Relationships between Language Structure and Language Learning: The Suffixing Preference and Grammatical Categorization

Michelle C. St. Clair

School of Psychological Sciences,

University of Manchester, Manchester, UK

Phone: +44 (0)161 2753530

Fax: +44 (0)161 2743965

michelle.stclair@manchester.ac.uk

Padraic Monaghan

Department of Psychology,

Lancaster University, Lancaster, UK

Phone: +44 (0)1524 593813

Fax: +44 (0)1524 593744

p.monaghan@lancaster.ac.uk

Michael Ramscar

Department of Psychology,

Stanford University, Stanford, CA, USA

Phone: +01 650 725 2417

Fax: +01 650 725 5699

ramscar@stanford.edu

Key words: language learning, language evolution, suffixing preference, grammatical categorization, corpus analysis, human experimentation

Abstract

It is a reasonable assumption that universal properties of natural languages are not accidental. They occur either because they are underwritten by genetic code, because they assist in language processing or language learning, or due to some combination of the two. In this paper we investigate one such language universal – the suffixing preference across the world's languages, whereby inflections tend to be added to the end of words. A corpus analysis of child-directed speech in English found that suffixes were more accurate at cuing the grammatical category of the root word than were prefixes. An artificial language experiment found that there was a learning advantage for suffixes over prefixes in terms of grammatical categorization within an artificial language. The results are consistent with an account of language universals that originate in general purpose learning mechanisms.

1. Introduction

There are many potential contributors to the occurrence of universal properties of natural languages. Language universals can be accounted for in terms of a common genetic encoding that predetermines language structure (e.g., Chomsky, 1957; Pinker, 1984). An alternative perspective is that they become incorporated within every language because they are propitious for the learning or the processing of the language as determined by general purpose learning mechanisms (Christiansen & Chater, 2008), though these two views are not necessarily mutually exclusive and some combination of genetic and learning constraints may converge to shape language. A common approach to arguing for genetic predisposition is to identify some aspect of language that is unlearnable, and then concluding that it must be innate. However, demonstrating unlearnability of language structure is a difficult, and perhaps impossible, task.

An alternative approach to questions of learnability, in the sense of differences in ease of learning, is to identify aspects of language that appear to be shaped by general cognitive constraints. Recent computational simulations, for instance, have shown that it is possible to learn far more about language structure from limited input than had been previously supposed (Reali & Christiansen, 2005). In this paper, we examine one well-established property of natural languages – the suffixing preference – as a candidate for assisting language learning based on general learning principles.

There is a universal tendency across natural languages for morphemes that modify either the grammatical or semantic properties of words to attach to the end (suffixing) rather than to the beginning (prefixing) or middle (infixing) of a word (Sapir, 1921). Hawkins and Gilligan (1988) noted that whereas suffixing only languages are common (74 of 203 languages studied), prefixing only languages are rare (9 of 203), although the most common pattern is a combination of prefixing and suffixing. This last class of languages still tends to have a predominance of suffixes: English, for instance, has 56 prefixes and 181 suffixes in Fudge's (1984) comprehensive list of inflectional and derivational morphemes.

Hawkins and Gilligan (1988) suggested that the suffixing preference is due to the preference across languages to have the part of phrases or words that determines the category at the end of the phrase or word (Hoeksema, 1985). As this head-final structure is more common than head-initial across languages (Hawkins & Gilligan, 1988), the consequence should be more suffixing than prefixing across the world's languages. Hawkins and Gilligan (1988) found that languages with object-verb word order and postpositions following noun phrases (NP+Po), characteristics of syntactic head-final languages, have more suffixing and less prefixing, as predicted. Languages with verb-object word order and prepositions preceding noun phrases indicating head-initial phrasal structure generally had both prefixing and suffixing. Languages which have only prefixing were invariably found to have verb-object word order, and only very rarely displayed the NP+Po pattern.

