

# MORPHEME SEGMENTATION IN SCHOOL-AGED CHILDREN

SARA FINLEY  
ELISSA L. NEWPORT  
*University of Rochester*

## 1 Introduction

One of the major questions in the cognitive science of language is how children learn the productive systems that encompass the grammar of our native languages. This paper addresses how children learn morphology, the systematic relationship between form and meaning. For example, the English plural can be made in one of many ways, most commonly with the addition of the suffix /-s/. While such patterns between sound and meaning are subject to exceptions, marking meanings through regular forms is one of many ways in which language is a productive, combinatorial system.

There are at least two possible ways in which morphology could be learned. One possibility is that the learner acquires relationships between forms and their meanings by grouping words that share a common component of meaning and then determining what part of the form is used to signal this meaning. For example, to acquire the English plural, the learner would take known plural words in the lexicon (*dogs, cats, bugs, buses*) and extract the commonality that the majority of these words end in the letter ‘s’ to form the rule that ‘s’ marks plural<sup>1</sup>. The other possibility is that the learner works in the opposite direction, observing that a number of words in the lexicon have commonalities of form and inferring a morphological or semantic relationship among words that sound similar. Discovering the regularities among forms, the learner can also determine the system that relates form to meaning. For example, the words *dogs, cats, bugs, buses* all end in ‘s’, and are therefore likely to be related in meaning. These routes to discovering morphological regularities have different advantages and disadvantages. Using semantic relations to discover regularities in form requires knowing the meanings of the words, so the learner would have to master several sets of word meanings before acquiring their morphological

---

<sup>1</sup> This is a simplification, as phonetic sound of plurals in English vary between /s/, /z/ and /əz/.

patterns. However, because on this route the learner knows the meaning of words, it is less likely that the learner will mis-parse words (e.g., *fuss* as a plural).

Previous research addressing how morphology might be learned has suggested that distributional cues alone are not sufficient to learn linguistic categories and sub-categories. Rather, they have suggested that phonological (Brooks, Braine, Catalano, & Brody, 1993; Frigo & McDonald, 1998; Gerken, Wilson, & Lewis, 2005) and semantic cues (Braine et al., 1990) must serve as the foundation to learning the morphological regularities of language (MacWhinney, Leinbach, Taranan, & McDonald, 1989; Maratsos & Chalkley, 1980). These studies suggest that it is impossible to learn morphologically marked categories without additional cues to category structure (Gomez & Gerken, 2000). However, recent evidence suggests that learners can use distributional information to acquire categories and subcategories, as long as the distributional regularities are rich enough (Reeder, Newport, & Aslin, 2009, 2010).

Outside of studies of category formation, research examining the learning of morphological patterns in language is relatively scarce. Much of the previous research focusing on morphology acquisition has tested how children learn the meaning of a morpheme rather than how children learn the systematic structure of the morphological patterns (Braine et al., 1990; MacWhinney, 1983). Braine et al. (1990) taught children inflectional locative affixes (e.g., *to*, *from*, *at*) in an artificial grammar learning setting. In this case, learning the form of the affix was dependent on the semantic context associated with the form. However, when form and meaning are coupled, it is impossible to differentiate between difficulties in learning the meaning of the affixes and difficulties in learning the sound patterns associated with those affixes.

The present study explores how morphological patterns are learned independently of meaning. One might think that without associated meanings, ‘morphological patterns’ are really just phonotactic or phonological patterns. However, it is important to note that morphological patterns differ systematically from phonotactic or phonological patterns in language, apart from the meanings that they signal. The phonotactics of a language are the restrictions on co-occurrence of sounds within a word, and they dictate possible words in a language. For example, syllables in English may not begin with two adjacent stop consonants (e.g., *\*ptik* is not a possible English word). Morphological patterns make use of the phonotactic restrictions of a language in order to express meanings through phonological regularities, but are independent of the phonology of the language. For example, in reduplication, a morpheme can be expressed by repeating all or part of the base word. In Marshallese, the final syllable of a base can be copied to the end of the word to signify the intensive morpheme (e.g., [ebbok-bok] ‘puffy’) (Byrd, 1983). If this were a purely phonological rule, one would expect that all words in the language would require repetition of the first syllable in all contexts, since phonotactic or phonological processes govern the required forms of words throughout the language. In contrast, morphological patterns are variations in form that can occur in a set of words. Such morphological changes occur in a productive way (that is, these changes occur for all words of a specific type), as part of patterned variations that the words undergo. In actual languages, of course, these patterned variations would signal variations of grammatical meanings as well. But in the present study we will investigate whether learners can acquire this type of variation in form, in the absence of information about meaning.

