Syllables
Trained syllables	-12	-11	-9
Novel syllables	-12	-12	-12

Table 13. Storage of syllables and/or rimes produces incomplete generalization of body associations: scores on body stimuli after Stage II

Thus, both models produce incomplete generalization of body associations and, interestingly, incomplete generalization of body associations is consistent with the existence of rime nodes. The observed absence of incomplete generalization for rime associations is expected if body nodes do not exist. Alternatively, it may indicate that syllable associations are not formed if the subjects are performing the task relatively accurately since overall accuracy is higher after training on rimes.

Increasing Constituent Associability

Constituent associability can be increased by increasing the activation level that the constituent node has during training. This can be accomplished by either increasing the resting activation of the constituent node (Figure 8), by increasing connection strength between parts of a constituent and the constituent node (Figure 9) and by increasing connection strength between parts of a constituent if those nodes will send some of their activation to the constituent node (Figure 10). In all the models here, nodes do not lose the activation that spreads from them to other nodes, i.e., $k=0$.

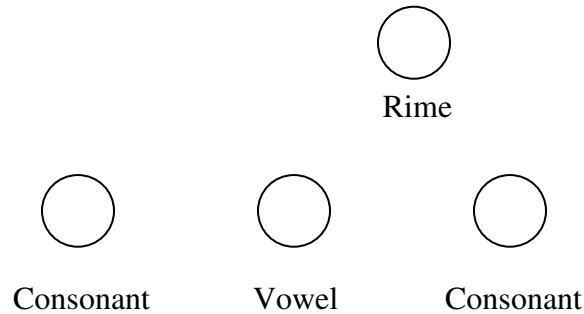


Figure 8. The parseability/salience model: the rime has a higher resting activation level than the body.

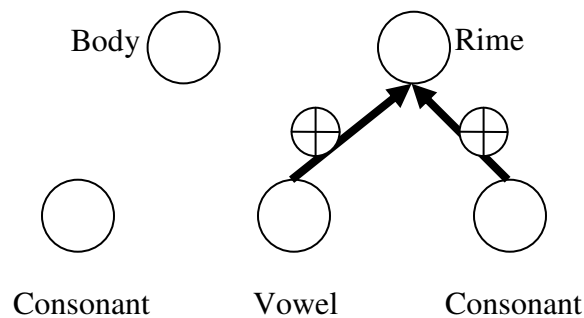


Figure 9. The part-to-whole connectivity model: segment-to-rime are stronger than segment-to-body connections where both body and rime exist.

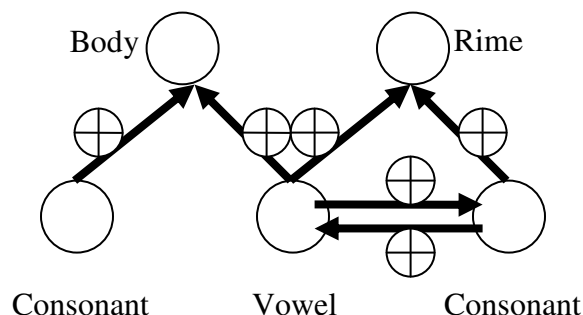


Figure 10. The part→part connectivity model: the onset←→nucleus connection is weaker than the nucleus←→coda connection where both body and rime exist, and both segment→body and segment→rime connections exist and are of equal strength.

It is important to note that the representation of the body in Figure 8 is the same as the representation of the rime in Figure 7, representation of the body in Figure 9 is the same as the representation of the rime in Figure 8, and representation of the body in Figure 10 is the same as the representation of the rime in Figure 9. Thus the way a model in figure number N-1 performs on rimes is equivalent to the way a model in figure number N performs on bodies.

Therefore, if we want to see whether one of the above models produces the rime-body asymmetry, we can simply compare its performance on rimes to the preceding model's performance on rimes. Table 14 shows performance of the distributed baseline model, the parseability/salience model, the part→whole connectivity model, and the part→part connectivity model on C, V, and rime stimuli after Stage I and after Stage II. Only generalization trials are shown.

	Distributed		Parseability-based		Part→whole		Part→part	
	Stage I	Stage II	Stage I	Stage II	Stage I	Stage II	Stage I	Stage II
Coda	18	5	22	14	23	17	25	17
Nucleus	18	7	20	9	21	10	24	12
Rime	-36	-12	-36	11	-36	19	-41	16

Table 14. Accuracy scores for generalization trials for the distributed baseline model, the parseability/salience model, the part→whole connectivity model, and the part→part connectivity model.

Table 14 shows that associability of the whole rises if the resting activation level of the whole is increased (compare the parseability-based model to the distributed model). Therefore, a possible account for why rimes are more associable than bodies is that rimes are more salient, or are more likely to be parsed out of the acoustic signal, than bodies.