A processing account of the suffixing preference was proposed by Cutler, Hawkins and Gilligan (1985; also Hawkins & Cutler, 1988), assuming that identifying each word in continuous speech as soon as possible provides a processing benefit, and hence a communicative advantage. Prefixes provide little information about the unique identity of the word, and on average will shift the uniqueness point (Marslen-Wilson, 1987) later in the word, as the first few phonemes are shared by every word containing the affix. In contrast, suffixing does not delay the identification of the word root, and the set of candidate words at each phoneme in the word is on average much smaller than if the word was prefixed.

An alternative view of the suffixing preference is that it facilitates language *learning* as well as language processing, in particular in terms of marking the grammatical category of the root word. Hawkins and Gilligan (1988) found five cross-linguistic examples of noun affixes that always occurred as suffixes, including case-marking and gender affixes (Kelly, 1996). An additional 12 cross-linguistic instances occurred significantly more frequently as suffixes than prefixes (Cutler et al., 1985; though see Enrique-Arias, 2002 for some affix types that seem to occur equally often as prefixes). These consistencies provide evidence that affixes can be reliable indicators of grammatical class, which is a necessary precursor to learning language structure. Additional support for the learning advantage of suffixes is the finding that children learn suffixes more readily than prefixes (Clark, 2007; Slobin, 1973), and that inflectional systems in suffix-only languages are learned earlier than in prefix-only languages (Kuczaj, 1979; Mithum, 1989; Slobin, 1973), though isolating suffixing as the key factor in ease of learning was not accomplished in these studies.

Though affixes are structurally related to grammatical properties of the language, it has not yet been established empirically that affixes provide reliable information about the grammatical categories of root words. Syncretism (the merging of distinct morphological forms into a single form) is well-attested and this could result in the same morpheme marking multiple grammatical categories, meaning the affix may not provide useful information for categorization of the root word. However, there are persuasive reasons for considering suffixes in particular as a candidate for promoting learning of grammatical categories. Greenberg (1957) suggested that suffixing and prefixing are equated to “convergent” and “divergent” hierarchies, respectively (Osgood, 1949). In terms of stimulus-response (S-R) associative learning, convergent hierarchies describe varied stimuli coupled with a functionally identical response (S_1, S_2, ...$S_x \rightarrow$ R), which is analogous to suffixing. Conversely, divergent hierarchies indicate similar stimuli but varied responses (S1 $\rightarrow$ R_1, R_2, ...R_x), similar to prefixing. Greater facilitation and positive transfer in learning the relationship between the stimuli and the response occurs for convergent hierarchies, whereas there is negative transfer and interference in learning for divergent hierarchies (Osgood, 1949). The different learning dependent on divergent and convergent hierarchies can be formally described in terms of error driven models of learning (e.g., Rescorla & Wagner, 1972), and can be implemented in a range of iterative learning mechanisms, such as recurrent neural networks (Ramscar & Yarlett, 2007; Ramscar, Yarlett, Dye, Denny, & Thorpe, submitted).

Error-driven learning frameworks predict, then, that informative associations between root words and suffixes will be more readily learned than those between prefixes and root words. Therefore, if natural languages are adapted for learning grammatical categories in terms of suffixing, then suffixes in natural languages ought to provide more reliable information about the category of the root word than prefixes, which provide information less easily available for learning. We provide the first empirical test of this in a corpus analysis of child-directed speech to determine whether suffixes provide more reliable information about grammatical categories than prefixes in English child-directed speech.

2. Corpus analysis: Categorization from affixes in child-directed speech

2.1 Method

2.1.1 Corpus preparation

We collected all speech spoken by adults in the presence of children from the English language section of the CHILDES database. The corpus contained 5.7 million words in 1.3 million utterances. Words were assigned a grammatical category using MOR, which has tagging accuracy of approximately 95% (Sague, MacWhinney, & Lavie, 2004). The grammatical category tags were grouped into 11 categories: noun, verb, adjective, adverb, numeral, pronoun, article, preposition, conjunction, interjection, and contraction (e.g., *it's, where're*).

