

Verbs are LookING Good in Language Acquisition

Jon A. Willits (wilits@wisc.edu)

Department of Psychology, 1202 W. Johnson Street
Madison, WI 53706 USA

Mark S. Seidenberg (seidenberg@wisc.edu)

Department of Psychology, 1202 W. Johnson Street
Madison, WI 53706 USA

Jenny R. Saffran (saffran@wisc.edu)

Department of Psychology, 1202 W. Johnson Street
Madison, WI 53706 USA

Abstract

Statistical learning is an important element of language acquisition. A basic unresolved question is, what are the units over which statistics are calculated? In a corpus study and two infant behavioral experiments, we show that varying the units that are used greatly affects learning. Using words as units, nouns are easier to segment from continuous speech than verbs. However, if a highly frequent morphological element such as ING is also treated as a unit, noun-verb differences disappear, in both corpus analysis and behavioral studies. These results suggest that infants can compute statistics over units other than words and syllables, and theories of statistical learning need better accounts for why some units are tracked and not others.

Keywords: Statistical learning; Language acquisition; Word segmentation; Morphology; Distributional statistics

What Do Infants Count?

Studies of infants and young children have established that statistical learning plays an important role in language acquisition (see Saffran & Sahni, 2007, for a review). A fundamental question for theories of statistical learning is, what units are statistics computed over? Many statistics can be derived from natural languages, a fact that could limit the role of statistical processes in acquisition, as this could make it more difficult for the language learner to figure out what statistics to use. For researchers in the area, the problem is to identify the units that *are* tracked and to determine why these units are tracked why others are not.

The literature on statistical language learning in infants, children, and adults has often focused on transition probabilities within and between words (e.g. the probability one unit will follow another). For example, Saffran, Aslin, and Newport (1996) manipulated transition probabilities between syllables within a word (which were high) compared to probabilities between syllables at word boundaries (which were low). These statistical heterogeneities provided a basis for identifying words in a simple artificial language. Many subsequent studies have focused on transition probabilities in both artificial and natural languages. Research has also begun to look at other types of dependencies, for example between non-adjacent syllables or words.]

All such experiments make assumptions about the units over which infants encode statistics such as frequency and transition probability. Syllables, for example, seem like obvious units given their fundamental role in speech production. Different units may be tracked at different points in development. As the child's vocabulary develops, so does the possibility of tracking word-level statistics. Moreover, statistical learning may occur at multiple levels of linguistic structure simultaneously. Thus, the question as to which units statistics are computed over is a central one. The answer will affect the extent to which statistical learning is implicated in acquisition.

We examined this question in the context of a puzzle in the language learning literature. In an important study, Jusczyk and Aslin (1995) found that 7.5-month-old infants could identify nouns from fluent, continuous speech. They played infants a two-minute corpus of typical, child-directed speech that repetitively used the same two nouns, and found during a test phase that infants discriminated between the nouns that had been played and frequency matched nouns that had not been played. However, in a later study using a similar procedure, Nazzi et al. (2005) found that verbs were not identifiable until between 13.5 and 17.5 months. Thus nouns and verbs appear to differ in ease of learning.

This difference could be because of intrinsic differences between nouns and verbs: verbs could be more complex because they encode relations, the relations can involve a variety of different elements in a sentence, and these relations can be expressed in a number of different syntactic structures (Gentner, 2006). It is also possible that the statistical properties of nouns and verbs differ, such that whereas nouns can be identified based on the immediate contexts in which they occur, verbs cannot. Identifying verbs might then require the use of other information such as syllabic stress which infants master at later ages than they do transition probability (Saffran & Thiessen, 2003).

Our study investigated this idea, but with an important twist: the learnability of nouns and verbs from statistical information (the frequencies of words and the immediate lexical contexts in which they occur) crucially depends on assumptions about the units over which the child computes such statistics. In particular, we examined the role of the highly frequent bound morpheme: ING. Typically, ING is

not treated as a unit. However, ING has many of the properties of words thought to make them salient units. We investigated whether infants' abilities to segment nouns and verbs from speech changed if the ING unit were taken into account.