Stronger part→whole connections also make a whole more associable, as shown by the fact that the rime is more associable under the part→whole model than under the parseability-based model. Therefore, a possible account of the high associability of rimes relative to bodies is that the rime node receives more activation from the segment nodes than the body node does.

Finally, part-part connections also influence associability of the whole if the parts are connected to the whole and thus can spread activation to it. However, strengthening part-part connections actually decreases the predicted accuracy on wholes after Stage II, as shown by the comparison of the part→whole model to the part→part model. This prediction is incorrect as is the prediction that subjects should be less accurate on rimes (whose parts are strongly interconnected) than on bodies (whose part-part connections are weak) following Stage I. In actuality, accuracy on rimes is higher than accuracy on bodies following both Stage I and Stage II. This finding indicates that the difference in associability between bodies and rimes cannot be accounted for between-segment connection strength.

It is important to note that to make a whole more associable than either of its parts, the activation level of the whole need not exceed the activation level of the parts. Associability of the whole in our model when the whole is allocated a separate node is equal to

$$(18) V_w = (13V_C + 11V_V + 23A_w) / 23 = 0.57V_C + 0.48V_V + A_w$$

Thus, for V_w to be greater than V_C and V_V ,

$$(19) \begin{aligned} 0.57V_C + 0.48V_V + A_w &> V_C \\ 0.57V_C + 0.48V_V + A_w &> V_V \end{aligned}$$

Since V_C and V_V are approximately equal to, respectively, A_C and A_V if the generalization test on parts does not involve wholes that have been extensively presented during training, the relations in (20) need to hold for the whole to be more associable than the parts. These do in fact hold in our experiment.

$$(20) \begin{aligned} A_w &> 0.43A_C - 0.48A_V \\ A_w &> 0.52A_V - 0.57A_C \end{aligned}$$

Interestingly, increasing associability of the whole also increases associability of the parts. Thus, predicted accuracy on consonants and vowels is higher under the parseability-based model than under the distributed model, and higher still under the part→whole model. This is because rimes presented during training on parts may be repeated during testing. This leads to the incorrect prediction that units that are part of a larger constituent should be more associable than parts that are not. If anything, the opposite is true in the experimental data: the coda is somewhat less associable than the onset if local associations are considered (coda-suffix associations are weaker than onset-prefix associations). In the next section, we consider ways of eliminating this problem by introducing top-down inhibition or loss of sent-off activation, which can be considered alternative ways of implementing between-level competition.

Decreasing Associability of Parts of Constituents

Figure 11 presents a way to modify the parseability-based model to account for the finding that parts of constituents are somewhat less associable than units that are not part of a constituent. The mechanism involves inhibitory connections from the whole to the part. If constituents have a higher resting activation level than non-constituents, the amount of inhibitory activation sent to the parts will be greater if the whole is a constituent than if it is not. An alternative way to reduce associability of constituent parts is to make parts lose the activation that they send off to the constituent node.

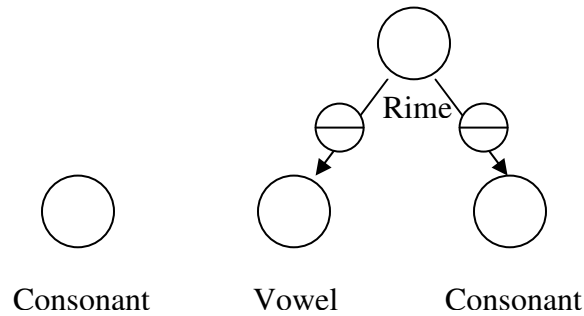


Figure 11. The parseability-based model with top-down inhibition.

Table 15 shows the predicted accuracy scores for the parseability-based model with and without top-down inhibition and the part→whole connectivity-based model with and without loss of sent-off activation. We do not consider the part→part connectivity-based model because the extra activation sent off by parts of a constituent is simply given to the other part of the constituent in the model. If the sent-off activation is lost to the sending node, there is no net increase in activation flowing into the constituent node relative to a model that does not have part-part connections, i.e., where the strength of the within-constituent part-part connections is zero.

	Parseability-based				Part→whole			
	with		without		with		without	
Stage	I	II	I	II	I	II	I	II
C	19	13	22	14	20	16	23	17
V	17	8	20	9	17	9	21	10
W	-31	13	-36	11	-30	19	-36	19

Table 15. Accuracy scores for generalization trials for the models with vs. without between-level competition

Table 15 shows that between-level competition effectively eliminates the problem of increased associability of parts following from greater associability of the whole. Models that include between-level competition and differences in activation level between constituents and non-constituents due to either resting activation level of the node representing the whole or the strength of part→whole connection correctly account for the experimental data. In these models, constituents are more associable than non-constituents but constituent parts are somewhat less associable than units that are not part of a larger constituent. An alternative possibility is to claim that the onset is inherently more associable than the coda due to temporal precedence. The two alternative hypotheses make different predictions regarding relative associability of onset and coda in syllables in which the body is a constituent while the rime is not.