There are two reasons why it is important to study morphological learning independently of meaning. First, understanding whether and how learners can infer morphological relatedness through form relatedness is important for understanding the structure of the lexicon as well as the restrictions on morphological systems in languages of the world. If learners hypothesize that

words with similar phonological forms might be morphologically related, they must have a way of differentiating these from the many morphologically unrelated words that also have similar phonological forms (e.g. word in English that end in 's'). We have noted above how these two are differentiated in natural languages. An important question is how learners distinguish these two kinds of similarity. While this is not the focus of the present study, it forms the basis for subsequent work on this topic. Second, infants do not begin to learn the meanings of words for many months, but they have access to the forms of the language from birth. Much of language learning may take place when the child only has access to non-semantic distributional cues to morpheme segmentation. For this reason, distributional cues should be a primary source of evidence in early morpheme segmentation; using distributional cues to find where semantic relations may lie takes advantage of the learner's ability to find patterns in the data, even before complex semantic relationships among words may be understood. In the present study, we test the hypothesis that distributional cues to morphologically related words are an important starting point to learning the morphological systems of language. We demonstrate that school-aged children are able to segment morphologically complex forms from distribution alone, without the help of semantic cues. This suggests that the distributional information that characterizes the patterns of form in morphological systems is important to the early steps in the acquisition of morphological systems.

The present study employs an artificial grammar learning paradigm to test the hypothesis that it is possible to learn morphological patterns by attending to the distributional cues of a morphologically rich language. Finley and Newport (2010 and in preparation) demonstrated that adults are able to segment morphologically complex words using the distributional cues in an artificial language. In these studies, adult listeners were exposed to a miniature language that involved words formed from a set of 24 stems and 4 affixes. Participants were able to parse the words into their stems and affixes and were able to differentiate between words that followed the affixation pattern versus words that did not.

However, because adults have more advanced cognitive abilities, greater motivation, and a larger set of test-taking strategies, it is possible that adults' ability to segment morphemes from distributional information does not translate to the child learner. It is also possible that child learners may require semantic information to learn the morphological patterns of their language, whereas adults may be able to employ additional learning strategies to use distributional cues to learn these morphological patterns. The present study uses a modified version of the artificial grammar used in Finley and Newport (2010 and in preparation) in order to examine whether child learners are able to use distributional information, without accompanying semantic information, in order to learn the morphological patterns of their language.

## **2 Methods**

Thirty-three school-aged children were recruited from afterschool programs in the Rochester, NY area. Participants were between the ages of 7 and 11. They were given stickers and bags of small toys for their participation. Eight participants were excluded because they opted not to complete the task ( $n = 6$ ) or because they did not understand the task ( $n = 2$ ). (One child repeated the words in the test rather than responding 'first' or 'second' as instructed, and the other child responded 'yes' or 'Silly Speak' to all options). Five participants did not participate in Day 2 of the experiment (because they were absent from daycare, or because they elected not to participate on the second day).

The design and procedure was based on Finley and Newport (2010 and in preparation). We created a miniature language (named Silly Speak) using the same stimulus set as Finley and Newport (2010 and Experiment 1, in preparation). The present experiment was designed to examine children's ability to segment morphologically complex words from exposure to the forms (but not meanings) of these words. The miniature languages each contained a number of morphologically complex words, but no sentences or meanings were provided. Each morphologically complex word consisted of a stem followed by a suffix. We created two sets of such items (hereafter called Languages A and B) that served to counterbalance any details of the stimuli that might inadvertently interfere with or influence learning. Each language contained 24 stems and four suffixes. Because we wanted to test participants' ability to remember the words heard in training as well as their ability to generalize the suffixing pattern to novel items, we presented each of the 24 stems with just two suffixes (with each of the four suffixes paired with twelve different stems), for a total of 48 different words. Over the language, however, all of the stems could (in principle) occur with any of the 4 different suffixes; the specific stem-suffix pairings presented during exposure did not contain any secondary patterns (subcategories) within the items. The stems were of the form CVCV and the suffixes were of the form CV, creating CV.CV.CV words. C was drawn from the set /p, t, k, b, d, g, m, n, f, v, s, z/ and V was drawn from the set /i, e, u, o/. Assignments of C and V to words were created semi-randomly under two constraints. First, English words were avoided. Second, the syllables used in the stems did not overlap with the syllables used in the suffixes. For example, Language A contained the suffix [-mu], but none of the stem items contained the syllable [mu].

