Regular Morphology and the Lexicon

Joan Bybee

Department of Linguistics, University of New Mexico, Albuquerque, New Mexico, USA

Three models of morphological storage and processing are compared: the dual-processing model of Pinker, Marcus and colleagues, the connectionist model of Marchman, Plunkett, Seidenberg and others, and the network model of Bybee and Langacker. In line with predictions made in the latter two frameworks, type frequency of a morphological pattern is shown to be important in determining productivity. In addition, the paper considers the nature of lexical schemas in the network model, which are of two types: source-oriented and product-oriented. The interaction of phonological properties of lexical patterns with frequency and the interaction of type and token frequency are shown to influence degree of productivity. Data are drawn from English, German, Arabic and Hausa.

THE CONTROVERSY

A long-standing debate in the linguistic and psychological literature centres around the representation of morphologically complex words in the grammar and lexicon. It seems as if every conceivable position on this issue has been argued for seriously and debated vigorously at some time in the last 30 years.[1] The view that is now emerging from a consideration of child language development, diachronic change of morphology in the history of languages and psycholinguistic experimentation, shows the representation

[1]In early generative phonology (Chomsky & Halle, 1968), the position was taken that only morphemes were stored in the lexicon and all words with more than one morpheme were formed by rules that concatenated morphemes and by phonological rules that changed feature values. Not long after the presentation of this view, a position that is the polar opposite of it was also defended: in this view, all words of a language, no matter how complex or how regularly inflected, are listed in the lexicon (Vennemann, 1974). Naturally, there are a number of possible positions to take between these two poles. We will consider arguments for one of these alternatives in this paper.

Requests for reprints should be addressed to Joan Bybee, Department of Linguistics, Humanities 526, University of New Mexico, Albuquerque, NM 87131–1196, USA. E-mail: jbybee@carina.unm.edu.

I am grateful to Martin Haspelmath, Klaus-Michael Köpcke, Beth Losiewicz and two anonymous reviewers for their comments on an earlier draft of the manuscript.

of morphology to be both complex and very interesting, reflecting basic cognitive organisational principles.

Studies of the inflectional system for English verbs show a highly productive regular pattern by which *-ed* is suffixed to verb bases, as well as a variety of vowel changes which affect relatively frequent verbs to form their past tense. Even though the latter set is deserving of the label "irregular", since there are several different vowel change patterns, certain classes of verbs within the irregulars show some degree of productivity—that is, the ability to be extended to new items. An important fact about the distribution of regular and irregular verbs in English is that the irregulars constitute a very small proportion of the entire set and they are mostly of very high token frequency. The regular class has by far the highest type frequency.

Several different proposals concerning how regular and irregular morphological formations are stored and processed have emerged in the recent literature. In this paper, I will describe and defend my own proposed model (Bybee, 1985; 1988) and compare it with connectionist proposals and with the dual-processing model of Pinker, Marcus and their colleagues. The main issue to be discussed is the role of type frequency in the determination of regularity or productivity. My model and the connectionist model claim that type frequency is an important determinant of productivity, whereas the dual-processing model denies any role of type frequency in productivity.

In the next section, I outline briefly the properties of the dual-processing and connectionist models and in somewhat more detail outline my own model, showing which properties it shares with these models and how it differs from each of them. I then go on to discuss the motivation for considering type frequency as pivotal in the determination of productivity and make certain predictions concerning the relationship between type and token frequency.

Next, I examine the evidence against type frequency as a major determinant of productivity, taking first the claims of Clahsen and Rothweiler concerning German particles, and showing that in fact under the most reasonable method of counting type frequency, the productive process does have the highest type frequency. I then discuss the case of the German *s*-plurals and argue that to a large extent the productivity patterns of German plurals also reflect their type frequency. I then go on to address the claim that the processing of regulars is of a qualitatively different type to the processing of irregulars. Here data from Hausa come into play, for the Hausa system shows clearly that even the productive patterns have the same properties as lexical patterns: they are organised in a prototype structure and they are product-oriented generalisations (Lobben, 1991). Next, I discuss the diachronic origins of productive morphological patterns which I argue explain why productive patterns tend to be more transparent and agglutinative in structure than non-productive ones. I then consider whether

the study of the acquisition of English past-tense verbs as presented in Marcus et al. (1992) demonstrates a difference between regular and irregular inflection. I will argue that it demonstrates just the opposite. Finally, I discuss evidence presented in Losiewicz (1992) that shows that regularly inflected forms have differential lexical strength, just as irregularly inflected forms do.

THE MODELS

Several recent studies have reached the conclusion that English irregular past-tense forms are listed in the lexicon and that their listing is structured in such a way that generalisations or schemas emerge from the similarities among verb forms (Bybee & Moder, 1983; Bybee & Slobin, 1982; Marcus et al., 1992; Prasada & Pinker, 1993). One group of researchers, whose model I will refer to as the dual-processing model (Clahsen & Rothweiler, 1992; Marcus et al., 1992; 1993; Pinker, 1991; Pinker & Prince, 1994; Prasada & Pinker, 1993), argue that irregular inflections are lexically represented but regular inflections are derived by a symbolic rule applying to an underlying form to produce the complex surface word. Thus they claim that two highly distinct types of processing are involved in dealing with regular versus irregular inflection. Unlike the other two approaches to be discussed here, these researchers argue that the distinction between regular and irregular morphology is a structural difference independent of the type frequency of the morphological patterns.

Connectionist models do not formulate autonomous symbolic rules for regular inflection, but rather treat all inflection in the same way: a network of mappings from base form to past-tense form is constructed. The network memorises individual patterns and their transformations and generalises on the basis of regularities found in the network. Rumelhart and McClelland (1986) show that a network so constructed can produce output similar to that of a child acquiring English past forms without formulating an explicit rule. In this simulation, type and token frequency are manipulated to reflect, at least approximately, the distribution of irregular and regular verbs in the English lexicon and the way children might encounter them: In the first phase, a small set of mostly irregular verbs is presented and learned, followed by a much larger set of mostly regular verbs.