General Discussion

Explanations for Differences in Constituency

A number of factors conspire to ensure that the rime is more cohesive than the body in the present stimuli, including 1) overall statistics of segment co-occurrence within English syllables, 2) statistics of segment co-occurrence within the particular syllables used in the present study, 3) potential lexicality, a.k.a., possible word status, and 4) actual lexicality. Thus, nuclei tend to co-occur with codas more than with onsets in English, and the rimes (/æʃ/ and /ʌg/) used in the present study are more frequent than the bodies (/ʃæ/ and /gʌ/).

In addition, /æf/ and /ʌg/ are well-formed words of English whereas /ʃæ/ and /gʌ/ are neither words nor well-formed in that they cannot stand on their own. It may be the case that for a difference in associability to be observed between a pair of segment strings, the segment strings must differ in potential or actual lexicality. For instance, Norris et al. (1997) and Cutler et al. (2001) argue that effects of syllable structure in monitoring tasks where subjects tend to detect syllables more easily than segment strings that cross syllable boundaries are due to the fact that syllables are possible words in the language and segment strings crossing constituent boundaries are often not. Thus, it may be the case that two segment strings that differ in linguistic constituency but do not differ in lexicality may not differ in associability either. Consequently, we would not be able to claim that such strings differ in parseability or salience and the difference in constituency would be more appropriately modeled by a difference in between-segment connection strength. Thus, it is important to manipulate lexicality, potential lexicality, and between-segment co-occurrence in future research.

Another possibility is that, while constituents are more parseable than non-constituents of the type investigated here, they are not parsed out in the course of normal language processing but can be parsed out if needed. A possible alternative to the present static models of constituent structure in which constituents are processing units (e.g., Mehler, 1981) is a dynamic distributed model that would dynamically create constituent nodes if they are needed for a learning task. Equipped with the assumption that a node is easier to create if the chunk it represents consists of two strongly connected parts than if it contains of weakly interconnected parts, such a model could potentially explain the results.

This model predicts that a constituent node will only be formed when it is needed for a configural learning task, i.e., a task that requires the subjects to associate the whole with a different response than either of its parts is associated with. Therefore, if subjects were exposed to a set of syllables sharing rimes paired with a particular affix, they would not be expected to form rime-affix associations, forming segment-affix associations instead. As a result, at least if syllable familiarity effect is not obtained, there should not be any rime familiarity effect. Thus, generalization to novel Cæf stimuli is no easier from a list of Cæf-mɪn training items than from a list of CVʃ-mɪn and CæC-mɪn training items that do not include Cæf-mɪn stimuli under this model.

Parseability/Salience vs. Part→Whole Connectivity

The difference in resting activation levels between rimes and bodies can have two manifestations: rimes should be more likely to be parsed out of the signal than bodies and, a body that has been parsed out of the signal should be less salient than a rime that has been parsed out.

The idea that a salient unit is more associable than a less salient unit has a long history in associative learning (e.g., Bush & Mosteller, 1951; Rescorla & Wagner, 1972). For instance, Rescorla and Wagner (1972) propose that associability of a stimulus is determined largely by how surprising it is (as well as how surprising its co-occurrence with the other stimulus is). It appears to be highly plausible that stimuli one attends to should be easier to associate than stimuli one does not attend to. While the high frequency of a stimulus makes its processing faster, it also makes its occurrence more predictable, thus reducing attention to the stimulus, making it less surprising.

However, saying that frequency reduces salience of a stimulus assumes that the stimulus is parsed out of the environment (cf. Kamin, 1969). One may be more likely to parse something out of the environment if its parts frequently co-occur, i.e., if it has high frequency (see Lee, 2006, for evidence regarding syllabic constituency). Because one can only associate something that one has perceived/parsed out, frequency is expected to also have positive effects on associability in the part of the frequency continuum that is below the part where it has negative effects.

The inverted U-shaped effect of predictability on processing speed and accuracy has been found in multiple linguistic domains. Several recent studies report U-shaped frequency effects on speed of word

recognition and production (Balota et al., 2004; Bien et al., 2005; Tabak et al., 2005) based on large-scale multiple regression analyses of existing collections of experimental data. Kapatsinski and Radicke (2007) find a U-shaped predictability effect in auditory particle recognition: *up* is detected more easily in medium-frequency verb-particle combinations than in low-frequency and high-frequency ones.

If constituency increases parseability, rimes are more likely to fall onto the part of the parseability continuum where increases in frequency are detrimental for associability than bodies are. Whether or not the model actually predicts a negative correlation between frequency and associability for rimes depends on whether the rimes are expected to reach a ceiling on the parseability continuum at a certain non-maximum point on the frequency continuum.