We present two types of data. First, we performed a corpus study of noun and verb statistics. This analysis indicates that when simple statistics are computed over words, nouns are more easily classified than verbs using the immediately adjacent units. However, the outcome changes if the computation of these statistics treats ING as a separate unit. ING occurs with high frequency in child directed speech and attaches to many verbs. This predicts that if the child is tracking ING in continuous speech, it should facilitate extracting the base words to which ING is often attached. In effect, by the same reasoning that led earlier researchers to conclude that transition probabilities would lead the infant to discover word boundaries, they should begin identifying units like ING, which facilitate learning verbs, making them more comparable to nouns with respect to age of acquisition. We then tested this prediction in a behavioral study with 7.5 and 9.5-month-old infants.

Corpus Study

In the corpus study we investigated noun-verb differences in distributional structure in child-directed speech with two primary purposes in mind: (1) in analysis 1, to replicate previous studies of the distributional differences between nouns and verbs; (2) in analysis 2, to see how these noun-verb differences change as a function of what gets counted as a unit when computing co-occurrence statistics.

A number of previous corpus studies have noted that nouns are easier than verbs to grammatically categorize based on their co-occurrence frames. For example, Redington, Chater, and Finch (1998) created context vectors for words in child-directed speech by computing the co-occurrence probability between a word and its adjacent neighbors in speech. They then compared the co-occurrence vectors of word pairs and used that similarity to try and predict whether or not the words were in the same grammatical category. For example, words like *truck*, *card*, and *hand* all tend to co-occur with the same set of words (like *the* and *my*), receive high similarity scores, and thus have a high likelihood of being considered in the "same" category compared to words like *crawl* and *eat*, which tend to co-occur with many other words. Redington et al. found that nouns were more easily classifiable than verbs based on their co-occurrence vectors (90% accurate vs. 72% accurate). Other corpus studies using different methods have replicated this noun-verb difference. For example, Mintz (2003) looked at the usefulness of frequent frames for grammatical categorization (such as "*the ___ is*"), finding that the 50 most frequent frames in child-directed speech are sufficient for good grammatical categorization of the intervening word, but that noun accuracy was much higher than verb accuracy.

The previous work looking at the effect of co-occurrence frames on the ability to grammatically categorize nouns and

verbs was done using traditionally defined words as co-occurrence units. In our study, we looked at whether treating the highly frequent unit ING as a unit affected the ability to grammatically categorize nouns and verbs, and whether it made the difference between nouns and verbs go away.

Method

Corpus. The corpus was derived from the CHILDES database of child-directed speech, using all samples of speech directed to children 24 months of age and younger. The samples were segmented into word units, and all punctuation and pauses were replaced with a PAUSE unit. In the first analysis, words were defined in the standard way, resulting in a corpus with 10,730 unique units, and a total corpus size of 1.207 million words. In the second analysis, all words that ended in the suffix ING were split such that the suffix was treated as a word-like unit, and the root words' spellings were normalized so that the root was counted as the same as occurrences of the word without the suffix. This resulted in a corpus with 10,551 unique units, and a total corpus size of 2.219 million words.

Stimuli. We used 150 words, the 75 most frequent nouns and verbs in child-directed speech according to the CHILDES corpus. The use of the most frequent words was done in order to facilitate making predictions about the following infant behavioral experiments, which would also use highly frequent words.

Statistics for Analysis 1. Our analyses were conducted in the same manner as Redington, Finch, and Chater (1998). First, we computed two, 10,730-element co-occurrence vectors for each of the 150 target words. The first vector was a record of the probability of each of the corpus's 10,730 words preceding the target word in the corpus; the second recorded the probability of those 10,730 words following the target word. These vectors were concatenated creating a 21,460-element vector for each of the 150 target words. Next we computed the Pearson correlation for all 150-by-150 target pairs, as a measure of the distributional similarity of each pair.