(1) Example Training Items

Language A	Language B
demebu, demedo	bovepa, bovegu
fibami, fibado	gisigu, gisino
tisebu, tiseke	finase, finapa
nobokey, nobomi	vemano, vemapa

There were two different types of test items: Stem Parsing and Suffix Parsing. Three sets of test items in the Stem Parsing condition were designed to test whether participants learned that the stem was a separate unit from the affix. The first set (referred to as OldStem-NewStem) probed whether participants could recognize stems that they had heard before (AB) (where A and B refer to syllables in the training set) versus stems that they had not heard before (AD) (where D refers to a syllable heard in the training set, but not within a stem combined with A). Participants chose between a familiar item (AB-X) (where X refers to an affix in the training set that was heard with this AB stem) and a novel stem-affix combination in which the first and last syllable of the word were the same as the alternative, but the second syllable came from another word (AD-X). The second set of test items probed whether learners were more likely to view the stem AB as a unit, as compared with BX (the second syllable of the stem followed by the affix X) (referred to as New-Nonword Hybrid). In these test items we compared a familiar stem containing a familiar affix that had not appeared with that stem (AB-Y) (where Y refers to suffix that was heard in training, but not paired with that particular stem) with a non-word hybrid composed of the first syllable of a familiar stem parsed with the final two syllables of a familiar affixed word (CB-X). Thus while both BX and AB were familiar to participants, if learners parse AB as a unit, they should choose AB over BX. The third set of test items probed the ability of

## Morpheme Segmentation in School-Aged Children

learners to generalize familiar stems to novel unfamiliar affixes (referred to as NewSuffix-Old(Scrambled)). Participants chose between a stem containing a novel affix (AB-Q) (where Q a syllable not heard in training) with its scrambled counterpart (AQB). Examples of the Stem Parsing test items can be found in (2).

### (2) Example Stem Parsing Test Items

Language	OldStem- NewStem	New-NonWord Hybrid	NewAffix- Old(Scrambled)
A	tisebu-sonodo	befado-mefabu	fibasi-fimiba
B	gisino-gimono	finagu-donagu	bovebu-bopave

The Suffix Parsing test items probed whether learners parsed the affixes as separate units from the stems. These test items were given on a different day (in order to reduce fatigue), with order of testing counterbalanced across children. As with the Stem Parsing test items, there were three different sets of test items. The first set tested familiarity of words, comparing a familiar word (ABX) with a scrambled familiar item (AXB) (referred to as Old-Old(Scrambled) items). The second set of test items tested the ability to generalize the suffixing pattern to a stem+suffix combination that fit the suffixing pattern but was not heard during the exposure phase. We compared a new stem-affix combination (ABY) with a scrambled familiar item (AXB) (referred to as New-Old(Scrambled)). If learners extracted the general form of words in the language to be Stem+Affix and parsed the affixes, they should choose both the ABX and ABY items significantly above chance. If learners have extracted a rule in which all words are of the form ABX, we expect that learners should be more likely to recognize a familiar word than a grammatical but unfamiliar word. The third set of test items probed whether learners recognize an item heard before (ABX) with a stem-suffix combination never heard before (ABY) (referred to as Old-New). Examples of the Stem Parsing test items can be found in (3).

### (3) Example Test Affix Parsing Test Items

Language	Old-Old(Scrambled)	New-Old(Scrambled)	Old-New
A	nobokey-nokebo	tisemi-tiseke	demebu-dememi
B	bovepa-bopave	vemagu-vepama	gisigu-gisise

Each child received identical training on both days of the experiment, but responded to a different set of test items on each day. The instructions were presented via computer, as a recorded set of instructions presented by a cartoon alien. The experimenter verbally presented additional clarifications and reminders as necessary. The instructions were set up as a game in which the child was working to help the Silly People to communicate with Earthlings. The child repeated each word spoken out loud. After the child spoke each word, a drawing of a Silly Person would appear on the computer screen for 500ms. To keep the children motivated, after each round of 48 exposure items, the child was instructed to place a sticker next to a paper drawing of a Silly Person. There were 5 drawings of Silly People on the paper, representing 5 rounds of exposure for the set of 48 items.