The corpus was then searched for words that began with each prefix or ended with each suffix from Fudge's (1984) list in terms of their orthography. Affixes that had alternate spellings (e.g., -itious/-itous) were included as separate affixes. For each affix, the number of words of each category that contained that affix was determined in both a type analysis (where each distinct word counted once for each affix) and token analysis (where each word contributed to the proportion according to its frequency). In order to determine whether each affix was a reliable indicator of the grammatical category of the root word, we counted the highest grammatical category proportion for each affix. If the affix was a reliable indicator then this proportion will be high. However, if the affix was a weak indicator and occurs in several word categories, then the proportion will be low. Words that were ambiguous with respect to category contributed to the proportions according to their frequency of usage as each category in the corpus. Table 1 shows an example analysis for two affixes: the suffix "-ness" and the prefix "im-".

------------------------------Insert Table 1 about here -------------------------------

In order to control for potential biases in the different orthographic indication of prefixes and suffixes that may not be expressed phonologically (in the form available to the child learning the language) we repeated the analyses except that we extracted the phonological form for each word in the corpus that also occurred in the CELEX database (Baayen, Pipenbrock, & Gulikers, 1995). We also determined the most frequent phonological form for each of the affixes in the Fudge (1984) list, and searched for each phonological prefix and suffix in a similar way to the orthographic analyses. Affixes with identical phonological expression were not repeated in the analyses (e.g., "-ious" and "-ous" were entered once as /əs/). There were 52 phonological prefixes and 151 phonological suffixes.

An additional bias in the corpus analyses may result from the potential of pseudo-affixes (such as "-ist" in "moist") to appear more as prefixes or as suffixes. We performed an additional analysis to minimize this potential bias by selecting only those words from the corpus classified in the CELEX database as polymorphemic, or derived from a lemma form with a different expression to the lemma (so, “divide” was not included, but “divided” was included, “be” was not, but “is” was included). We repeated both the orthographic and the phonological form analyses for the polymorphemic corpora.

2.2 Results and discussion

The difference in the maximal proportion of each grammatical category for prefixes and suffixes was assessed using the Mann-Whitney U-test. Affixes that did not appear in the corpus were omitted from the analyses. Table 2 reports the results of the analysis for orthographically and phonologically expressed affixes for the corpus containing all words and the corpus containing only words classified as polymorphemic. For all the analyses, the prefixes were significantly less accurate than suffixes at defining grammatical categories. Measuring the affixes orthographically or phonologically made little difference in the relative accuracies of the categorizations. Furthermore, the polymorphemic words analyses seemed to indicate that pseudo-affixes contributed to the differences in proportion of categorizations by slightly reducing the accuracy of the suffix categorizations. So, if the language learner can detect whether an affix is a morphological component then accuracy of categorizations based on suffixes can further improve accuracy.

-----------------------------Insert Table 2 about here ------------------------------

To test whether the results were dependent on all the inflectional morphemes being suffixes, whereas derivational morphemes may be prefixes or suffixes, we repeated all the analyses omitting the inflections “–s”, “-ed”, “-t”, “-er”, “-est”, and “–ing”. The results were very similar and are shown in Table 3.

------------------------------Insert Table 3 about here -------------------------------

The corpus analyses indicated that, as predicted, there was more reliable grammatical category information present in suffixes than prefixes in English. Models of associative learning predict that learning the association between prefixes and root words would be more difficult than learning the root-suffix association. In English, this suffix advantage is a feature of the language in terms of reflecting grammatical categories. However, though this information for categorization may be present in natural language, it has not yet been shown directly that suffixing presents a learning advantage over prefixing. Note that categorization is a computational process that goes beyond merely learning associations, as the similarity between different words that occur with the same affix has to be identified in addition to learning the affix-root word pairing. The next study directly tested the learning of grammatical categories from suffixes compared to prefixes.

3. Experiment: Testing the suffix advantage for categorization

In order to test the differential quality of prefixes and suffixes in aiding grammatical categorization an artificial language learning experiment was conducted. This paradigm has been used in numerous studies to investigate grammatical categorization (i.e., Braine et al., 1990; Monaghan, Chater, & Christiansen, 2005). Of particular relevance, Valian and Coulson (1988) found that affix-like cues that preceded words resulted in the assignment of words into categories according to co-occurrence with the affix. These previous studies have investigated the effectiveness of distributional cues for categorization, though none has yet compared the relative effectiveness of preceding and succeeding cues. This experiment directly compared learning categories from either prefixes or suffixes.