By contrast, the model that assumes that constituency influences part→whole connectivity predicts that part-whole co-occurrence should have a monotonic positive correlation with associability, under the standard Hebbian assumption that high frequency of co-occurrence between parts and wholes strengthens connections.

A Specificity-based Alternative

All of the models discussed above assume that the differences between constituents and non-constituents stem from how easy it is to extract constituents vs. non-constituents during an encounter with a linguistic signal that contains it. A major alternative explanation is that the differences in associability come from how easy it is to detect that two words share the same rime as opposed to how easy it is to detect that they share the same body.

If listeners do not automatically categorize incoming speech into phonemic categories automatically (e.g., Port & Leary, 2005), and variations in the coda have a greater impact on vowel quality than variations in the onset, the equivalence of different tokens of the same rime may be easier to detect than the equivalence of different tokens of the same body (cf. Geudens et al., 2005). As a result, acquiring rime associations would be easier than acquiring body associations. Thus, if phonemic perception is not assumed, the results seem consistent with a full-listing approach in which there are no differences between bodies and rimes, except for how acoustically similar the tokens of a given chunk type are to each other. The specificity-based approach may be able to account for the results of Lee (2006) as well. Given that there are more codas than onsets in Korean, a given rime is more acoustically variable than a given body, leading Koreans to treat the body as a constituent. Assuming that categorization is easier when the category has many exemplars, one would also predict that high-frequency units should be more associable.

In addition, generalization of rime associations to novel syllables would be easier than generalization of body associations simply because the generalization stimuli would be (subphonemically) more similar to the training stimuli in the rime condition relative to the body condition (cf. Cutler et al., 2001). However, this hypothesis necessarily predicts incomplete generalization in both conditions, while we have observed full generalization of rime dependencies and partial generalization of body dependencies.

In order to obtain this pattern of results, one would have to propose that the phonemic-level category to which new tokens of a body are compared is influenced by the tokens of that body previously encountered in the experiment more than the phonemic-level category to which new tokens of a rime are compared is. That is, the representation of the rime on the phonemic level would be more stable and less dependent on recent experience.

Categories are more stable when they have a large number of members. Thus, a prediction this model makes is that complete generalization is more likely for dependencies involving high-frequency units. Rimes are more frequent than bodies in the present stimuli. However, if the difference in

generalization is found when frequency is controlled, the categorization-based model would have to adopt the traditional model's assumption that, at least at the phonemic level, rimes are more likely to be parsed out of the signal than bodies are. In that case, the only difference between this model and the traditional model would be in what is stored at the subphonemic level with essentially the same explanation for constituency effects.

The Garner interference paradigm (Garner, 1974; Garner & Felfoldy, 1970) provides a possible further source of evidence on this hypothesis. In this paradigm, the subject is asked to classify stimuli along one dimension while some other dimension is either held constant, varies randomly, or is correlated with the attended dimension. If the subject cannot separate the dimensions, i.e., cannot attend to the dimension that is relevant for categorization without attending to the other, random variation on the 'unattended' dimension may make classification along the relevant dimension harder than if all stimuli have the same value on the 'unattended' dimension.

In our case, the subphonemic perception hypothesis predicts that since the vowel is supposed to be influenced by the coda more than by the onset, random variation in the coda should make classifying stimuli based on the vowel harder than random variation in the onset. Furthermore, given that onsets and codas do not differ significantly in associability, ease of vowel classification should account for most of the variation in accuracy and reaction time that the rime-body distinction accounts for.

Consonants vs. Vowels

Vowels were found to be less associable than consonants. These results are consistent with previous findings of Bonatti et al. (2005) who found that French speakers can learn statistical dependencies between non-adjacent consonants better than they can learn dependencies between non-adjacent vowels. However, it is not clear what is responsible for these results.

One possibility is that there is more variation in vowels than in consonants (Creel et al., 2006). This could make vowel associations harder to acquire since the difference between the speaker's and the listener's vowel categories may be greater than the difference between their consonant categories. In addition, the speaker's vowel productions may vary more than his consonant productions. One possible way to control for this factor is to have the stimuli synthesized anew for each listener recalibrating the synthesizer to match the listener's vowels.

Another possibility is that consonants are more associable because there are more irrelevant consonant types than vowel types: the consonants occurring in training are: /b/, /d/, /g/, /θ/, /ð/, /s/, /z/, /f/, /v/, /ʃ/, /ʒ/, /dʒ/, and /tʃ/. By contrast, the only vowels that occur are /i/, /ɪ/, /o/, /u/, /æ/, and /ʌ/. As a result, the subjects might have a harder time noticing which vowels are predictive (/æ/ and /ʌ/) since all vowels are relatively frequent in the experiment. Conversely, given that a vowel occurs in a greater variety of contexts than a consonant does, it may be harder to notice that all tokens of the vowel belong to the same phoneme. Equalizing the numbers of consonants and vowels used as distractors in the experiment may eliminate the consonant-vowel asymmetry. However, Creel et al. (2006) investigated this issue by teaching native English speakers artificial lexicons consisting of VCVC or CVCV words and found that words sharing consonants were more confusable for English speakers than words sharing vowels even if the number of vowels in the artificial language exceeded the number of consonants.