Similarity scores were used to predict the grammatical category of the words (as either nouns or verbs), again in the same manner as Redington et al. First, a similarity threshold was chosen as a cutoff, such that similarities between word pairs that were greater than the cutoff were predicted to be of the same category, and similarities that are below the cutoff were predicted to be of different categories. These distributional similarity-based guesses were compared to whether or not the actual pairs were both nouns, both verbs, or were of different categories. These comparisons are used to calculate signal detection measures: hits (guessed same category, was same category), misses (guessed different, was same), false alarms (guessed same, was different) and correct rejections (guessed different, was different). The signal detection measures are then used to calculate overall accuracy as $(\text{hits} / (\text{hits} + \text{false alarms}))$ and completeness as $(\text{hits} / (\text{hits} + \text{misses}))$.

Statistics for Analysis 2. The statistics for analysis 2 were the same as the previous analysis, except that the target words' vectors were only 10,551 elements long, reflecting the smaller number of words in that corpus due to the consolidation of the nouns' and verbs' different forms.

Results and Discussion

In analysis 1, the mean noun-noun similarity score was $r = 0.82$ (SE = 0.04), the mean verb-verb similarity score was $r = 0.73$ (SE = 0.07), and the mean noun-verb similarity score was $r = 0.38$ (SE = 0.18). Accuracy and completion scores for nouns and verbs were computed as described above, and plotted in Figure 1.

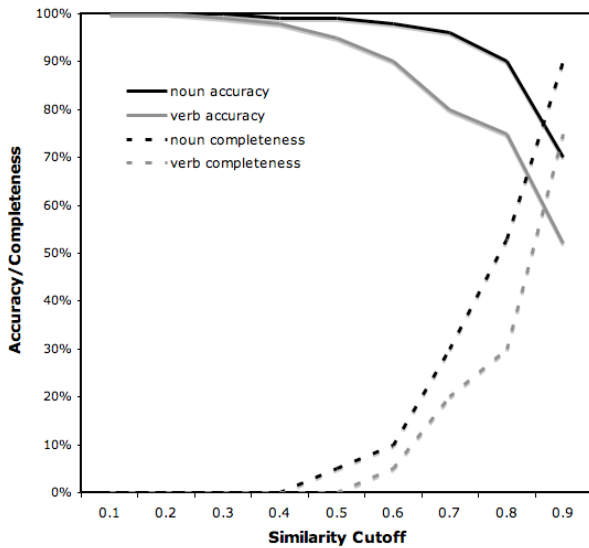


Figure 1: Grammatical categorization accuracy and completeness scores using 1-word distributional frames

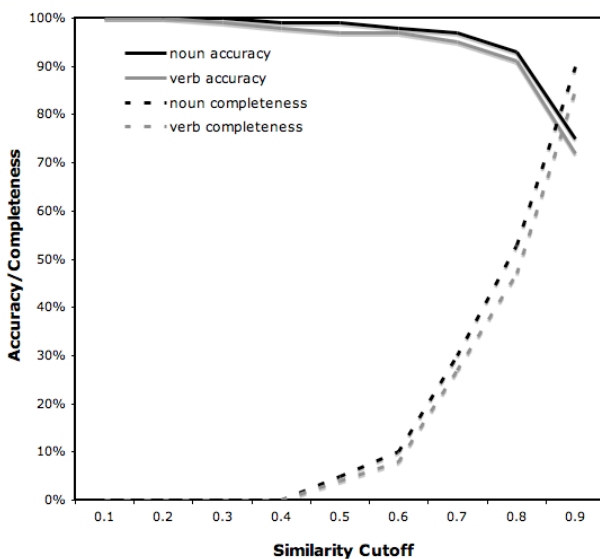


Figure 2: Grammatical categorization accuracy and completeness including ING

In analysis 2, the mean noun-noun similarity score was $r = 0.78$ (SE = 0.04), the mean verb-verb similarity score was $r = 0.75$ (SE = 0.05), and the mean noun-verb similarity score

was $r = 0.38$ (SE = 0.22). Accuracy and completion plots for nouns and for verbs are shown in Figure 2.