3.1 Method

3.1.1 Subjects

Twenty-four University of York undergraduates (age 18-22; 19 female) participated in this experiment. All were native English speakers and were paid £3 or given course credit.

3.1.2 Materials

The artificial language comprised 12 category words divided between two categories, *A* and *B*. There were two affixes in the language (*a* and *b*), one always co-occurring with category *A* words and the other co-occurring with category *B* words. We refer to these as affixes as they are phonologically similar to affixes in natural language (see below) and because, as in natural language, they provide reliable grammatical category information about the root word. Category *A* words contained onset and offset consonant clusters, unrounded high vowels, nasals and stops. Category *B* words had no consonant clusters, rounded low vowels, and fricatives. These phonological regularities have been found to be necessary to induce categorical learning in this and similar paradigms (Brooks, Braine, Catalano, Brody, & Sudhalter, 1993; Monaghan et al., 2005) and so we included them in the materials to maximize the possibility of observing category learning. Note that the presence of the phonological information was identical for the prefix and suffix conditions. Table 4 lists the category words for both categories. Half of the category words were high frequency and occurred twice as often as the low frequency category words. For each participant, two affixes were selected from four possibilities (/gæ/, /mI/, /vɛ/, and /dʌ/) to ensure that the results were not unduly influenced by any particular affix. The four possible affix words were chosen to be similar in terms of voicing and vowel quality to English affixes (St. Clair, 2007). However, the ending vowels in the affixes were lax vowels, consistent with word-internal vowels in English (Norris, McQueen, Cutler, Butterfield, & Kearns, 2001), so learning with these affixes favored the prefixing condition, rather than the suffixing condition. Advantage in the suffixing condition, therefore, can to be taken as robust evidence of the strength of the suffixing advantage in language learning.

-----------------------------Insert Table 4 about here ------------------------------

There were 18 sentences for training, each containing two category word-plus-affix pairs, with the sentences produced in monotone using the Festival speech synthesizer with a British English voice (Black, Clark, Richmond, King, & Zen, 2004). In the prefix condition, the affix preceded the category words (*aAbB*) whereas the affix succeeded the category words in the suffix condition (*AaBb*). *A* phrases (*aA* or *Aa*) appeared equally often in the first and second position.

Participants were tested on 24 sentences. Twelve were compatible sentences which had not occurred during training but conformed to the artificial language. The remaining 12 sentences were incompatible, containing affixes that co-occurred with the wrong category word (e.g., *aBbA*, or *AbBa*).

3.1.3 Procedure

Participants were instructed to pay attention to the patterns within the language. To familiarize participants to the language, they first heard all the root words in the language without affixes, as well as the affixes. Participants then heard the 18 training sentences presented in random order, repeated in four blocks. Each trial consisted of two presentations of the training sentence with a one second interval between presentations. Participants were instructed to repeat the sentence aloud. The next trial was presented after a five second interval.

During the test phase, participants were instructed that half the sentences were similar to the training language and half dissimilar, and responded with the “Y” keyboard key for similar, and “N” for not similar. The participants were correct if they accepted the compatible sentences and rejected the incompatible sentences. The training and testing sessions were then repeated, using another set of compatible and incompatible sentences for the second test. The test set was counterbalanced.

In order to test for categorization over and above associative learning of the sequences, after the second test session the participants were given 12 cards each with one category word without the accompanying affix printed on it. Participants were then asked to sort them into two equal groups according to which words they thought went together. Accuracy was determined by the number of cards of the same category sorted together, in the range [3,6]. Chance level for this task was calculated from the proportion of choosing all six words of one category by chance, which is 3.91 cards (see Monaghan et al., 2005, for further details of the chance calculation).

3.2 Results and discussion

We performed an ANOVA with test session and frequency as within subjects factors and affixing condition as a between subjects factor. A dummy variable (test counterbalance) was included in the ANOVA as a between subjects factor (Pollatsek & Well, 1995). The dependent variable was the proportion of correctly accepted or correctly rejected test sentences.