Subsequent connectionist models have responded to criticisms of this first simulation and have brought connectionist modelling closer to representing the actual situation encountered by children acquiring their language. The models of Plunkett and Marchman (1991) and MacWhinney and Leinbach (1991) each used input that was closer to the actual English input to children. Plunkett and Marchman (1991) also showed that learning was enhanced when phonological sub-regularities (such as those found in *ring*, *rang* and

sing, *sang*) could be taken into account by the model. MacWhinney and Leinbach (1991) incorporated semantic features into their model so that it could differentiate homophonous pairs such as *ring*, *rang* versus *wring*, *wrung*. Cottrell and Plunkett (1991) devised a model which does not just learn mappings between base and past forms, but also can learn and access past forms directly. Each improvement in modelling makes more plausible the claim that all English past tense could be acquired and produced by an associative network without the formulation of a symbolic rule.

The model of Bybee (1985; 1988) was developed to account for cross-linguistic, diachronic and acquisition patterns in complex morphological systems. The basic proposal is that morphological properties of words, paradigms and morphological patterns once described as rules emerge from associations made among related words in lexical representation. A major difference between this model, which I will call the *network* model, and structuralist models containing rules, is that actual usage in terms of both type and token frequency plays an important role in establishing and maintaining representation. Langacker (1987; 1988) has also proposed a model of grammar based on usage with the same properties as the model I will describe here. Connectionist simulations could be thought of as testing some of the properties of the network model and Langacker's cognitive grammar, but the model itself is more complex and accounts for more phenomena than any existing connectionist model. The relevant properties of this model are described here.

Words entered in the lexicon have varying degrees of *lexical strength*, due primarily to their token frequency. Words with high lexical strength are easy to access, serve as the bases of morphological relations and exhibit an autonomy that makes them resistant to change and prone to semantic independence. Thus lexical strength explains why irregular formations are usually of high token frequency: In a language such as English, with both irregulars and a strong regular pattern, irregulars will tend to regularise unless they are sufficiently available in the input to create a strong lexical representation. Thus if the irregular past has low token frequency and is thus more difficult to access, a regular form might be created. Note that in this model it is to be expected that a few paradigms might have inflectional doublets such as *weeped*, *wept* and *creaped*, *crept*. Within paradigms, words with higher lexical strength serve as the basis for the formation of new words (Bybee, 1985).

Words entered in the lexicon are related to other words via sets of lexical connections between identical and similar phonological and semantic features. These connections among items have the effect of yielding an internal morphological analysis of complex words, as shown in Fig. 1. Even though words entered in the lexicon are not broken up into their constituent morphemes, their morphological structure emerges from the connections

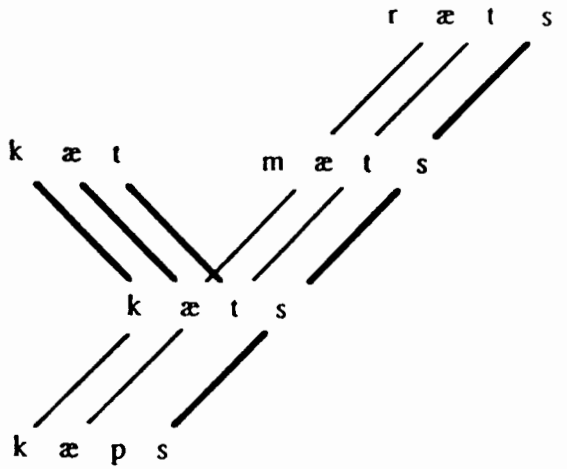


FIG. 1. Sets of lexical connections yielding word-internal morphological structure.

they make with other words in the lexicon. Parallel sets of phonological and semantic connections, if they are repeated across multiple sets of words, constitute morphological relations (in Fig. 1, these are represented by heavier lines). Note that in Fig. 1 connections between base ([kæt]) and complex form ([kæts]) exist, as well as connections among complex forms ([ræts], [mæts], [kæts], [kæps]).

Lexical connections can vary in strength according to the type and number of features shared. Variations in strength of connections due to semantic features in verbal paradigms are discussed in Bybee (1985). Weaker semantic connections are reflected in phonological form by a greater degree of stem change or even suppletion among related forms (Bybee, 1985). For example, stem changes and suppletion are more common among forms that differ in tense or aspect (*break*, *broke*) than among forms that differ only by person/number (*break*, *breaks*).

Strength of lexical connections can also be affected by token frequency, as reflected in lexical strength. Words that have high token frequency have greater lexical autonomy and one reflection of this is that such words form weaker connections with other items. This relation is based on the common-sense observation that items that are of high frequency in the input can be learned on their own terms, while lower-frequency items are better learned in relation to existing items. One of the facts accounted for by the hypothesis that high-frequency words form weaker connections is suppletion. True suppletion (paradigms consisting of stems of different etymologies, such as *go*, *went* or *am*, *was*) occurs only in the most frequent of paradigms.

Suppletion requires the splitting of paradigms—*went* had to separate from *wend* in order to become the past tense of *go*. My hypothesis is that the increase in the frequency of *went* and its consequent greater lexical autonomy contributed to its split from *wend*.