This corpus study replicates previous work by Redington et al. and Mintz, demonstrating a significant noun advantage when using co-occurrence frames in order to guess the grammatical category of a word. Except when at ceiling or at floor, noun categorization was both more accurate (more hits with fewer false alarms) and more complete (more hits with fewer misses) than verbs. In analysis 2, this difference goes away. There is no significant difference between noun accuracy and verb accuracy, or noun completeness and verb completeness, regardless of the cutoff that is used.

This shows that if ING is treated as a unit, it makes verbs more similar to each other. It also makes most nouns less confusable with verbs, since they very rarely occur with an ING. In general, noun and verb transition probabilities have different characteristics. When only words are considered, nouns have more consistent transitions. This makes them easier to classify and easier to segment. However, when verbs occur in their ING frame, this makes them very easy to classify, and this occurs often enough in child-directed speech (around 15% of a verb's occurrences) to make general noun-verb differences disappear, assuming ING gets treated as a separate unit.

The corpus study suggests that young infants have enough exposure to verbs in ING form to use it as a grammatical categorization cue. And because grammatical categorization is relying on information very similar in nature to word segmentation (namely, transitional probability statistics), this difference may also impact infants' ability to segment verbs as easily as they can segment nouns. We test this hypothesis in Experiments 1 and 2.

Experiment 1

Previous research has found that 7.5-month-old infants can segment nouns from fluent speech, but that infants cannot segment verbs until 13.5 months. As we have shown in the corpus study, this may be because the co-occurrence frames for verbs are typically much less frequent and much less consistent than for nouns.

In fact, in Nazzi et al.'s study showing that infants cannot recognize verbs in fluent speech until 13.5 months of age, the verbs and co-occurrence frames that were used were quite low in frequency (such as "boss *permits* everyone"). Nazzi et al. suggested that the reason infants have more trouble with verbs is due to phonological and prosodic factors. In English, most words are consonant initial and of strong-weak stress, but that verbs often are not. They manipulated these factors, and found this affected whether infants were successful and 13.5 months vs. 17.5 months.

It is possible that infants' difficulty in their study was due to low familiarity with the target verbs (*permit*, *discount*, *import*, and *incite*) compared to the high familiarity of Jusczyk and Aslin's target nouns (*dog*, *cup*, *bike*, and *feet*). It is also possible that the difference was due to Nazzi et al.'s relatively low frequency frames ("boss ___ everyone"). However, our corpus analysis suggests that even if high frequency verbs and verb frames are used, infants may still

fail because distributional differences between typical nouns and verbs are large, and that infants may be failing at these early ages because they have not yet learned about typical verb frames and are not yet good at using them to extract verbs from fluent speech.

Experiment 1 tested this hypothesis. If infants' failure to segment verbs in Nazzi et al was due mainly to their use of low frequency verbs and verb frames, then they should succeed in Experiment 1, in which we used high frequency verbs and verb frames. However, if they still fail, it could be attributed to infants' lack of knowledge about verb frames at these young ages.

Method

Participants. We tested 24 American infants (12 males, 12 females) from English-speaking families, between the ages of 9 and 10 months (with an average age of 39 weeks, 6 days; and with a range 37 weeks, 0 days to 42 weeks, 0 days). Eight additional infants were tested but not included due to crying during the procedure.

Procedure. Infants were tested individually in a 2-meter x 2-meter soundproof booth while seated in their caregiver's lap. The caregiver was given headphones playing music so that they could not hear the experimental stimuli and thus could not influence the infant's looking or listening to particular sounds. The experiment was run using the Headturn Preference Procedure.

The experimental procedure contained a familiarization phase and a test phase. During the familiarization phase, each infant heard 12 simple sentences, six using a single verb, and another six using a second verb. The presentation of the sentences was blocked by verb (all six sentences for a particular verb were played in a row), with a pause of one second between each sentence. The alternating blocks of six sentences were repeated four times. The total duration of the each block was approximately 23 seconds, and the total time during familiarization was approximately 140 seconds. In addition to hearing the sentences played during familiarization, infants were also exposed to flashing lights that were contingent on where they were looking. At the beginning of the familiarization phase, a center light in front of the infant flashed, and when the infant looked at it, a light on either the left wall or the right wall (chosen randomly) began to flash. Once the infant looked at the sidelight, it continued to flash until the infant looked away for more than two seconds, upon which the center light would begin to flash and the sequence of events would repeat. This process of light-flashing contingent on the infant's looking went on while the familiarization sounds played continuously in the background. During the familiarization phase the flashing lights and the sounds were not contingent on one another.

