St. Clair, M., Monaghan, P. & Christiansen, M.H. (2014). Acquisition of grammatical categories. In P. Brooks & V. Kempe (Eds.), Encyclopedia of Language Development (pp. 253-255). Thousand Oaks, CA: Sage Publications.

Grammatical Categories (Acquisition of)

Grammatical categories, also referred to as word classes, or parts of speech, are groups of words that have a similar syntactic role in sentences. Thus, words belonging to the same category can be interchanged within a sentence without the sentence becoming ungrammatical. Over the last 60 years in the psychology of language, one of the critical questions has been how grammatical structure of language is acquired. At one extreme, nativists propose that the grammatical categories are pre-specified, and the child acquires grammatical categories by mapping particular words onto these categories. A weaker nativist position is that categories are not specified, but semantic features corresponding to grammatical categories are innately determined. Thus, from the nativist perspective, learning entails matching sets of words to the innate semantic features of grammatical categories. At the other extreme, empiricists contend that words' membership of categories as well as the categories themselves can be learned from the child's input alone using general purpose learning mechanisms. The debate has focused on whether grammatical categories can be learned from the input, or whether additional enriched structure is required to explain learning in the face of an apparent poverty of the stimulus. In this entry, we review the experimental and corpus analysis work investigating the extent to which categories can be acquired from input. In particular, we discuss how grammatical categories could be learnt from co-occurrence information across words in addition to reviewing what other linguistic information can differentiate grammatical categories.

In 1952, Charles C. Fries noted that much information about grammatical categories could be acquired from the distributional contexts in which instances of grammatical categories occur. However, these ideas could not be tested until technological advances and improvements in data collection allowed large-scale investigations into how reliably distributional information can predict grammatical categories. Recent investigations have found that information within the structure of language may be very *useful* for identifying and grouping words into grammatical categories. Parallel experimental work has also found the same cues to be *useable* for learning. However, there has been debate over the precise nature of the distributional cues used to differentiate grammatical categories. The main distinction within the contemporary literature is whether grammatical category learning progresses based on frequently occurring non-adjacent dependencies within language ("you___it", where ___ indicates an intervening word), otherwise known as frequent frames, or more simple adjacent dependencies ("you___" or "___it").

Michael P. Maratsos and Mary Anne Chalkley were pioneers in identifying that succeeding and preceding words could have powerful predictive power for the grammatical category of the intervening words. This was later substantiated and empirically tested by large scale corpus analyses. Key concepts in thinking about the usefulness of distributional information in predicting grammatical categories are *precision* and *recall*. *Precision* refers to how accurately any particular distributional cue groups words of the same category together whereas *recall* refers to the extent to which the grouping covers all words of the same category. Additionally, the *coverage* refers to the proportion of words the child is exposed to that are classified – certain cues which provide high precision and recall may at the same time only capture a small percentage of words in the child's language. The usefulness of grouping words by a specific distributional cue can thus be empirically evaluated using these metrics.

At the centre of the debate over which distributional cues are used, Toben Mintz proposed frequent frames as an important cue for learning grammatical categories. He tested the precision and recall of the 45 most frequent frames found within corpora of child directed speech (e.g., "you ___ to", "the ___ is", where the intervening words are grouped together as a candidate grammatical category). Frequent frames result in very high precision, but very low recall, which was due to the same words being captured by multiple frames (e.g., verbs occur in both "you ___ to" and "I ___ it"). To resolve this, merging groups based on the extent to which they overlap across similar distributional contexts can improve recall. However, the coverage of the frequent frames is poor, due to the specificity of the contexts: Words from the same category may not occur in common contexts, as attested to by constructionist grammar approaches and the possibility of item-based learning by children acquiring appropriate word usage. Subsequent experimental studies tested whether human learners could use these cues to infer distinct categories of words. Using an artificial language paradigm, both adult and infant participants aged 12 to 13 months were found to learn a category of nonsense words based on distinct frequent frames, again made up from nonsense words. This was tested by presenting learners with sentences comprising sequences of words that were compatible with the grammatical categories of the artificial language but that had not been heard during training and comparing to sentences comprising unheard sequences where words violated the grammatical category structure. Thus, it has been shown that these frequent frames are both useful and usable; in other words, they are informative and we can learn from them.

Though frequent frames have been shown to induce category learning, it is not clear that the non-adjacent frame information is necessary for this learning. Frequent frames are, by nature, made up of two adjacent cues. So learning categories from "you ____ it" could be due to the "you-it" frame, or alternatively could be due to only the first bigram cue "you ____" or only the second bigram cue "____ it", or some dynamic combination of the

two. Corpus analyses have examined how useful information from only a single adjacent word – bigram cues – are in differentiating grammatical categories. Michelle St Clair and colleagues found that by only looking at the preceding or succeeding word recall was higher, indicating a better capture of all category words with less specific cues, but precision was slightly lower, indicating the groups were not as well defined by grammatical category. However, the coverage differed enormously as bigram cues were able to capture a much higher proportion of the corpora when comparing non-adjacent frame information. Best performance of all was achieved by categorising words based on a flexible combination of both preceding and succeeding bigrams (where "you ___" and "___ it" are simultaneously available for categorisation). Using these flexible frames, the overlap between apparently distinct fixed frames ("you ___ it" and "I ___ it") is immediately available to the learner. As with the frame cues, artificial language studies have found that both adults and children can learn categories from both preceding and succeeding bigram cues.

