

Use impacts morphological representation

Joan Bybee

Linguistics Department, University of New Mexico, Albuquerque, NM 81731.
jbybee@unm.edu www.unm.edu/~jbybee/

Abstract: The distinction between regular and irregular morphology is not clear-cut enough to suggest two distinct modular structures. Instead, regularity is tied directly to the type frequency of a pattern. Evidence from experiments as well as from naturally occurring sound change suggests that even regular forms have lexical storage. Finally, the development trajectory entailed by the dual-processing model is much more complex than that entailed by associative network models.

The separation of usage from structure is a traditional practice in linguistics dating back to Saussure's (1916/1973) distinction between *langue* and *parole*. Clahsen subscribes to this dichotomy,

seeing linguistic analysis as totally driven by innate structures rather than by linguistic experience. In recent years, many linguists have made great headway by considering an alternative hypothesis, which is that language structure is built up from patterns of language use. For instance, our understanding of the nature of ergativity (DuBois 1987), subjects and passive (Givón 1979), relative clauses (Fox 1987; Fox & Thompson 1990), and tense, aspect, and modality systems (Bybee et al. 1994) has been greatly expanded by attention to the way language use conditions grammatical change over time. These studies demonstrate that no innate structures are necessary to explain the grammar of human language.

The difficulty Clahsen encounters in trying to argue for innate modular components to account for structural differences is that properties of usage often correspond to the proposed properties of structure, and disentangling the two is a major challenge. For instance, the regular "rules" of morphology usually have high type frequency; that is, they apply to a large number of different forms. Clahsen has addressed this challenge by trying to show that there are regular rules that do not have high type frequency. However, there are problems with his proposals, as described below.

First, Clahsen takes the German *-t* participle and *-s* plurals as cases in which frequency and structure do not correspond, claiming that both of these affixes constitute the "regular" alternative but have a lower type frequency than competing "irregular" allomorphs. However, neither of these cases is clear-cut: (1) the claim that "three different frequency measures revealed that (in contrast to English) regular and irregular verb forms have similar frequencies" is based on counting German verbs differently from the way English verbs are counted (sect. 4.6, para. 8). For English we count *write* as one irregular verb even though it occurs with different particles in *write out*, *write up*, *write down*, but Clahsen counts the comparable structures in German with *schreiben*, "to write," such as *aus-schreiben* "to write out, announce," as a different verb from *schreiben*. Insofar as many "irregulars" occur with the prefixed particles (or separable prefixes), this greatly inflates the number of irregulars (Bybee 1995b). In addition, when considering the claim that type frequency conditions productivity, it is important to note that the irregulars cannot be grouped into one class because they have a number of different types of vowel changes (in English as well as in German). Thus the type frequency of the English *-ed* past tense or the German *-t* participle is much higher than the type frequency of any particular class of irregulars. (2) The *-s* plurals do have a very low type frequency compared to any other class of noun plurals. As would be predicted from their low type frequency, they are not free of lexically based similarity effects, contrary to the claim made by Clahsen (see Table 4). Köpcke (1988) showed in a nonce-probe task that subjects tended to use the *-s* plural on nouns that resembled existing *-s* plurals, in particular those ending in full vowels, such as *Autos* and *Pizzas*.

Second, considerable evidence is accumulating to show that even regularly inflected forms show word frequency effects, which suggests that high-frequency regulars are stored in the lexicon rather than derived by rule. Stemberger and MacWhinney (1986) show that high-frequency regular forms are less prone to error than low-frequency regulars, in both naturally occurring and experimentally induced errors. Losiewicz (1992) found that the *-ed* affix on low-frequency regular past tense verbs was significantly longer in acoustic duration than the same affix on high-frequency regulars. Bybee (1999) found that the rate of deletion of final /t/ and /d/ on regular English past tense verbs was higher for high-frequency verbs than for low-frequency verbs. All of these frequency effects are compatible with the hypothesis that high-frequency regulars are stored in the lexicon and accessed directly, whereas low-frequency regulars require some access to the regular affixation schema, in other words, that it is frequency of use that determines the nature of storage and access, not structure.

Finally, consider the development stages entailed by the dual-

processing model. Presumably, at first, all items are stored in memory, for otherwise it would be impossible to segment them into stem and suffix. That is, *play* – *played*, *spill* – *spilled*, and a large number of parallel items must be stored and associated in memory before the suffix can be discovered. Once the suffix is segmented, the question arises of how children identify it as "regular," if it is not because of the strength it gains from type frequency. In Clahsen's model the normal forms of the language are not sufficient for this categorization; the child has to wait until he or she has heard strange forms such as *ringed* (as in *they ringed the city*). Having established the rule, the child not only reorganizes thoroughly by moving the rule to a different module, but he or she must now ensure that it does not apply to irregulars, by also establishing the blocking device that prevents this (Marcus et al. 1992). This is not the end of the reorganization. The regular forms that were previously stored must now be purged from memory so that only irregulars remain.

The alternative is much simpler and to my mind more plausible: The child learns specific words, some with affixes, some without. They are stored in memory and a network of associations among them begins to develop. These associations eventually reveal recurrent subparts, such as *play* in *plays*, *playing*, *played* and *-ed* in *played*, *spilled*, *wanted*. Subparts that occur in more combinations (that is, have a high type frequency) are reinforced more, both because their occurrence in more combinations makes them more segmentable, and because their reuse increases their levels of resting activation, making them more accessible. Words with high frequency are more autonomous from the networks of associations, which means that if they are irregular they can resist regularization (Bybee 1985; 1995b). No blocking devices or major reorganizations into modules are needed, just a growing network of associations with individual items differentiated by accessibility, which is a result of how often they occur in the child's experience with language.