After the 140-second familiarization period, the test phase began. The test phase had 12 test trials. At the beginning of each test trial, the center light started flashing, and when the infant looked at it, a sidelight would begin to flash. When the infant looked at the sidelight, one of four verbs (the two verbs the infant heard during familiarization, and two "new"

verbs) began playing from a speaker mounted next to the flashing light. The verb played repeatedly with a 600 ms interval between each repetition until the infant looked away from the flashing light for more than 2 seconds, or until a maximum trial length (15 repetitions) was reached.

Materials and Design. The experiment used four verbs: *kiss*, *drink*, *give*, and *walk*. Each infant was exposed to only two of the four verbs during the familiarization phase, and then they heard all four during the test phase. The verbs were counterbalanced such that half the infants were exposed to *drink* and *give* during familiarization, and the other half were exposed to *kiss* and *walk*.

We chose the four targets by selecting verbs that were similar to the nouns from Jusczyk and Aslin (1995) in terms of infants' typical level of comprehension and exposure to the words. Measures of children's comprehension were obtained from the MacArthur-Bates Communicative Development Inventory norms (Dale & Fenson, 1996). Frequency in child-directed speech was measured by obtaining frequency counts from all corpora in the CHILDES database (MacWhinney, 2004) directed at infants 12 months and younger.

For each verb, we used the CHILDES database to find the six most frequent frames for each verb, using the same method as Mintz (2003) and the previous corpus analyses. An example of one of these frames is "to X it." We used the six frames to construct six sentences for each verb. We varied the location of the verb in each sentence such that it occurred once as the first word, once as the last word, and four times at varying points in the middle of the sentence. The verbs were the only words that were repeated more than once in any of a verb's six sentences, and the average sentence and syllable length were kept the same for each of the verbs' sentences. We recorded the target verbs and sentences, spoken in an infant-directed manner by a native-English-speaking adult. The individual verbs and the sentences as a whole were standardized in volume using Adobe Audition.

The design of the experiment was a thus 2 (verb familiarity: familiarized during training vs. not familiarized during training) x 2 (list 1 vs. list 2) mixed design, with the list variable between-subjects and the verb familiarity condition within-subjects. The dependent variable was the amount of time the infant spent looking at a flashing light while a verb was playing during the test phase.

Results and Discussion

A scatter plot of the looking times for Experiment 1 is shown in Figure 3. The figure shows each participant's mean novelty preference: their looking time for unexposed words minus the looking time for words they heard during familiarization. There was not a significant effect or interaction involving the list variable (all F 's < 1) so this variable was removed from further analyses.

In Experiment 1, 9.5-month-old infants showed no evidence of discriminating verbs they had been exposed to, compared to the verbs they had not ($F(1,23) = 0.60$, $p = 0.446$). The verbs were all very high frequency, in high

frequency frames, consonant initial, and of strong-weak stress, and infants were still not able to segment the verbs from fluent speech. This contrasts with other studies in which infants were successful with high frequency nouns in high frequency frames at 7.5 months.