Related Methods

XOR learning is a subtype of configural learning. That is, it is one of multiple paradigms in which associations of the whole are not predictable from associations of the parts. Other configural learning paradigms involve the classical XOR distribution outlined in the introduction as well as negative patterning and biconditional discrimination.

In negative patterning, subjects need to learn that presentation of either stimulus A or stimulus B on its own is followed by a reward while the presentation of A and B together is not (Woodbury, 1943), as shown in Figure 12. The argument is that if subjects succeed at the task, they must be treating AB as not just A and B next to each other. In other words, there is a node corresponding to AB (Wagner, 1971; Pearce, 1987, 1994).

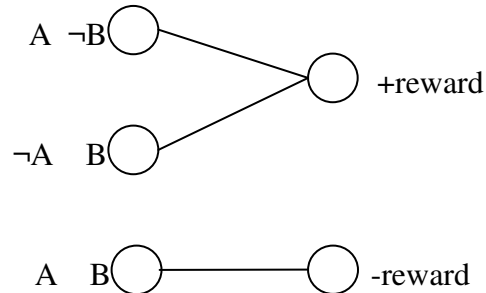


Figure 12. Hypothetical representation of stimulus associations in a negative patterning task: when A and B are presented together, +reward and -reward are activated equally strongly, cancelling each other.

In biconditional discrimination, introduced by Saavedra (1975) and illustrated in Figure 13, four stimuli (A, B, C, and D) are arranged into four compounds (AB, CD, AC, BD), two of which are associated with response X and two with response Y. Note that the compounds are arranged in such a way that any individual cue (A, B, C, and D) is just as likely to be paired with X as with Y. The discrimination is easy to encode if AB, CD, AC, and BC have a node corresponding to each of them as shown in Figure 5.

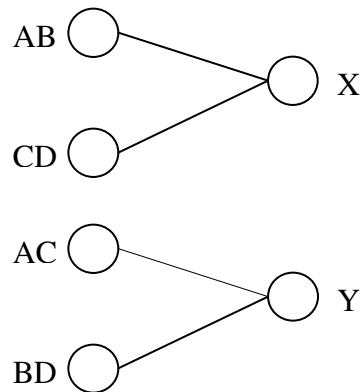


Figure 13. Hypothetical representation of stimulus associations in biconditional discrimination: AB and CD are associated with response X while AC and BD are associated with response Y.

The advantage of biconditional discrimination compared to the present task is that in the present paradigm there needs to be a large amount of variability in each position within the stimulus (e.g., the vowel /æ/ needs to be paired with a wide variety of codas to ensure that subjects learn /æ/-affix rather than rime-affix associations). It is sometimes impossible to create sufficient variability. For instance, if one wanted to examine the nature of constituency in article-noun strings in English, there is only a limited number of articles that could be paired with a noun. In biconditional discrimination, the parts of a whole whose existence is being tested need not be associated with anything. Therefore, one could compare, for instance, the learning of 'the cat-Y, a dog-Y, a cat - X, the dog -X' to the learning of 'cat the-Y, dog a-Y,

cat a - X, dog the - X' by exposing subjects to sentences like 'He gave the cat a blanket'. The disadvantage of biconditional discrimination is that it does not provide a way to examine associability of wholes and associability of parts within a single experiment.

The issue of psychological reality of complex units can also be addressed through typological research by comparing the frequencies of patterns whose acquisition requires configural learning to distributions that can be acquired elementally. If two units are likely to be chunked together, then learning distributions that require such chunking should not be much harder than learning distributions that do not require the units to be chunked. For instance, Pertsova (2007) has looked at syncretism distributions in personal pronouns across a sample of world languages, focusing specifically on cases in which the first and second person, singular and plural are represented by only two pronouns. Given this restriction, only the arrangements in Table 18 are possible. Pertsova found that distributions that require associating each pronoun with a combination of a person feature and a number feature ("person-number pronouns" in the table) were less frequent than distributions in which each pronoun could be associated with either just a person feature or just a number feature. In the present framework, these results suggest that person and number features are unlikely to be chunked together into a single complex feature unit since if they were, configural learning would be just as easy as elemental learning.

	Number Pronouns		Person Pronouns		Person-number Pronouns	
	Singular	Plural	Singular	Plural	Singular	Plural
1 st Person	X	Y	X	X	X	Y
2 nd Person	X	Y	Y	Y	Y	X

Table 16. Possible syncretism patterns in the domain of 1st and 2nd person personal pronouns, given that there are only two pronouns (X and Y) in the domain. Based on Pertsova (2007).