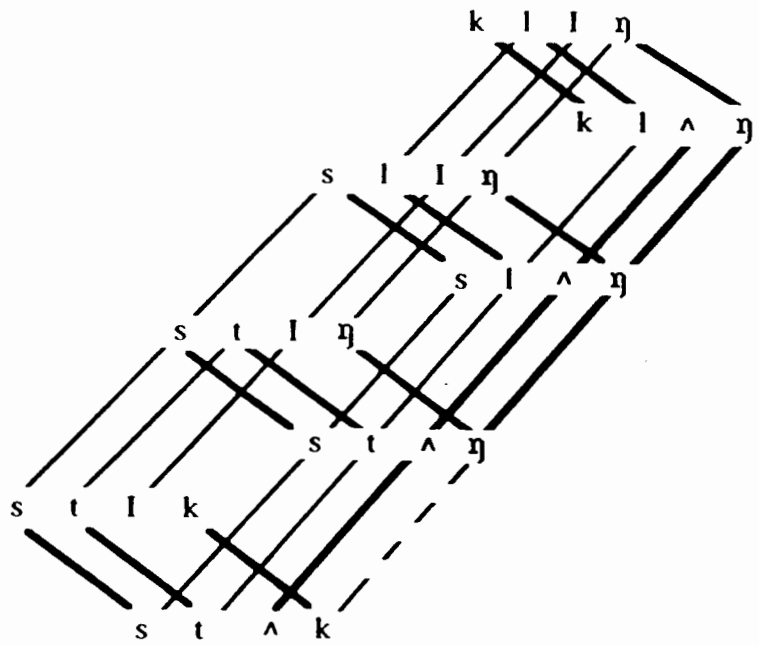
Sets of words having similar patterns of semantic and phonological connections reinforce one another and create emergent generalisations describable as schemas. New items or items whose connections are not known or are weak can be fitted into these schemas. The likelihood of the schema being extended to new items is directly dependent upon two factors: (i) the defining properties of the schema and (ii) its strength, the latter property being derivable from the number of items that reinforce the schema. If the defining properties of the schema are very specific, the schema will be restricted in its application to new forms, and result in lower productivity. If the schema is very open, placing few restrictions on the items to which it can apply, its productivity will be greater. The other determinant of productivity is the strength of the schema, which is based directly on its type frequency—the higher the type frequency of the pattern described in the schema, the greater are its chances of applying to new items.

There are two types of schemas, corresponding to the two ways that morphologically complex forms can relate to other forms. The first, source-oriented schemas, are generalisations over pairs of basic and derived forms, such as *wait, waited*. These correspond roughly to generative rules, since they can be thought of as instructions for how to modify one form in order to derive another. The second, product-oriented schemas, have no counterpart in generative theory. They are generalisations over sets of complex or "derived" forms, such as *strung, stung, flung, hung*, etc., which show what features these derived forms have but without stipulating the operations it takes to produce such forms (Zager, 1980). Membership in product-oriented classes is based on family resemblances: There are more and less central members of the classes, with the central members, such as *strung*, sharing more phonological features with other members of the class than the more marginal ones, such as *dug*.

Product-oriented generalisations were studied by Bybee and Moder (1983) via the class of verbs that form their past tense like *strung*, and it was found that among nonce verbs the least important feature of the input string was the vowel, even though it appears to be the vowel that undergoes a change. In other words, the defining features of the class are not based on the present form, *string*, but rather on the past, *strung*. The data from diachronic change support this view—new members of this class (in certain dialects) include verbs that do not have [ɪ] as their base vowel: *strike* [ai], *struck*; *sneak* [i], *snuck*; *drag* [æ], *drug*. Wang and Derwing (1994) showed in a nonce-probe task that generalisations governing vowel changes in English past tense, *-ity* nominalisations and pluralisation are based on product-oriented

schemas, since certain vowels are favoured for each morphological pattern without regard to the quality of the input vowel.

Since in this model lexical connections relate basic and derived forms as well as derived forms of different paradigms, and since there is no attempt to avoid redundancy, both source-oriented and product-oriented schemas may exist for the same morphological relation. Above I described the regular English past-tense suffixation as a source-oriented schema, since there is a standard operation for forming a past from a base stem ("add /t/, /d/ or /ɨd/"). But this pattern also appears to be product-oriented in the case of the set of verbs that undergo no change in the past tense (e.g. *put, set, quit, cut, spread*, etc.). The verbs of this class all end in /t/ or /d/ and have lax vowels (with the exception of *beat*, which has a tense vowel). They could be described as fitting a product-oriented schema that requires past tense to end in a lax vowel and alveoar stop.



$$\left[\text{C (C) (C) } \wedge \left\{\begin{matrix}\text{velar}\\ \text{nasal}\end{matrix}\right\}\right]_{\text{past}}$$

FIG. 2. Lexical connections and the schema for a part of the *strung* class.

I will now compare the network model with the dual-processing model and connectionist models.[2] The connectionist models and the network model are similar in not postulating a discrete cut-off point between regular and irregular morphology, as the dual-processing model does. Connectionist and network models attribute differences between regular and irregular morphology to quantitative differences, in particular to the relatively greater type frequency of regular patterns.

In the dual-processing model, irregular past-tense forms are listed lexically but regular ones are not (except in special circumstances; see Pinker & Prince, 1994). The regular/irregular distinction, then, corresponds to an important structural distinction that is part of the innate architecture of the language acquisition device: this model has a rules component that is separate from the lexicon. In my network model (and in connectionist models), regular and irregular forms and patterns are treated in the same way and there is no separate component for morphological rules. Differences in degrees of productivity are not attributed to different processing types, but to differences in type frequency and the openness of defining schemas.

Another difference between the dual-processing model and the network model concerns the treatment of irregulars. The dual-processing model is said to handle these in associative networks, but the nature of the generalisations made about them is not specified. In particular, I find no mention of product-oriented generalisations in the literature on dual-processing.

For the purposes of the present paper, connectionist models and the network model make essentially the same claim: type frequency is a major determinant of productivity. However, there are differences between my model and existing connectionist ones. In particular, the treatment of token frequency in connectionist models is taken to be the frequency of the mapping between base and derived form (Hare & Elman, 1995; Plunkett & Marchman, 1991). In the network model, it is the frequency of the derived word itself that is significant and is represented as lexical strength, not a mapping between two forms. In fact, the higher the frequency of the derived form, the weaker the mapping between it and the basic form. High-frequency irregulars are resistant to regularisation not because their connections with their base are better established, but because they are themselves lexically stronger. As I explained above, postulating that high-frequency irregulars have high lexical strength explains not only their

[2] The network model is compared with level-ordered phonology in Bybee (forthcoming). In the final section of this paper, I explain how the network model handles the regular inflection of verbs derived from nouns.

resistance to change, but also their propensity to undergo the paradigm splits that lead to suppletion.