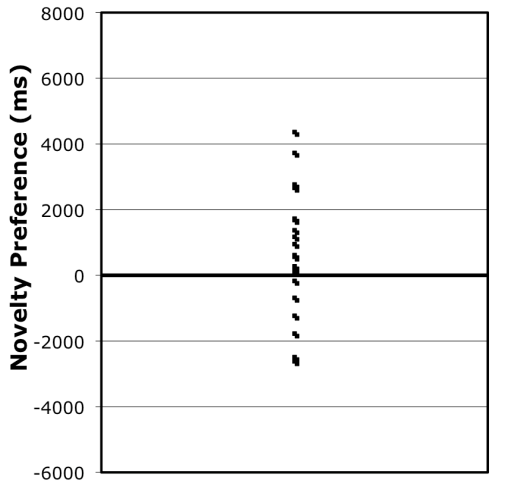


Figure 3. Infants' Novelty Preference in Experiment 1

There are several possible reasons for this result, such as semantic or pragmatic factors (e.g., infant-directed speech is more often using nouns in isolation and calling attention to specific objects at the same time). However, it is still possible that the difference is due to transition probability differences between nouns and verbs. As the corpus study showed, lexical verb frames are less diagnostic than lexical noun frames overall, and thus lesser knowledge about verb frames could be making it more difficult for infants to segment verbs. The corpus study also suggests that using inflectional forms like ING as units in computing transition probabilities would ameliorate this difference. Doing so would allow infants to make use of verbs' most frequent and most diagnostic frame. In Experiment 2 we tested this hypothesis.

Experiment 2

7.5 and 9.5-month-old infants were tested to determine if they could segment verbs if frame was the most informative one, which includes ING. If so, this would show that infants' difficulty with verbs (in our Experiment 1, and in Nazzi et al) was due to differences in transition probabilities between nouns and verbs, and that when this difference is equalized, the noun-verb difference is eliminated.

Method

Participants. We tested 48 American infants (24 males, 24 females) from English-speaking families. Half were between the ages of 7 and 8 months (with an average age of 30 weeks, 6 days; and with a range 37 weeks, 0 days to 35 weeks, 0 days). The other half were between the ages of 9 and 10 months (with an average age of 39 weeks, 6 days; and with a range 37 weeks, 0 days to 42 weeks, 0 days). Thirteen additional infants were tested but not included due to crying during the procedure.

Procedure. The procedure for Experiment 2 was exactly the same as for Experiment 1, with the exception of the training materials, as described below.

Materials and Design. The same four target verbs were used for Experiment 2 as for Experiment 1, and the test phase was exactly the same (infants heard the root form of the verb only). During the training phase, infants heard 12 sentences just as in Experiment 1, except that this time the verbs occurred in a frame that included the *-ing* inflection, such as "you X ing". The exact frames were once again chosen using the CHILDES corpus to identify the six most frequent frames for the verbs when they were used in their *-ing* form. Six grammatical sentences were constructed for each verb using these frames.

The design for Experiment 2 was a 2 (age: 7.5-months vs. 9.5-months) x 2 (verb familiarized during training vs. verb not familiarized during training) x 2 (list 1 vs. list 2) mixed design, with the age and list variables between-subjects and the verb familiarity condition within-subjects. The dependent variable was again the amount of time the infant spent looking at a flashing light while a verb was playing during the test phase.

Results and Discussion

A scatter plot of the listening times for Experiment 2 is shown in Figure 4. The figure again shows infants' mean novelty preference (e.g. their looking time for word to which they were not exposed, minus their looking time for the words to which they were exposed).

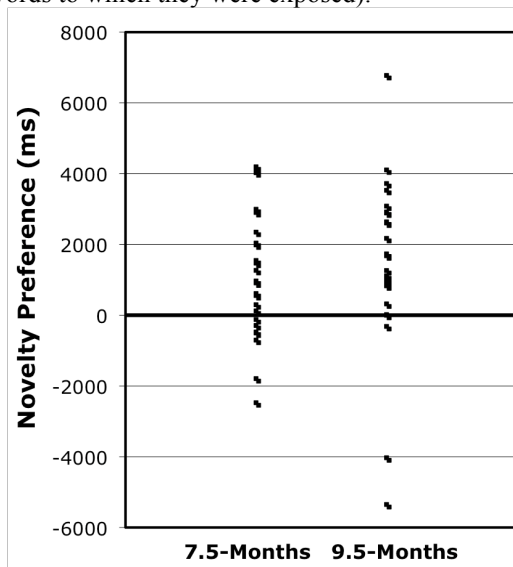


Figure 4. Infants' Novelty Preference in Experiment 2

There was not a significant effect or interaction involving the list variable (all F 's < 1) so this variable was removed from further analyses. There was also no significant effect of age ($F(1,47) = 0.85, p > 0.05$) nor age-by-familiarization interaction ($F(1,47) = 1.45, p = 0.234$). There was a main effect of familiarization, such that the verbs to which the infants had not been exposed were preferred during the test phase ($F(1,47) = 7.27, p = 0.009$). Follow-up tests show that

this difference was significant both for 7.5-month-olds ($p = 0.006$) and for 9.5-month-olds ($p = 0.009$).