As predicted, there was a significant main effect of affixing condition, with the suffix condition performing more accurately than the prefix condition, $M = .80$, $SD = .15$ and $M = .67$, $SD = .12$, respectively, $F(1, 20) = 4.92$, $p < .05$, $\eta_p^2 = .20$. There were no other significant main effects or interactions: test session by condition: $F(1,20) = 2.04$, $p = .17$; test session by condition by frequency: $F(1,20) = 1.28$, $p = .27$; all others $F < 1$. Both the prefix and suffix condition were significantly above chance level, $t(11) = 4.96$, $p < .001$, $d = 2.99$ and $t(11) = 6.92$, $p < .001$, $d = 4.17$, respectively.

For the card sorting results, there was no significant difference between the prefix and suffix conditions, $t(22) = -.48$, $p = .63$. However, the prefix condition did not differ from the chance level of 3.91, $t(11) = 1.56$, $p = .30$, $M = 4.5$, $SD = 1.31$, whereas the suffix condition was significantly above chance level, $t(11) = 2.39$, $p < .05$, $d = 1.44$, $M = 4.75$, $SD = 1.22$, indicating that category learning in the suffix condition was more robust.

The sentence judgment and the categorization task results taken together support the Rescorla-Wagner model prediction that suffix cues should be better predictors than prefix cues for learning the relationship between the marker words and the category words. Although both groups learned above chance level, performance for the suffix language was substantially more accurate than the prefix language. This difference was found even though for both the affixing conditions some suffix information was present. In the prefixing condition, for example, category *B* words were half the time succeeded by the *a* marker word, and half the time succeeded by no marker word. Yet, the difference in reliability of the suffixing information between the prefixing and suffixing condition was sufficient to elicit a difference in learning, both in terms of recognizing consistent and inconsistent sentences, and in forming categories of words in the card-sorting task. These results support the corpus analysis indicating that suffixes provide more consistent cues to the grammatical category of the root word.

4. General Discussion

We have examined the hypothesis that suffixes provide a learning advantage for learning of the grammatical categories of root words. The advantage of succeeding cues for providing associations has a long and prestigious history in the associative learning literature (e.g., Rescorla & Wagner, 1972), and has previously been viewed as a contributory factor to the suffixing preference found across the world's languages (Greenberg, 1957). This view is perfectly consistent with claims that the suffixing preference results in more efficient language processing, such as the importance of early word identification (Hawkins & Cutler, 1988). However, in this paper, we have considered suffixing to be an advantage for the *learning* and not only the *processing* of the language, as supported by the corpus analysis and the language learning study.

However, the statistical mechanism required to learn the category structure of the artificial language, and, by extension, the grammatical categories of natural languages, must be more sophisticated than just learning the associations between suffixes and particular root words. The language learner needs to recognize that the root words have some shared property as a consequence of their distribution with respect to the affixes in the language. Statistical clustering of words according to their context of occurrence with other words occurring immediately before or after them has proven to be a valuable mechanism for clustering grammatical categories together (Redington, Chater, & Finch, 1998). The corpus analyses indicated that suffixes provided the most information about grammatical category, and we contend that a similar clustering mechanism that focuses on suffixing co-occurrence would provide a powerful mechanism for learning accurate categories.

We have seen that succeeding morphological cues are better for learning than preceding cues, yet there are many preceding lexical distributional cues that also potentially provide information about grammatical categories, such as "the" preceding nouns in English. However, these highly frequent cues can also be viewed as reliable succeeding cues as well. For instance, "the" serves both as a preceding cue for nouns and as a succeeding cue for verbs, and in a computational model of grammatical category learning succeeding word cues were found to be more accurate (St. Clair, Monaghan, & Christiansen, submitted).

The corpus analysis indicates that language may indeed be shaped through language evolution to exploit the use of suffixes for categorization during language learning. The Rescorla-Wagner model provides a domain general mechanism for why suffixing cues are more informative than prefixing cues and predicts the dominance of suffixes throughout the world's languages and the rarity of prefix only and infix languages (Sapir, 1921).