Another difference between existing connectionist models and the network model is that connectionist models at present form generalisation only over relations between base and derived form and do not form generalisations over sets of derived forms. That is, connectionist models do not form product-oriented generalisations. Moreover, I see no sense in which connectionist models abstract schemas from relations among words, as proposed in the network and cognitive grammar models.

THE ROLE OF TYPE AND TOKEN FREQUENCY IN DETERMINING PRODUCTIVITY

The evidence for a pivotal role for type frequency in determining productivity comes from child language and from examining the distribution of morphological patterns in many different languages. In Spanish, French and other Romance languages, verbs are distributed across several conjugation classes, with the largest number of types belonging to what is termed the "first conjugation" (Spanish *-ar* verbs and French verbs such as *chanter*, "to sing"). The other conjugations, though not necessarily irregular, tend to contain fewer verbs and these verbs are typically of high token frequency. Guillaume (1927/1973) observed that the innovations in verb forms created by children acquiring French most commonly involved the use of the first conjugation inflections on verbs belonging to other conjugations in the adult language. Counting the verbs used by nursery school children during play revealed that they were generalising not the verb class that was used most often, but the class that had the highest type frequency, as shown in Table 1. All of the children studied over-generalised the first conjugation. Furthermore, in the adult language, the first conjugation is the one most often used for inflecting new verbs.

Forms with high individual token frequency can be learned by rote, and they can be autonomous, even from other members of their own paradigms.

TABLE 1
Count of Verbs Used by French Nursery School Children During Play

Conjugation Class	*Number of Uses*	*Number of Verbs*
First (*chanter*)	1060 (36.2%)	124 (76.0%)
Second (*finir*)	173 (6.0%)	10 (6.1%)
Third (*vendre*)	1706 (57.8%)	29 (17.9%)

From Guillaume (1927/1973).

With sufficient availability in the input, they can be acquired without forming relations with other items, and without undergoing internal analysis (many high-frequency forms are more fused than low-frequency ones) (Bybee, 1985; Guillaume, 1927/1973).

Forms with lower token frequency will be learned more easily if they can be related to other stored forms. If the input contains a large number of distinct items which share an affix, these words will be related to one another in the lexicon, and the existence of the affix will emerge, as shown in Fig. 1 (Bybee, 1985). The more forms that bear the affix, the stronger the representation of that affix. The stronger the representation of the affix, the easier it will be to access when a new word needs to be inflected, and the greater likelihood that that affix will be productive.

If these hypotheses are correct, then forms of high token frequency will be more autonomous and more likely to be unanalysed, and less likely to participate in schemas; high token frequency forms will thus not contribute to the productivity of a pattern (Bybee, 1985; Guillaume, 1927/1973; Moder, 1992). For example, two English irregular classes of approximately the same type frequency exhibit differences in productivity because one has more high-frequency members than the other. The *strung* class with 13 members, whose total token frequency according to Francis and Kučera (1982) is 199, is much more productive than the *swept* class, which has 14 members with a total token frequency of 656 (Moder, 1992). Baayen and Lieber (1991) also argue, based on intuitive judgements of productivity of derivational morphology, that productivity is governed by type frequency, and that high token frequency can actually detract from productivity. Wang and Derwing (1994) on English and Lobben (1991) on Hausa (to be discussed below) find the likelihood of the use of a pattern in nonce-probe tasks corresponds to the type frequency of that pattern in existing forms.

Given this interaction of type and token frequency, it follows that a high-frequency member of a (semi-)productive class does not contribute to that productivity. Thus *begin*, the highest frequency member of the *ring*, *rang*, *rung* class, may not actually be associated with this class in the mental lexicon of real speakers. (We can observe that no new members of this class have two syllables or end in the alveolar nasal.) Moder (1992) tested this hypothesis by priming subjects with high- and medium-frequency members of English irregular verb classes before presenting them with lists of nonce verbs resembling members of these classes. When primed with medium-frequency irregulars, the subjects produced more nonce irregular past-tense forms than when primed with high-frequency irregulars. While it might seem intuitively that individual high-frequency irregular forms would better serve as models for new formations, the fact that high-frequency items tend to be stored unanalysed makes them less usable than items that enter into relations with other lexical items.

Having now explained how high type frequency contributes to productivity, it is necessary to add that an additional factor needs to be taken into account. A pattern cannot attain full productivity if there are restrictions—phonological, semantic or morphological—on its applicability. The *strung* pattern favoured for verbs beginning in consonant clusters and ending in a velar and/or nasal will never be fully productive, since most verbs do not meet these conditions. Thus productivity also depends upon the openness of the defining schema. Only a schema that is totally open, such as the English past in *-ed* or plural in *-s*, can attain full productivity. This second factor explains how new patterns can attain high type frequency: if they have an open schema, they will gradually increase in type frequency. Two cases where openness of the schema and high type frequency do not correspond are discussed on pp. 438–443. We turn now to the arguments that have been offered against the role of type frequency in determining productivity.

THE EVIDENCE FROM GERMAN PARTICIPLES

Clahsen and Rothweiler (1992) discuss the German system of past-participle formation in which there are two possible suffixes, *-en* and *-t*. The former suffix occurs on verbs that also undergo some stem changes, and the latter suffix usually entails no stem changes, although a small set of 13 verbs has both the *-t* suffix and stem changes. Clahsen and Rothweiler present considerable evidence that the productive or regular affix for past-participles in German is the *-t* affix. They claim that this fact is theoretically interesting because the *-t* affix does not have a higher type frequency than the *-en* affix. Thus they claim that the determinant of productivity or regularity is not type frequency.

The figures for type and token frequency that they present from Ruoff (1990) are from the first 1000 verbs he lists, which constitute 96% of all tokens, as shown in Table 2. If mixed and weak verbs (all those with the *-t* suffix for the past-participle) are taken together, their type and token frequencies are very similar to those of the strong verbs. Thus, Clahsen and Rothweiler argue, the type frequency of the weak and mixed verbs cannot be responsible for the productivity of the *-t* suffix.