When the exposure phase ended, the child was asked to play the ‘Silly Guessing Game’ in which the participant was asked to decide which of two words was Silly Speak. They were told that they would hear two words, one from Silly Speak and the other not from Silly Speak, and

their job was to tell the experimenter which item they believed was the word from Silly Speak (the first or the second).

### 3 Results

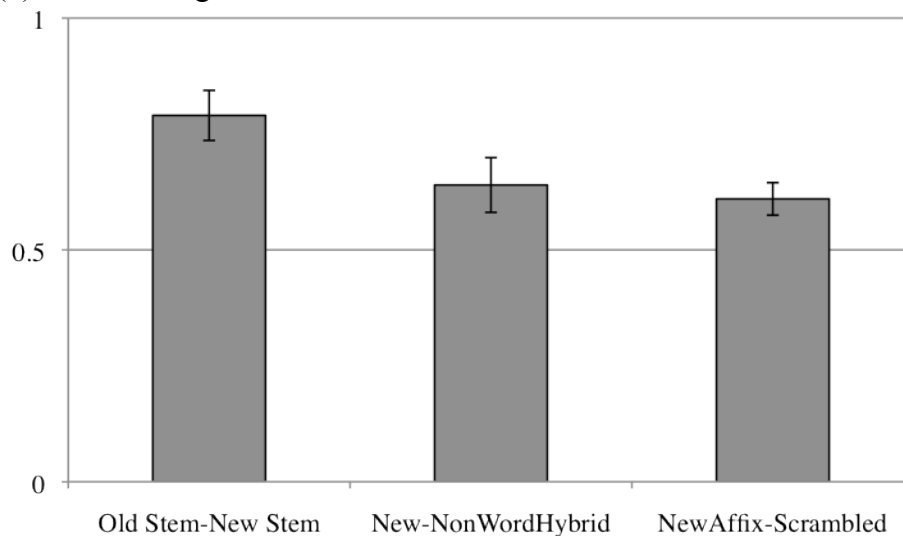
We performed analyses separately for the Stem Parsing tests and the Suffix Parsing tests.

#### 3.1 Stem Parsing

We combined the data from Languages A and B, as they were not significantly different from each other,  $F < 1$ . There was no effect of test item type,  $F(2,40) = 2.45$ ,  $p = 0.10$  and no significant interaction,  $F(2,40) = 1.85$ ,  $p = 0.17$ . Results for Stem Parsing Test items appear in (4).

We compared each of the means for the Stem-Parsing test items to chance (50%) via Bonferroni corrected one-sample t-tests. All types of test items were significantly above chance, suggesting that participants had learned to parse the stems. The NewSuffix-Scrambled items (mean = 0.61,  $CI \pm 0.071$ ) (ABQ vs. AQB) were significantly above chance,  $t(21)=3.16$ ,  $p < 0.05$ . The NewStem-OldStems items (ABX vs. ADX) were significantly above chance (mean = 0.71,  $CI \pm 0.081$ ),  $t(21)=5.40$ ,  $p < 0.001$ . The New-Non-Word Hybrid test items (ABY vs. CBX) were significantly above chance (mean = 0.67,  $CI \pm 0.081$ ),  $t(21)=3.63$ ,  $p < 0.01$ .