There are other ways to approach the issue in addition to configural learning methods. Wilson (2006) describes what he calls the "poverty of the stimulus method" (PSM), which involves exposing learners to input that is consistent with more than one possible generalization. By looking at which generalization subjects choose, one can determine which one is more natural.

It appears possible to apply this method to the study of constituency in the following way. If rimes are units and bodies are not, one can expect that given a choice between an onset-affix association and a body-affix one, the subjects would make the onset-affix association, while given the choice between a coda-affix association and a rime-affix association, the subjects would choose the rime-affix association (cf. Shanks et al., 1998, for a non-linguistic example). Such an experiment would also be helpful to test for possible biases in favor of generalizations made on the basis of larger units.

There are at least two possible reasons for such a bias: 1) generalizations based on larger units are more likely to be appropriately constrained (e.g., Williams et al., 1994) and 2) syllabic constituents may be easier to extract from syllables than segments because syllables are more similar to syllabic constituents than to segments (McNeill & Lindig, 1973). For instance, McNeill and Lindig showed that when subjects are presented with a sequence of words, it is easier for them to detect syllables than segments in the words. However, given that listeners usually do not hear subsyllabic units in isolation, syllabic constituents would be expected to be more detectable during normal speech perception than segmental units in general, making distributed models of syllabic constituency doubtful.

In order to test this hypothesis, all training tokens might involve the same rime or body while testing tokens would include novel syllables with either novel or familiar rimes or bodies. If subjects make a rime-based association, they should display imperfect generalization to novel rimes. By contrast, perfect generalization to novel bodies might be expected.

Another potential way to test associability is to expose subjects to training in which the presence of a C and a V as either the rime or the body is associated with a certain outcome, e.g., both *ka-* and *-ak* would be associated with mɪn. Then one can test generalization to novel syllables containing either *ka-* or *-ak*. We expect more generalization to syllables with the known rime than to those with a known body.

Another prediction of constituency is that generalization should be more likely when the generalization stimuli share the relevant constituent with the training stimuli. That is, VC rime associations should be hard to transfer to testing items that contain the same VC sequence that is not the rime in the testing item. Thus, there should be incomplete generalization of -æf-mɪn associations to CæfV testing items.

Finally, a possible way to examine differences between tree-based and dependency-based models of constituency is to look for pairs of stimuli that have the same structure under one view and different structures under the other view. If such a pair is found, one can look for the presence of structural priming between the stimuli. If structural priming is observed, the view of constituency that assigns the same structure to the two stimuli is supported (see Snider, 2007, for an example from syntax).

Conclusion

This paper introduces a domain-general method for implementing and testing models of constituency. We have applied this method to a test case in which multiple factors conspire to make constituency particularly clear and uncontroversial, the case of English syllables containing a lax vowel. In this particular case, the constituency difference corresponds to a difference in associability. That is, a constituent is easier to associate with an affix than a non-constituent of the same length regardless of whether the affix is a suffix or a prefix. We have also shown that the effect is not explained by similarity relations between syllables because subjects were found to be as accurate with novel syllables as with familiar ones. Finally, subjects assigned to different groups were found to be equally good learners and there were no significant differences in associability between the parts that were not shared between a constituent and a non-constituent. We have examined a number of models of constituency to determine which of them could account for the findings. Under the assumption of phonemic speech perception the data were best accounted for by localist models that assume an activation level difference between constituents and non-constituents, implemented either as a difference in parseability, or in part→whole connection strength. Thus, rimes of English syllables with lax vowels are more salient than bodies to English speakers. It is an open question whether all proposed linguistic constituents are more associable than comparable non-constituents and whether all types of constituency are profitably modeled by a localist, tree-structural representation. XOR learning, and configural learning more generally, provides a method to address this question.

References

- Alvarez, C.J., Carreiras, M., & Perea, M. (2004). Are syllables phonological units in visual word recognition? *Language and Cognitive Processes, 19*, 427-52.
- Anderson, J.M., & Ewen, C.J. (1987). *Principles of Dependency Phonology*. Cambridge, UK: Cambridge University Press.
- Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes, 19*, 391-426.
- Balota, D.A., Cortese, M.J., Sergent-Marshall, S.D., Spieler, D.H., & Yap, M.J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*, 283-316.
- Bendrien, T.A. (1992). Sound similarity judgments in English CVC's. BA Honors Thesis: U of Alberta.