Yet, language learnability is just one of several forces influencing language structure, and the presence of prefixes in some languages is testimony to the variety of pressures that influence morphological change. For instance, diachronic language studies indicate that affixes retain the same position as the ancestral unbound form with respect to the root word (Givón, 1979), consequently, languages with head-initial and/or verb-initial phrase structure are likely to contain some prefixes (Bybee, Pagliuca, & Perkins, 1990). Additionally, the likelihood of prefixation increases with morphological complexity. This is because prefixation is more likely than additional suffixation in words that already contain suffixes, which results in a combination of prefixing and suffixing in languages with a high affix to root ratio (Enrique-Arias, 2002). It is evidently not the case that all languages are equally learnable, and structural properties of languages produce variance in this respect (Christiansen & Chater, 2008; Evans & Levinson, in press), however, our results do suggest that in learning grammatical category structure, suffixing is an advantageous property.

Another potential contributor to the advantage for suffixing is the possibility that listeners were perceiving trochaic stress patterns in our stimuli, which is the most frequent stress pattern in English (Cutler & Carter, 1987), even though the speech stimuli were produced in a monotone. Such a bias in perception could have mitigated against participants processing the prefixed words as a whole. Morphology has an effect on stress position –

several affixes alter the stress pattern of the root word (Fudge, 1984) – and difficulties in segmentation of words with a de-stressed prefix may be an additional contributor to the suffixing preference. Such a view is consistent with our own examination of the suffixing advantage in corpus analyses (where no perceptual biases were taken into account) and our experimental study, and, as with the psycholinguistic explanations of the processing advantage, provide converging accounts of the suffixing preference.

We have characterised the suffixing preference as a language universal that is driven by constraints in a general-purpose cognitive system, yet there remains the possibility that the suffixing advantage in the experimental study may be a transfer effect from the dominant mode of affixation in the participants' native language, English. Ideally, the experiment would have to be replicated with monolingual speakers from a prefixing language, but this raises practical difficulties due to the rarity and often geographical isolation of speakers of prefixing languages, and additional problems associated with bilingualism in such populations (e.g., Kmer or Congolese and French; Navajo and English). However, there is converging evidence supporting our account of the suffixing advantage as an advantageous feature for language learning. The results from the natural experiment of the corpus analysis show that the suffixing advantage applies for learning English grammatical categories. As mentioned previously, in first language learning studies, suffixes are learned earlier than prefixes (Clark, 2007; Slobin, 1973), and inflectional systems are learned quicker in suffix-only languages than prefix-only languages (Kuczaj, 1979; Mithum, 1989; Slobin, 1973). In addition, Hupp, Sloutsky and Culicover (in press) investigated the suffixing preference using multimodal stimuli, and found a suffixing preference for linguistic, tonal, and visual shape stimuli. Transfer effects from a first language are unlikely to extend to learning of tonal or shape sequences (see also Ramscar et al., submitted, for convergent evidence). Thus, we contend that there is growing evidence that the suffixing advantage for learning is not just a transfer effect but is a reflection of general learning principles that also apply to language acquisition.

This view of the suffixing preference as a key property for language learnability is one instance of the general view advanced by Christiansen and Chater (2008) of languages adapting to fit the cognitive constraints of the brain, in the absence of a stable language environment to which evolutionary changes in the brain can adapt. We have here identified one such property of learnability for constraining the language – differences in the information that can be learned associatively from preceding compared to succeeding information. Learning the grammatical categories of a language is an important precursor to learning the full syntactic complexity of the language, and we have indicated, through a combination of corpus analysis and artificial language learning, the characteristics of the learning mechanism that drive such learning.

Acknowledgements

We thank Professors Gareth Gaskell and Graham Hitch for their helpful comments on this research. The artificial language learning experiment was submitted as part of the first authors PhD thesis at the University of York.

References

Baayen, R. H., Pipenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Black, A. W., Clark, R., Richmond, K., King, S., & Zen, H. (2004). Festival Speech Synthesizer (Version 1.95). Edinburgh: University of Edinburgh.