This argument depends to a great extent on the way type frequency is counted for German verbs, and I argue that the method chosen by Clahsen and Rothweiler (1992) artificially inflates the type frequency of the irregular verbs by counting the same verbs with productive prefixation as instances of different types. When they first introduce the strong verbs, they note that there are approximately 160 of them. However, by counting verbs with productive prefixes as different types, they substantially increase the number of strong verb types. In addition, by cutting off the count after the first 1000 verbs, they count only one-quarter of the total number of verbs in

TABLE 2
Type and Token Frequency of the Most Frequent 1000 German Verbs

	Tokens	*Types*
Strong verbs	47%	50% (502)
Mixed verbs	32%	5% (50)
Weak verbs	17%	45% (448)
(Mixed verbs have the *-t* suffix and some internal stem changes)		
Examples:		
Strong: *schlafen, schlief, geschlafen*		"sleep"
Mixed: *rennen, rannte, gerannt*		"run"
Weak: *kaufen, kaufte, gekauft*		"buy"

Based on Ruoff (1990).

the Ruoff database, which substantially lowers the number of weak verb types because they tend to be of lower frequency than the strong verbs. Let us consider these two problems.

The productive prefixes on German verbs are known as "separable prefixes" because they occur in two positions: prefixed to the infinitive or participial forms of the verb, or at the end of the clause when the verb is in a finite form of the present, past or imperative, as shown in the following examples:

1. Seine Frau wird morgen *abreisen.*
   his wife FUT tomorrow away-travel
   "His wife will set out tomorrow".

2. ER *reist* *heute* *ab.*
   he travels today away
   "He sets out today".

Even when prefixed, the separable prefix comes before the past-participle prefix *ge-*, as in the form *abgereist*. This German formation is productive, even though not all combinations still have predictable meaning. Thus the prefix *aus*, meaning "out", occurs in combinations such as: *ausgehen* (literally "out-to go"), with the meaning "go out" or "end, finish"; *ausdenken* ("out-to think"), with the meaning "think out" or "elaborate"; and *ausschreiben* ("out-to write"), meaning "write out", "announce" or "advertise". This German construction is fully analogous to the English verb–particle combinations which are distinguishable by the fact that the particle may be separated from the verb by an object pronoun or noun phrase. The following are some examples:

3. I *wrote* the number *down.*

4. We *thought* it *out* carefully.

Even though the combination of the verb and particle constitutes a lexical item by virtue of its semantic coherence, there is no question that the verbs *write* and *think* in these examples are identified as instances of the simplex verbs *write* and *think* and not different lexical types. Part of the evidence for this is that they have the same morphological properties as the verbs which occur without the particles, while verbs homophonous with strong verbs but constituting different types are regular: English *ring, ringed* (as in "to form a ring around") and *grandstanded*. The same is true for German verbs with and without their separable prefixes: strong verbs stay strong with and without the prefixes.

In their count of the type frequency of German verbs, Clahsen and Rothweiler (1992) counted the same verbs with different prefixes as different types. The verb *schreiben* ("to write"), for instance, was counted 19 times because it occurred as a simplex verb and in 18 combinations with separable prefixes. This would be comparable to counting each instance of English *write* in *write up, write down, write over*, etc., as distinct verb types rather than as the same verb.

The frequency count used by Clahsen and Rothweiler, that of Ruoff (1990), also gives a count of *Grundverben*—that is, the base verbs, in which all instances of a verb, with and without separable prefixes, are counted as one type. In this list, there are 1258 base verbs, of which 150 are strong verbs and have the *-en* suffix. The remaining 1108 are weak or mixed and have the *-t* suffix. Thus when only base verbs are counted, weak and mixed verbs constitute 88% of the total and strong verbs only 12%. Contrary to the claim made by Clahsen and Rothweiler, the relative distribution of strong to weak verbs in German is very similar to that of English. It is no surprise, then, that the weak verb participle suffix *-t* is the productive suffix in German.[3]

As for their token frequency, if we exclude the auxiliaries *haben* and *sein*, which account for about one-half of the verb tokens, there are 56,331 verb tokens in Ruoff's database of 500,000 words. Of these, 23,608 are strong verbs, leaving 32,723 weak verbs (and a few mixed ones). Given only 150 strong verb types, it is clear that, as in English, the strong verbs have relatively high token frequency. As argued above, high token frequency actually detracts from productivity, because lexical items that are extremely frequent can be learned more easily and stored without forming lexical connections with other related items. Thus forms with high token frequency do not necessarily contribute to the strength of schemas even if by their form

[3]The implication of the token frequency distribution of base verbs versus verbs with separable prefixes is that strong verbs occur with separable prefixes more often than weak verbs do. Use with prefixes is one of the factors contributing to their high frequency, which in turn contributes to the preservation of their irregularities.

they appear to fit these schemas. It is, then, those schemas that are reinforced by medium- and low-frequency items that tend to be the strongest and have the greater productivity. For this reason, it is not legitimate to cut off the count of German verb types after the most frequent 1000. Ruoff (1990) lists 4414 verbs. Among the 3414 that were not considered by Clahsen and Rothweiler, there would be many more weak verb types than strong verb types, since the former tend to be less frequent. Thus by counting only one-quarter of the database, they have not given an accurate representation of the relative number of weak verb types.

The conclusion here, then, is that the German past participles do not provide evidence against the proposal that productivity of a morphological pattern is directly related to the number of lexical items which display the pattern. In fact, this case constitutes another instance in which productivity corresponds to high type frequency.

GERMAN AND ARABIC PLURALS

As mentioned above, there appear to be at least two main factors that determine the productivity of a morphological pattern: the type frequency and the extent to which the defining schema of the pattern is open to a wide variety of phonological types (see Janda, 1990). In the case of English past-tense *-ed* and the German past-participle in *-t*, both of these factors are operative—both patterns have high type frequency and a completely open schema. In the cases to which we now turn, only the latter factor is operative. The German plural in *-s* and the Arabic "sound" plural have open schemas, but relatively low type frequency. These cases allow us to assess the role of type frequency in determining productivity.