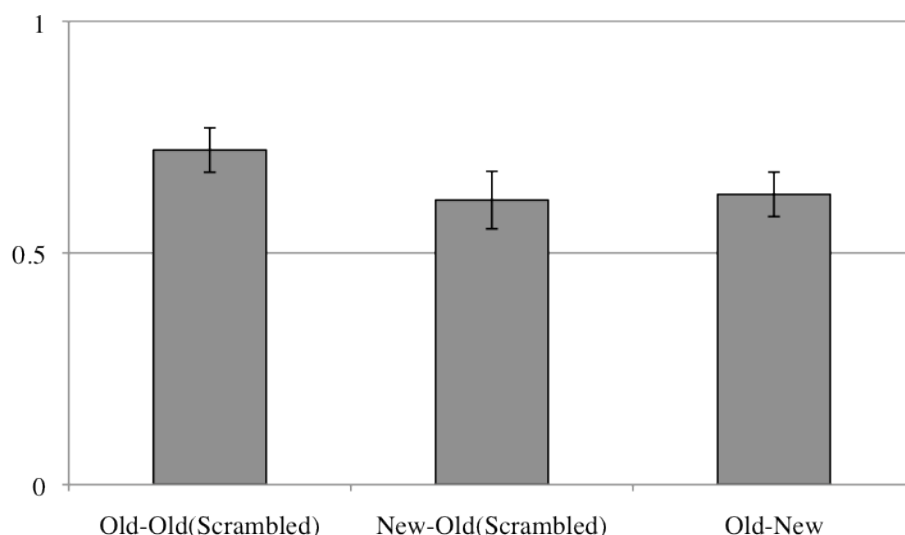
(4) Stem Parsing Results



#### 3.2 Suffix Parsing

Results for Suffix Parsing Test items appear in (5). We combined the data from Languages A and B, as they were not significantly different from each other,  $F < 1$ . There was no interaction,  $F < 1$ . There was a significant effect of Test Item type,  $F(2, 42) = 3.26$ ,  $p < 0.05$ . This reflects the fact that there were significantly fewer correct responses to Old-New and New-Old(Scrambled) test items compared to Old-Old(Scrambled) test items,  $F(1, 21) = 6.32$ ,  $p < 0.05$ .

(5) Suffix Parsing Results



We compared each of the means for the Suffix Parsing test items to chance (50%) via Bonferroni corrected one-sample t-tests. The Old-Old(Scrambled) (ABX-AXB) were significantly above chance (mean = 0.72,  $CI \pm 0.084$ ),  $t(22)=5.53$ ,  $p < 0.0001$ . The New-Old(Scrambled) (ABY vs. AXB) were significantly above chance (mean = 0.61,  $CI \pm 0.090$ ),  $t(22)=2.66$ ,  $p < 0.05$ . The Old-New test items (ABX vs. ABY) were significantly above chance (mean = 0.63,  $CI \pm 0.058$ ),  $t(22)=4.49$ ,  $p < 0.0001$ . All three types of test items were thus significantly above chance, suggesting that participants had learned to parse the suffix items. Because the Old-New items (ABX vs. ABY) were significantly above chance, it appears that learners did have a preference for familiar words over grammatical words not presented during training.

Overall, however, if participants simply responded based on familiarity to forms heard previously in training to select their answer, rather than learning a suffixing ‘rule,’ we would expect no difference between the Old-New and the Old-Old(Scrambled) test items. However, there were significantly more correct items Old-Old(Scrambled) test items compared to Old-New test items,  $t(23)=2.39$ ,  $p < 0.05$ , suggesting that learners made a distinction between items that were correct because they were heard in the training set and items that were correct because they follow a suffixing pattern. They did not simply memorize the forms heard in training and respond using word familiarity.

## 4 Discussion and Conclusions

In the present experiment, school-aged children did use the distributional information of words in an artificial language to reliably segment three-syllable words into component stems and suffixes. Learners apparently were able to infer a productive morphological pattern and did not simply memorize wordforms heard during training. Participants were substantially more likely to reject an ungrammatical form than a grammatical but unfamiliar form.

The results of the present study parallel those of previous findings with adults (Finley and Newport, 2010 and in preparation). Both children and adults were able to parse the morphological patterns using distributional cues. The adults’ results differed from the children’s

results in that adults were relatively more accurate than children on all test items (except for Old-New test items). That is, while both the adults and the children showed results greater than chance, adults' responses were slightly higher. There are two possible reasons for this. One is that the children in the present experiment were highly variable, but it is unclear whether the variation was due to learning ability or task related reasons (e.g., motivation, interest and understanding of the task). The other is that children may require longer amounts of time to learn the pattern. Adults in Finley and Newport (in preparation) were given approximately 20 minutes of exposure, which is about the same amount of time that the children were given on the two days combined.