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., Catalano, L., Brody, R. E., & Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language, 32*, 76-95.

Bybee, J. L., Pagliuca, W., & Perkins, R. D. (1990). On the asymmetries of the affixation of grammatical material. In W. Croft, K. Denning & S. Kemmer (Eds.), *Studies in Typology and Diachrony*. Amsterdam: John Benjamins Publishing Company.

Chomsky, N. (1957). *Syntactic structures*. New York: Mouton.

Christiansen, M., & Chater, N. (2008). Language as shaped by the brain. *Behavioural and Brain Sciences, 31*, 489-509.

Clark, E. V. (2007). Morphology in language acquisition. In A. Spencer & A. M. Zwicky (Eds.), *The Handbook of Morphology* (pp. 374-389). Oxford: Blackwell Publishing.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2*, 133-142.

Cutler, A., Hawkins, J. A., & Gilligan, G. (1985). The Suffixing Preference: A Processing Explanation. *Linguistics, 23*, 723-758.

Enrique-Arias, A. (2002). Accounting for the position of verbal agreement morphology with psycholinguistic and diachronic explanatory factors. *Studies in Language, 26*, 1-31.

Evans, N., & Levinson, S. (in press). The myth of language universals: Language diversity and its importance for cognitive science. *Brain and Behavioural Sciences*.

Fudge, E. (1984). *English Word-Stress*. London: George Allen & Unwin.

Givón, T. (1979). *On Understanding Grammar*. New York: Academic Press.

Greenberg, J. H. (1957). Order of affixing: A study in general linguistics. In *Essays in linguistics*. New York: Wenner-Gren Foundation for Anthropological Research.

Hawkins, J. A., & Cutler, A. (1988). Psycholinguistic factors in morphological asymmtery. In J. A. Hawkins (Ed.), *Explaining Language Universals*. Oxford: Basil Blackwell.

Hawkins, J. A., & Gilligan, G. (1988). Prefixing and suffixing univerals in relation to basic word order. *Lingua, 74*, 219-259.

Hoeksema, J. (1985). *Categorical Morphology*. New York: Garland Publishing Inc.

Hupp, J. M., Sloutsky, V. M., & Culicover, P. W. (in press). Evidence for a domain-general mechanism underlying the suffixation preference in language. *Language and Cognitive Processes*.

Kelly, M. (1996). The Role of Phonology in Grammatical Category Assignments. In J. L. Morgan & K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 249-261). Mahwah, NJ: Lawrence Erlbaum Associates

Kuczaj, S. A. (1979). Evidence for a language learning strategy: On the relative ease of acquisition of prefixes and suffixes. *Child Development, 50*, 1-13.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word recognition. *Cognition, 25*, 71-102.

Mithum, M. (1989). The acquisition of polysynthesis. *Journal of Child Language, 16*, 285-312.

Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition, 96*, 143-182.

Norris, D., McQueen, J. M., Cutler, A., Butterfield, S., & Kearns, R. (2001). Language-universal constraints on speech segmentation. *Language and Cognitive Processes, 16*, 637-660.

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review, 56*, 132-143.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology, 21*, 785-794.

Ramscar, M., & Yarlett, D. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science, 31*, 927-960.

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (submitted). The feature label-order effect and its implication for symbolic learning.

Reali, F., & Christiansen, M. H. (2005). Uncovering the richness of the stimulus: Structure dependence and indirect statistical evidence. *Cognition, 96*, 143-182.

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science, 22*, 425-469.

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 Theory and Research* (pp. 64-99). New York: Appleton-Century-Crofts.

Sague, K., MacWhinney, B., & Lavie, A. (2004). *Adding syntactic annotations to transcripts of parent-child dialogs*. Lisbon, Portugal.

Sapir, E. (1921). *Language*. New York: Harcourt, Brace and Company.

Slobin, D. I. (1973). Cognitive prerequisites for the acquisition of grammar. In C. A. Ferguson & D. I. Slobin (Eds.), *Studies of child language development* (pp. 175-208). New York: Holt, Rinehard & Winston.