Marcus et al. (1993) argue that these two morphological formations are governed by "symbolic rules", since they demonstrate productivity, and that their productivity cannot be the result of type frequency. However, their experimental results on German, those of Köpcke (1988) and German child language data show that the German plural *-s* has a much lower degree of productivity than English past-tense *-ed*. I will argue that the difference in productivity is a direct result of the difference in type frequency.

For the following discussion, it is useful to distinguish between the terms "productive" and "default". I have been using a gradient notion of productivity coinciding with the ability of a pattern to apply to novel forms. The term "default" as used by Marcus et al. (1993) and Prasada and Pinker (1993) is different: it is not intended to be gradient, but is meant to single out the methods of inflection that are used in various "emergency" circumstances when a plural or other inflected form is non-existent or unknown. The default pattern must be capable of occurring with words of any phonological make-up, so it must be completely open in the way it

defines the items to which it may apply. It is this property that leads Marcus et al. (1993) and Prasada and Pinker (1993) to regard default patterns as processed by symbolic rules. Marcus et al. (1993, p. 7) say: "The hallmark of a symbol is that it can uniformly represent an entire class of individuals, suppressing the distinctions among them". Thus, they would regard the *-s* suffix of German pluralisation as applying only to words marked as nouns, with no further specification. Symbolic rules are qualitatively different from schemas that emerge from lexical patterns. Symbolic rules are independent of any of the forms to which they apply and are unaffected by lexical patterning.

In the following, we will see that the German *-s* pluralisation is used in various emergency situations as a default. However, this does not mean that this pattern is processed by symbolic rule. Rather, evidence indicates even this default pattern functions like a lexical schema—its application is highly affected by existing lexical items.

Table 3 shows the five German suffixes used for plural formation and indicates their type frequency according to a study by Janda (1990) based on a corpus of 600,000 words in tape-recorded interviews (Pfeffer, 1964). The percentages given are from a count of the 200 most frequent nouns in that corpus. The table shows that there are two fairly robust patterns of pluralisation for German nouns. What it does not show is that there is also a strong association of pluralisation in *-(e)n* with feminine nouns, of *-e* with

TABLE 3
German Noun Plural Formation in the 200 Most Frequent Nouns

	Singular	*Plural*	*Gloss*
-(e)n 42%	die Strasse	die Strassen	"the street"
	die Frau	die Frauen	"the woman"
	das Bett	die Betten	"the bed"
-e 35%	der Hund	die Hunde	"the dog"
(+ umlaut)	die Kuh	die Kühe	"the cow"
zero 12%	der Daumen	die Daumen	"the thumb"
(+ umlaut)[a]	die Mutter	die Mütter	"the mother"
	das Leben	die Leben	"the life"
-er 10%	das Kind	die Kinder	"the child"
(+ umlaut)	der Wald	die Wälder	"the forest"
-s 1%	das Auto	die Autos	"the car"
	der Park	die Parks	"the park"

[a]Umlaut (the fronting of a back vowel) is a marker of plural sometimes used in conjunction with a suffix.

Based on Janda (1990).

masculine nouns, and of *-er* with neuter nouns (Köpcke, 1988). In general, pluralisation in German is heavily influenced by gender assignment. The conditioning for the zero allomorph is phonological: it most commonly occurs with nouns ending in *-en*, *-el* and *-er*. However, again gender comes into play: the zero occurs with masculine or neuter nouns with these endings, but not with feminines.

The *-s* alternative has extremely low type frequency in this sample of 200 nouns. It should be noted that 200 nouns is an extremely small sample, and may not be relevant at all for determining productivity. (German and English both have nearly 200 strong verbs each, but these do not include very many productive patterns.) In a larger corpus of 6 million words, where there are 4571 nouns stems, the type frequency of *-s* increases to 7.2% of the types (Marcus et al., 1993). Such an increase is expected of a pluralisation type that tends to occur with infrequent words. The type frequency is still very low, but as noted above, high token frequency detracts from productivity, so that it is the type frequency among infrequent words that is the most important determinant of productivity. However, for the sake of the argument, we will assume that the type frequency of *-s* is too low to account for its productivity, and we now turn to the question of just how productive *-s* is.

The evidence for productivity shows some competition among the pluralisation patterns. Children generalise the *-(e)n* plural most frequently (Mills, 1985; Park, 1978; but see also Clahsen, Rothweiler, & Woest, 1992). Recent loan words favour *-(e)n* for feminines and *-s* with masculines and neuters. The plural *-e* is also strongly represented in masculine gender borrowings (Köpcke, 1988, p. 325). Marcus et al. (1993) summarise Köpcke's analysis by saying that about half of the 182 recent loan words he analysed took *-s* as the plural suffix. In Köpcke's nonce-probe task, *-n* was strongly favoured for nouns ending in schwa in all genders, and for monosyllabic nouns *-en* was favoured for feminines and *-e* for masculines; neuters were split between *-en* and *-e*. Fewer than 15% of monosyllables had responses with *-s*. The *-s* suffix was favoured only in nouns ending in a full vowel, an environment which almost always takes *-s* in the German lexicon (Janda, 1990).