One resolution to the debate over which distributional cue is used for categorisation may be that bigrams and frequent frames are used successively by the child, with more accessible bigrams providing an initial parse for grammatical category acquisition, providing broad coverage but lower precision, and more specific frames applied later in acquisition to fine tune the bigram categorisation into more precise sub-categories. Such a view would also work well cross-linguistically, as there is debate about whether non-adjacent fixed frames apply to languages with greater word order variability. Corpus analyses of German and Dutch, which have SVO word order in main clauses and SOV word order in subordinate clauses, result in lower levels of precision and recall than analyses of English due to these word order changes. However, bigrams have similar precision and recall in all these languages, and also are useful in languages with free word order, where distributional regularities are determined by conventions of usage rather than rules about word order. Later in development, more language-specific distributional information can then refine the child's initial guesses about the membership and usage of grammatical categories in her language.

Thus far, we have only considered distributional cues for acquisition of grammatical categories. However, a range of other cues have also been shown to be relevant to grammatical categorisation. In languages with a richer morphology than English, it may be that the important distributional information is present within the word through correlations between the word stem and affixes. This seems particularly true in gendered languages, where nuanced distinctions within grammatical categories have to be formed, so nouns not only have to be differentiated from other types of words, but also segregated into masculine, feminine and (sometimes) neutral genders. For example, in Italian morphological markers distinguish many masculine and feminine nouns, with feminine nouns often ending in –a and masculine nouns ending in –o. However, in other gendered languages, such as German, the cues with which these categories are differentiated are more opaque and are only revealed by statistical analyses of language corpora, where numerous, interrelated cues, from phonological, morphological, and semantic sources, are shown to be clustered within gendered categories. For German, each of these cues has a very weak predictive power individually, but when combined are highly predictive. Experimental studies using artificial languages suggest that such phonological cues may be necessary in order to learn these category distinctions. It also appears that the distinctiveness of the gendered cues influences how quickly gender is consistently learned. Within a relatively transparent language, such as Italian, this is learned fairly easily. German gender, as described above, is much less transparent and is not fully acquired until around 8-9 years of age.

In addition to noun gender subcategories, phonological differences between different grammatical categories have also recently been investigated. Corpus analyses of grammatical categories across many different languages have found consistent within-category phonological characteristics, such as reduced vowels, syllabic complexity and syllable length, for distinguishing nouns from verbs, and content from function words. Furthermore, the way in which cues of different types interact suggests a subtle serendipitous arrangement between information sources. For verbs, phonological cues are more reliable, but distributional information is less available, whereas for nouns the opposite pattern emerges: phonological cues are weaker, but distributional cues are stronger for indicating category. Experimental studies combining distributional cues with phonological cues in artificial language learning results in more accurate grammatical categorisation. Indeed, many experiments fail to find evidence of grammatical categorisation from distributional information without additional phonological cues helping to scaffold learning.

Additional cues for deriving grammatical categories, and membership of specific lexical items within a category, are provided by the environment in which language is spoken. Gestural, attentional, and semantic cues all contribute to determining categories. For instance, if a child is confronted with a novel object with a novel texture and then hears "see the wug" they may infer that "wug" refers to the object itself. But if they hear "this feels wuggish" they may then infer that wug refers to the novel texture. This form of syntactic bootstrapping would likely proceed based on the both the –ish suffix as well as the semantic information contained with the verb "feel", which implies that the novel word "wug" may refer to the texture more than the object itself. Additional referential cues to category are available in the same situation, with the caregiver directing attention

to the object, through eye gaze or deictic gesture. Thus, social and semantic context can influence grammatical category inferences about novel words.

Although a full picture of how grammatical categories are acquired has yet to emerge, the current state of the field demonstrates a dynamic process involving the online combination of multiple cues from a variety of linguistic sources. Empirical investigations suggest that highly accurate grammatical categories are detectable from the child's linguistic and social environment, though this does not preclude the possibility of innately specified structure interacting with these environmental information sources. The challenge for future research is to create new and innovative ways of testing various combinations of cues across differing linguistic domains – e.g., syntactic, semantic, pragmatic, social and gestural – and determining whether innate structure is involved in category acquisition.

Michelle C St Clair
University of Cambridge

Padraic Monaghan
Lancaster University

Morten H Christiansen
Cornell University

See also: Grammatical Gender (Acquisition of); Multiple cues in language acquisition; Syntactic development: Construction grammar perspective; Distributional knowledge and language learning

Further Reading

Cartwright, T., & Brent, M. (1997). Syntactic categorization in early language acquisition: formalizing the role of distributional analysis. *Cognition, 63*, 121–170.

Fries, C.C. (1952). *The Structure of English: An Introduction to the Construction of English Sentences*. New York: Harcourt, Brace & Co.

Kelly, M. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99*, 349–363.

Maratsos, M. P., & Chalkley, M. A. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. E. Nelson (Ed.). *Children's language* (Vol. 2, pp. 125–214). New York: Gardner Press.

Mintz, T.H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*, 91-117.

Pinker, S. (1999). *Words and rules: The ingredients of language*. New York: Basic Books.

St Clair, M.C., Monaghan, P., & Christiansen, M.H. (2010). Learning grammatical categories from distributional cues: Flexible frames for language acquisition. *Cognition, 116*, 341-360.