- Benki, J. (2003). Quantitative evaluation of lexical status, word frequency, and neighborhood density as context effects in spoken word recognition. *Journal of the Acoustical Society of America*, 113, 1689-1705.
- Bien, H., Levelt, W.M.J., & Baayen, R.H. (2005). Frequency effects in compound production. *Proceedings of the National Academy of Sciences*, 102, 17876-81.
- Bonatti, L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, 16, 451-9.
- Boothroyd, A., & Nittrouer, S. (1988). Mathematical treatment of context effects in phoneme and word recognition. *Journal of the Acoustical Society of America*, 84, 101-14.
- Bush, R.R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, 58, 313-23.
- Bybee, J., & Brewer, M.A. (1980). Explanation in morphophonemics: Changes in Provençal and Spanish preterite forms. *Lingua*, 52, 201-42.
- Carreiras, M., & Perea, M. (2002). Masked priming effects with syllabic neighbors in a lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1228-42.
- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33A, 497-505.
- Creel, S.C., Aslin, R.N., & Tanenhaus, M.K. (2006). Acquiring an artificial lexicon: Segment type and order information in early lexical entries. *Journal of Memory and Language*, 54, 1-19.
- Cutler, A., McQueen, J.M., Norris, D., & Somejuan, A. (2001). The roll of the silly ball. In E. Dupoux (ed.), *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler* (pp.181-94). Cambridge, MA: MIT Press.
- Davis, S. (1989). On a non-argument for the rhyme. *Journal of Linguistics*, 25, 211-5.
- Derwing, B.L. (1987). A cross-linguistic experimental investigation of syllable structure. Part I: Background and methodology. *Proceedings of the Third Annual Meeting of the Pacific Linguistics Conference*, 93-102.
- Derwing, B.L., & Nearey, T.M. (1986). Experimental phonology at the University of Alberta. In J. Ohala (ed.), *Experimental phonology* (pp.187-209). San Diego: Academic Press.
- Felty, R. 2007. Context effects in spoken word recognition of English and German by native and non-native listeners, Unpublished Doctoral Dissertation, University of Michigan.
- Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of words and picture naming: The role of syllabic units. *Journal of Memory and Language*, 35, 708-23.
- Fudge, E. (1987). Branching structure within the syllable. *Journal of Linguistics*, 23, 359-377.
- Garner, W.R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.
- Garner, W.R., & Felfoldy, G.L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, 1, 225-41.
- Geudens, A., Sandra, D., & Martensen, H. (2005). Rhyming words and onset-rime constituents: An inquiry into structural breaking points and emergent boundaries in the syllable. *Journal of Experimental Child Psychology*, 92, 366-87.
- Hall, G. (2003). Learned changes in the sensitivity of stimulus representations: Associative and nonassociative mechanisms. *Quarterly Journal of Experimental Psychology*, 56B, 43-55.
- Hockema, S.A. (2006). Finding words in speech: An investigation of American English. *Language Learning and Development*, 2, 119-46.
- Jaeger, J.J. (2005). *Kids' slips: What young children's slips of the tongue reveal about language development*. Mahwah, NJ: Erlbaum.
- Kamin, L.J. (1969). Predictability, surprise, attention and conditioning. In B.A. Campbell & R.M. Church (eds.), *Punishment and aversive behavior*. (pp. 279-296). New York: Appleton-Century-Crofts.
- Kapatsinski, V.M. (2006). Frequency and cohesion: Evidence from repair. Paper presented at High Desert Linguistics Society VII. Albuquerque, NM.
- Kapatsinski, V.M., & Radicke, J. (2007). Frequency and the emergence of prefabs: Evidence from monitoring. This volume.