Compared with English past-tense *-ed* and German past-participle in *-t*, which strongly outweigh other alternatives in borrowed words and nonce forms, the German plural *-s* shows only weak productivity, and in fact is used less often than other alternatives. It does seem to be favoured under certain conditions, however. In a task in which subjects evaluated the naturalness of the plurals of nonce nouns, a distinction was made among ordinary nonce nouns, proper names and borrowings, by introducing the nonce words in appropriate contexts that identified their status (Marcus et al., 1993). The results showed that for nonce words that were intended to be German common nouns, a suffix other than *-s* was preferred; for proper names, the *-s* suffix was preferred; and for borrowings, a suffix other than *-s* was preferred if the nonce root rhymed with an existing root, and *-s* was preferred if the nonce root did not rhyme with an existing root. These results can be summarised by saying that if the nonce form could be treated as an ordinary common noun, then the preferred method of pluralisation was one of the suffixes other than *-s*, but if it could not be regarded as a German common noun, the suffix *-s* was preferred.[4]

From these facts, Marcus et al. (1993) appropriately conclude that *-s* is the "default" marker of plural in German. They go on to argue that since *-s* does not have a high type frequency, its status as the default allomorph cannot be due to its having formed a strong schema in an associative network. They argue that the pluralisation process that uses *-s* is a symbolic rule which is qualitatively different from the lexical patterns that account for the productivity of the other pluralisation patterns. As such, it is on a par with the English past-tense *-ed* and the German past-participle *-t*, even though it does not have a comparable type frequency. Their point is that type frequency is not a determinant of productivity or "default" status.

While Marcus et al. have made an adequate case for regarding the *-s* of German plural as an "emergency" or default pluralisation process, the data do not support the conclusion that morphological formations with *-s* are unaffected by type frequency nor that *-s* pluralisation is qualitatively different from the other pluralisation patterns, in that it is an independent, symbolic rule. The data on German *-s* pluralisation confirm the prediction that the openness of the phonological definition of a schema is independent of the type frequency of the schema, and that it is type frequency which is important in determining productivity. Even though the *-s* pluralisation pattern has totally open defining conditions, it shows reduced productivity compared with the English past-tense *-ed* or the German past-participle in *-t*, because of its low type frequency. In particular, since it is restricted to loan words and proper names, it is productive only within these lexical domains and for certain special formations that have no established plural. Since type frequency affects lexical schemas but not symbolic rules, the evidence suggests that the *-s* pluralisation pattern is a lexical pattern, represented by a schema rather than a symbolic rule.

If *-s* pluralisation is a lexical schema rather than a symbolic rule, then existing lexical items will have an effect on its application to nonce items. This prediction is also borne out in Köpcke's (1988) experiment: When he

[4] In fact, Köpcke (pers. comm.) reports that recent loans, which are not yet integrated into the phonological, morphological and graphemic system, take *-s*. But as soon as they are integrated, the same nouns change from using *-s*, to *-en* (*Rivals* > *Rivalen*, "rival"), *-e* (*Barons* > *Barone*, "baron"), or -0 (*Computers* > *Computer*, "computer"), depending upon their gender and phonological shape.

asked subjects to supply plural forms for nonce words that were intended to be normal German nouns, the percentage of *-s* plurals was very low (ranging from 1 to 20%, depending upon the phonological shape and gender of the stem), except for nouns that ended in a full vowel, where *-s* plurals were used 69% of the time. This result reflects the fact that almost all existing German nouns that end in a full vowel form their plural with *-s*.[5] Prasada and Pinker (1993) and Marcus et al. (1993) state that a primary difference between associative lexical patterns and symbolic rules is that the former are affected by the distribution of items in the lexicon while the latter are not. Since *-s* pluralisation is favoured for nonce nouns matching the phonological shape of a cluster of real nouns that take *-s*, the conclusion must be that *-s* pluralisation is not a symbolic rule, but rather a lexical pattern. Even though it can be used as the default plural, it is not qualitatively different from other pluralisation mechanisms.

Another case that has been invoked to argue in favour of a qualitative difference between schemas and rules that is not based on type frequency is the pluralisation process in Arabic. The productive pattern in Arabic, called the "iambic broken plural", requires changes internal to the noun root, and affects nouns of different phonological shape in different ways (McCarthy & Prince, 1990). The following are some examples from McCarthy and Prince (1990, p. 217):

5.

singular	*plural*	*gloss*
nafs	nufuus	"soul"
rajul	rijaal	"man"
jundub	janaadib	"locust"

McCarthy and Prince (1990, p. 214) cite this pattern as "truly productive" because it will apply to any loan word that has a stem that meets its canonical shape criteria. However, there is another pattern, involving suffixation (called the "sound plural"), that can be applied in what might be regarded as "default" circumstances, such as with proper names, transparently derived nouns or adjectives, diminutives, non-canonical or unassimilated loans, and the names of the letters of the alphabet, which are mostly non-canonical (McCarthy & Prince, 1990, p. 212). The sound plural has a very low type frequency, since it is only used with non-canonical stems. As might be expected, however, children over-generalise both patterns (Omar, 1973).

This situation appears to be very similar to the situation with German pluralisation. There is no argument about the sound plural being the default pattern, but it does appear that its productivity is extremely limited, due to its low type frequency. Again, it appears that type frequency does affect the degree of productivity of a default pattern.

[5]Köpcke (1993, pp. 128–133) shows further that the choice of *-s* versus *-en* pluralisation depends upon the quality of the final vowel and the gender assignment.

In this section, we have seen that the two cases put forth as evidence that type frequency does not determine productivity are both cases where the morphological pattern used in emergency situations has very low type frequency. In both cases, however, the patterns in question have very limited productivity compared with patterns with high type frequency in other languages (e.g. English plural *-s*, past-tense *-ed*, and German weak participles). I have argued that this limited productivity is due to the low type frequency. Furthermore, in the case of the German *-s* plural, there is experimental evidence that shows that existing lexical patterns (the fact that nouns ending in full vowels pluralise with *-s*) affect the application of the default pattern to new items. Thus even though the *-s* plural is the default pattern, it is represented in a lexical schema rather than as a symbolic rule.

PRODUCT-ORIENTATION AND HAUSA PLURALS

As explained above, lexical schemas involve both product-oriented and source-oriented generalisations. Product-oriented schemas are generalisations over non-basic forms rather than generalisations about the relation of a non-basic form to some underlying stem or base form. The product-oriented schema specifies only the phonological shape of the member of the category, but does not specify how that shape is to be achieved. In the nonce experiment of Köpcke (1988), some subjects pluralised nouns ending in full vowels, such as *der Treika* and *die Kafti*, by replacing the final vowel with the suffix *-en*, yielding *die Treiken* and *die Kaften*. The resulting plurals matched those in the lexicon and yet the procedure required more than just adding a suffix to the input noun. The subjects were aiming at a product that matches other lexical items and they adjusted the nonce words to achieve this result.