St. Clair, M. C. (2007). *Language structure and language acquisition: Grammatical categorization using phonological and distributional information.* Unpublished Thesis, University of York, York, UK.

St. Clair, M. C., Monaghan, P., & Christiansen, M. H. (submitted). Learning grammatical categories from distributional cues: What information is useful and when? *Cognition*.

Valian, V., & Coulson, S. (1988). Anchor Points in Language Learning: The Role of Marker Frequency. *Journal of Memory and Language, 27*, 71-86.

Table 1. Example corpus analysis of two affixes for determining the classification of words of the same category in the orthographic type and token analyses.

Affix	Words (Category: Count)[a]	Proportion Same Category[b] Types	Tokens
-ness	babiness (N: 1), badness (N: 1), blackness (N: 1), brightness (N: 1), business (N: 98), crookedness (N: 1), dizziness (N: 1), exclusiveness (N: 1), goodness (N: 1133), harness (V: 13), happiness (N: 11), laziness (N: 2), likeness (N: 2), madness (N: 6), meanness (N: 1), ness (N: 2), nosiness (N: 1), numbness (N: 2), seriousness (N: 1), sickness (N: 6), sweetness (N: 3), wetness (N: 1)	**.95 (N)** .05 (V)	**.99 (N)** .01 (V)
im-	Im (N: 2), Ima (N: 1), image (N: 5), images (N: 2), imaginary (Adj: 5), imagination (N: 29), imaginations (N: 1), imagine (V: 111), imagines (V: 1), imagining (V: 3), imitate (V: 9), imitating (V: 6), imitation (N: 6), immaculate (Adj: 2), immediately (Adv: 13), immigrant (N: 1), immortalize (V: 1), imp (N: 11), impede (V: 1), impetigo (N: 1), implication (N: 1), import (V: 1), importance (N: 1), important (Adj: 123), impossible (Adj: 24), impressed (V: 26), impression (N: 14), impressive (Adj: 11), improvement (N: 4), improves (V: 1), improvise (V: 4), improvising (V: 1), impulse (N: 2)	**.45 (N)** .36 (V) .15 (Adj) .03 (Adv)	.19 (N) **.39 (V)** **.39 (Adj)** .03 (Adv)

N = noun, V = verb, Adj = adjective, Adv = adverb.

[a]For Words column, labels and values in parenthesis indicate category and count for each type in the corpus containing the affix.

[b]For Proportion Same Category column, label in parenthesis indicates the category with the proportion of types or tokens, respectively containing the affix. Entries in bold indicate the highest proportion.

Table 2. Corpus analyses of proportion of words of the same category classified by prefixes and suffixes for orthographic and phonological analyses of all words and only polymorphemic words.

Analysis	Corpus	Types			Tokens		
		Prefix	Suffix	U	Prefix	Suffix	U
Orthographic	All Words	.499	.772	953.5***	.675	.794	2666.0***
	Polymorphemic	.494	.803	854.0***	.625	.852	1466.0***
Phonological	All Words	.499	.739	1031.5***	.702	.784	2262.0***
	Polymorphemic	.517	.788	809.5***	.666	.841	1412.5***

*** indicates $p < .001$.

Table 3. Corpus analyses of proportion of the same category classified by prefixes and non-inflection suffixes for orthographic and phonological analyses of all words and only polymorphemic words.

Analysis	Corpus	Types			Tokens		
		Prefix	Suffix	U	Prefix	Suffix	U
Orthographic	All Words	.499	.776	916.5***	.675	.812	2305.0***
	Polymorphemic	.494	.806	801.5***	.625	.867	1228.0***
Phonological	All Words	.499	.741	996.5***	.702	.789	2154.0***
	Polymorphemic	.517	.790	777.5***	.666	.843	1368.5***

*** indicates $p < .001$.

Table 4. Experimental materials: Category words by frequency group.

Frequency	*Category A*	*Category B*
High	Tweand	Foth
	Dreng	Vawse
	Klimp	Suwch
Low	Gwemb	Zodge
	Prienk	Thorsh
	Blint	Shufe