- Kessler, B., & Treiman, R. (1997). Syllable structure and the distribution of phonemes in English syllables. *Journal of Memory and Language*, 37, 295-311.
- Lee, Y. (2006). Sub-syllabic constituency in Korean and English, Unpublished Doctoral Dissertation, Northwestern University.
- McNeill, D., & Lindig, K. (1973). The perceptual reality of phonemes, syllables, words, and sentences. *Journal of Verbal Learning and Verbal Behavior*, 12, 419-30.
- Mehler, J. (1981). The role of syllables in speech processing: Infant and adult data. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 295, 333-52.
- Minsky, M.L., & Papert, S.A. (1969). *Perceptrons*. Cambridge, MA: MIT Press.
- Moder, C.L. (1992). Productivity and categorization in morphological classes, Unpublished Doctoral Dissertation, SUNY Buffalo.
- Nearey, T. (1990). The segment as a unit of speech perception. *Journal of Phonetics*, 18, 347-73.
- Nearey, T. (2003). On the factorability of phonological units in speech perception. In R. Ogden, J. Local, & R. Temple (eds.), *Laboratory Phonology 6*. (pp. 197-221). Cambridge, UK: Cambridge University Press.
- Nelson, D.L., & Nelson, L.D. (1970). Rated acoustic (articulatory) similarity for word pairs varying in number and ordinal position of common letters. *Psychonomic Science*, 19, 81-2.
- Newport, E.L., & Aslin, R.N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48, 127-62.
- 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.
- Nusbaum, H.G., Pisoni, D.B., & Davis, C.K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. In *Research on Speech Perception Progress Report No. 10* (pp. 357-76). Bloomington, IN: Speech Research Laboratory, Indiana University.
- Pearce, J.M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, 94, 61-73.
- Pearce, J.M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, 101, 587-607.
- Pertsova, K. (2007). Towards learning form-meaning correspondences of inflectional morphemes. Paper presented at LSA Annual Meeting, Anaheim, CA.
- Port, R.F., & Leary, A.P. (2005). Against formal phonology. *Language*, 81, 927-64.
- Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black & W.F. Prokasy (eds.), *Classical conditioning II: Current research and theory* (pp.64-99). New York: Appleton Century Crofts.
- Saavedra, M.A. (1975). Pavlovian conditioning in the rabbit. *Learning and Motivation*, 6, 314-26.
- Selkirk, E.O. (1982). The syllable. In H. Van der Hulst & N. Smith (eds.), *The structure of phonological representations (part II)* (pp.337-83). Dordrecht: Foris.
- Skousen, R. (1989). *Analogical modeling of language*. Dordrecht: Kluwer.
- Snider, N. (2007). Evidence from priming for hierarchical representation in syntactic structure. Paper presented at LSA Annual Meeting, Anaheim, CA. Available online at <http://www.stanford.edu/~snider/pubs/snider,lsa,2007.pdf>.
- Stemberger, J. (1983). *Speech errors and theoretical phonology: A review*. Bloomington: Indiana University Linguistics Club.
- Tabak, W., Schreuder, R., & Baayen, R.H. (2005). Lexical statistics and lexical processing: semantic density, information complexity, sex, and irregularity in Dutch. In M. Reis & S. Kepsers (eds.), *Linguistic evidence: Empirical, theoretical and computational perspectives* (pp.529-55). Berlin: Mouton de Gruyter.
- Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition*, 15, 49-74.
- Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language*, 25, 476-491.

- Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 145-52.
- Treiman, R., Kessler, B., Knewasser, S., Tincoff, R., & Bowman, M. (2000). English speakers' sensitivity to phonotactic patterns. In M.B. Broe & J.B. Pierrehumbert (eds.), *Laboratory Phonology V: Acquisition and the Lexicon* (pp.269-82). Cambridge, UK: Cambridge University Press.
- Vennemann, T. (1988). The rule dependence of syllable structure. In C. Duncan-Rose & T. Vennemann (eds.), *On language: Rhetorica, phonologica, syntactica: A festschrift for Robert P. Stockwell from his friends and colleagues* (pp.257-83). London: Routledge.
- Vitz, P.C., & Winkler, B.S. (1973). Predicting the judged "similarity of sound" of English words. *Journal of Verbal Learning and Verbal Behavior*, *12*, 373-88.
- Wagner, A.R. (1971). Elementary associations. In H.H. Kendler & J.T. Spence (eds.), *Essays in Neobehaviorism: A memorial volume to Kenneth W. Spence* (pp.187-213). New York: Appleton-Century-Crofts.
- Williams, D.A., Sagness, K.E., & McPhee, J.E.. (1994). Configural and elemental strategies in predictive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 694-709.
- Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, *30*, 945-82.
- Woodbury, C.B. (1943). The learning of stimulus patterns by dogs. *Journal of Comparative Psychology*, *35*, 29-40.
- Yoon, Y.B., & Derwing, B.L. (2001). A language without a rhyme: Syllable structure experiments in Korean. *Canadian Journal of Linguistics*, *46*, 187-237.

Appendix: Instructions

Learning Toy Languages Instructions

There are three parts to this experiment.

Within each part, you will first listen to words paired with either 'min' or 'noom'. Whether you hear 'min' or 'noom' depends on what the other word sounds like. This will last for about three minutes.

Then you will be presented with some more word pairs in which the min/noom part has been replaced by noise. You will be asked to guess whether the noise replaces 'min' or 'noom' by pressing a button on the button box. Once you have made your guess, the correct answer will be revealed to you. PLEASE, TRY TO RESPOND AS ACCURATELY AS POSSIBLE. This will last for about five minutes.

Finally, you will be presented with some more words paired with noise. You will again be asked to guess whether the noise replaces 'min' or 'noom'. This time, though, you will not get feedback. PLEASE, TRY TO RESPOND AS ACCURATELY AS POSSIBLE. This will take about four minutes.

This entire three-part sequence will be repeated three times. The whole experiment will last for about 40 minutes.

The following messages appeared on the screen:

Prior to each training session (each subject saw either 'preceding' or 'following'):

Whether you hear 'min' or 'noom' depends on what the preceding/following word is like.
Press any button to begin

Prior to each feedback session:

You will now be given feedback.
Once you make your guess the correct answer is going to be pronounced.

Prior to each testing session:

You will now be tested on what you know
Press any button to begin

End of experiment:

The experiment is over.
Thank you for your participation.