Experiment 2 confirms the importance of units other than words in computing statistics in early language learning. Our corpus study suggested that treating morphological elements like ING as units in statistical computations has a large effect, eliminating the learnability difference between nouns and verbs. Experiment 2 demonstrated that this makes an important difference in infants' word segmentation. When verbs occur in the presence of ING, the verb can be segmented at the same age as nouns can be.

General Discussion

The corpus analyses and behavioral experiments reported in this article have important consequences for theories of language acquisition that invoke statistical learning. Most important is that the units that are used to calculate the statistics have a major impact on what is learnable. When high frequency morphemes such as ING do not get treated as units in statistical calculations, nouns have a significant advantage in terms of segmentation from fluent speech, and are likely much easier to classify in terms of their grammatical category and semantics. When high frequency morphemes such as ING are counted as units, verbs are just as easy to segment, and grammatically classify. Future work will need to address the question of how units are determined for different tasks at different ages. ING appears to function as a unit because it exhibits statistical properties similar to words.

Our research suggests that verb forms are not necessarily segmented and learned later than nouns; some verbs can be learned early if certain conditions are met. The infant must be treating morphological frames such as ING as units; the verb must occur frequently in this frame; and the verb itself must occur with high frequency (in work not reported here, infants could not use the ING frame to segment nonwords). Other studies such as Nazzi et al. deviated from these conditions and thus found learning for verbs to only occur at much later ages. The factors they studied (effects of prosody and phonotactic regularities) are relevant to how children learn lower frequency words (consistent with a corpus analysis of noun-verb differences by Monaghan, Chater, & Christiansen, 2005). These kinds of information are clearly very important for language learning and comprehension (Johnson & Jusczyk, 2001), but may not come online until infants are older (Saffran & Thiessen, 2003).

The present work also contributes to our understanding of the onset of morphological learning. Principles that were established in connection with the word learning/segmentation problem extend to the identification of subword units such as ING. These units will eventually be morphemes: units that have consistent phonological and semantic content, which combine and recombine with other units in systematic ways and are the basis of productivity in the lexicon.

Our results also bear on the debate about whether or infants use statistics or more formal, structured types of knowledge during language acquisition. In an early reply to

Chomsky's criticism of behaviorist approaches to language learning, Osgood (1963) pointed out that many of Chomsky's statements about the limitations of statistical approaches apply only if the level at which the statistics are being computed is the word. Acknowledging that statistics are computed at many levels, including subword, word, and supraword, is likely to expand the range of language learning phenomena that statistical learning can explain.

Acknowledgments

This research was supported by NIH grant R01HD47425. J. Willits was supported by NIH Institutional Training grant 5T32HD049899 to the University of Wisconsin-Madison.

References

- Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, 28, 125-127.
- Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek, & R. Golinkoff, (Eds.) *Action meets word: How children learn verbs* (pp.544-564). Oxford University Press.
- Johnson, E.K. & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44, 1-20.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29, 1-23.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk. Third Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, 90, 91-117.
- Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorization. *Cognition*, 96, 143-182.
- Nazzi, T., Dille, L. C., Jusczyk, A. M., Shattuck-Hufnagel, S., & Jusczyk, P. W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech*, 48, 279-298.
- Osgood, C. (1963). On understanding and creating sentences. *American Psychologist*, 18, 735-751.
- Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, 22, 425-469.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.
- Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, 39, 484-494.
- Saffran, J. R., & Sahni, S.D. (2007). Learning the sounds of language. In M. Joanisse, M. Spivey, and K. McRae (Eds.), *Cambridge Handbook of Psycholinguistics*, Cambridge University Press.