According to the critical period hypothesis for language learning (Lenneberg, 1967), children are much more likely to reach native proficiency when learning a new language than are adults. This might lead one might expect that children should be faster at learning new languages than adults. While the children in the present study did not appear to be faster learners, there are important considerations to keep in mind. First, children were able to learn the same morphological pattern that adults learned in our previous studies. This is important because adults in Finley and Newport (2010 and in preparation) were given longer listening times and were able to rely on more extensive cognitive resources, more years of schooling, and more practiced test-taking strategies to perform well on the task. The children were at a disadvantage in all of these respects, but were still able to learn the pattern, perhaps suggesting language learning abilities that are different from the adults. Second, the exposure phases in both the adult and the child studies were very short. It is possible that children may be better language learners than adults in terms of building structural regularities over input, but this learning may not be any faster at the very initial stages (which is the focus of the present paper).

While the results of the present study demonstrate that suffixing patterns can be learned through distributional information, more work is needed to understand the precise role of this information in learning morphological systems. Finley and Newport (in preparation) demonstrated that adult learners are able to parse stems from prefixes as well as suffixes from input like that in the present experiment, but that parsing more complex morphological patterns, such as infixation and non-concatenative morphology, requires exposure to a larger set of distributional patterns. Importantly, when non-concatenative patterns were augmented to be more like natural languages (e.g., including additional affixes and larger amounts of variation), adult learners can learn to parse these complex patterns. Future work will investigate these more complex patterns in children as well.

Understanding the role of distribution in morphological learning is important for developing a theory of the mechanisms that underlie language development. As we begin to understand the biases of learners in using various statistical cues, we can uncover the ways in which these learning biases shape the patterns of languages throughout the world.

## References

- Braine, M. D. S., Brody, R. E., Brooks, P. J., Sudhalter, V., Ross, J. A., Catalano, L., et al. (1990). Exploring language acquisition in children with a miniature artificial language: Effects of item and pattern frequency, arbitrary subclasses, and correction. *Journal of Memory and Language*, 29, 591-610.
- Brooks, P. J., Braine, M. D. S., Catalano, L., & Brody, R. E. (1993). Acquisition of gender-like



- noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language*, 32, 76-95.
- Byrd, Dani. (1993). Marshallese suffixal reduplication. In J Mead (Ed.) Proceedings of the Eleventh West Coast Conference on Formal Linguistics, pp. 61-77.
- Finley, Sara, and Newport, Elissa L. (2010). Morpheme segmentation from distributional information. In Boston University Conference on Language Development (BUCLD) Online Proceedings Supplement.
- Finley, Sara, and Newport, Elissa L. (in preparation). Morpheme segmentation in artificial grammar learning. Manuscript, University of Rochester.
- Frigo, L., & McDonald, J. L. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language*, 39, 218-245.
- Gerken, L., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, 32, 249-268.
- Gomez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Science*, 4(5), 178-186.
- Lenneberg, E. H. (1967). *Biological foundations of language*. New York: John Wiley
- MacWhinney, B. (1983). Miniature Linguistic Systems as Tests of the Use of Universal Operating Principles in Second-Language Learning by Children and Adults. *Journal of Psycholinguistic Research*, 12(5), 467-478.
- MacWhinney, B., Leinbach, J., Taranan, R., & McDonald, J. L. (1989). Language learning: Cues or rules? *Journal of Memory and Language*, 28, 255-277.
- Maratsos, M. P., & Chalkley, M. A. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. Nelson (Ed.), *Children's language* (Vol. 2, pp. 127-189). NY: Gardner Press.
- Reeder, P., Newport, E.L., & Aslin, R. N. (2009). The role of distributional information in linguistic category formation. In N. Taatgen and H. van Rijn (eds), Proceedings of the 31st Annual Meeting of the Cognitive Science Society.
- Reeder, P.A., Newport, E.L., & Aslin, R.N. (2010). Novel words in novel contexts: The role of distributional information in form-class category learning. In S. Ohlsson & R. Catrambone (Eds.), Proceedings of the 32nd Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society.

## **Acknowledgments**

We are grateful to Patricia Reeder, Neil Bardhan, Neil Snider, Carrie Miller, Kelly Johnston, Lily Schieber, Anna States, and to members of the Aslin-Newport lab. We would also like to thank the children, parents, and staff at the daycare centers where we conducted this study. This research was supported in part by NIH grant DC00167 to E. Newport, HD37082 to R. Aslin & E. Newport, and NIH training grant T32DC000035.