A generative "symbolic" rule corresponds to a source-oriented schema in that it specifies the input shape and details the procedure for changing it. An affixation rule that makes no changes to the stem could be described by a source-oriented rule; for example, for German plurals, suffix *-s* to the end of the stem. A symbolic rule cannot describe a product-oriented generalisation. It appears that the dual-processing model does not allow for product-oriented generalisations, as they are not describable by symbolic rules and the lexical network used for irregulars in this model generalises over base–past pairs such as *swing–swung* and *string–strung* (Marcus et al., 1992, pp. 122–123).

In the cases we have examined so far, the default inflection has been more agglutinative or concatenative than the other lexical patterns. Thus the

English irregular verbs are formed with vowel and sometimes consonant changes to the stem, while the regular formation involves a suffix and no stem change; the German past-participle in *-en* often involves a vowel change, while the pattern with *-t* rarely does; some of the German plural suffixes are accompanied by umlaut, but the suffix *-s* never is; and the Arabic iambic plural involves stem changes while the sound plural is suffixation. Marcus et al. (1993, p. 7) imply that they always expect the default pattern to involve agglutination by calling the processing type for the default a "symbol-concatenating rule". If this is the correct interpretation of Marcus et al., then another proposed qualitative difference between lexical patterns and symbolic rules has emerged: lexical patterns may involve affixation or internal changes to the stem, but symbolic rules are strictly affixing. Since stem changes typically involve product-oriented schemas, and symbol-concatenation is necessarily source-oriented, then we can conclude that default rules in their sense must be source-oriented. In this section and the next, I will argue that default status or productivity is not necessarily associated with source-oriented rules. First, I will present data from Hausa that show that the productive morphological patterns for plural are all product-oriented. Then, in the next section, I will argue that the usual association of source-oriented patterns with default status is due to the way new affixes tend to develop diachronically.

As just mentioned, in Hausa, even the most productive plural patterns are product-oriented. One analysis of Hausa nouns proposes 10 lexical classes, using a combination of criteria which include properties of the singular as well as the plural (Kraft & Kirk-Greene, 1973). More recent analyses by Haspelmath (1989) and Lobben (1991) point out that there are a number of arguments for analysing Hausa plurals in terms of product-oriented schemas that reference the plural forms only. First, there is little predictability between singular and plural form. Just to cite one problem, Lobben (1991, pp. 112–113) notes that of 95 nouns that end in *-aa* and have low–high–low tone patterns, 48 take a plural in *-uu* with a low... high tone pattern and 47 take a plural in *-ooCii* with a high tone.

Second, the plural suffix replaces part of the end of the singular noun, yet the part replaced cannot be considered a singular suffix. For instance, in the singular/plural pairs *tàmbáayàa*, *támbáyóoyíi* ("question") and *táagàa*, *táagóogíi* ("window"), the final vowels *àa* are replaced by *óo* + a copy of the last consonant + *íi* (examples from Lobben, 1991, p. 3). In these examples, the acute accent (*á*) represents a high tone, and the grave accent (*à*) represents a low tone.

But the strongest evidence for product-oriented schemas over plural forms is the fact that for both real and nonce nouns, a single plural pattern can be arrived at in a variety of ways. For example, Haspelmath (1989) gives the following examples of the plural pattern HL *-úCàa*:

6a.	rìigáa	ríigúnàa	"gown"
	kàntíi	kántúnàa	"store"
b.	záurèe	záurúkàa	"porch"
	hánnúu	hánnúwàa	"hand"
c.	árzìkíi	árzúkàa	"prosperity"
	tákòobii	tákúbàa	"sword"

As can be seen in (6), the plurals all have a consistent tone pattern of high–high–low, and they all end in *-úCàa*, where *C* is /n/, /k/ or /w/ or a consonant that is part of the root. In (6a) and (6b), the final vowels of the singular appear to have been removed and a suffix added. But in (6c), the plural morpheme is fitted like a template over the last two syllables of the singular form. Thus it is possible to formulate a single product-oriented schema that describes the shape of the plural form, but it would take numerous source-oriented rules to describe how to derive a plural from a singular.

Since a similar situation holds for all the plural patterns of Hausa, and since Lobben found strong evidence for product-oriented schemas in her nonce experiments with Hausa speakers, she argues that Hausa pluralisation is processed entirely by product-oriented schemas. A further argument, put forward by Haspelmath, is that the individual plural schemas have properties in common, so that it is possible to formulate a meta-schema that unites the various methods of plural formation. Generalisation at this level is not possible using source-oriented rules.

Lobben's nonce word task showed that there is not just one productive pattern, but rather three patterns that show some degree of productivity, each one accounting for more than 20% of the nonce responses. Lobben (1991, p. 223) states that the degree of productivity of a pattern is in direct correspondence with the type frequency of that pattern in Hausa nouns, which is precisely what we would expect from lexical patterning. It is not clear whether or not there is just one "emergency" plural formation strategy in Hausa. The most productive schema in the nonce experiment (high *-óoCíi*), garnering 29% of responses, is also the one used with recent loan words, especially from English. For example, *ásìbítìi* ("hospital") has the plural *ásíbítóocíi*, and *téebùr* ("table") has the plural *téebúróorii*. According to Lobben (1991, p. 223), this schema also has the highest type frequency, and it apparently also has the most open phonological definition for singulars (Lobben, 1991, p. 113), occurring with the highest number of combinations of finals and tone patterns (at least for trisyllabic nouns). It is also the only schema that applies to words (which are all loans) that end in consonants (see the word for "table" above).

The evidence, then, points to this one schema as being (or becoming) comparable to the "default" pattern in other languages. Yet this pattern

cannot be captured in a "symbol-concatenating" or source-oriented rule; it has the same properties as other patterns that are represented in lexical associations, including a prototype structure. In the nonce experiment, Lobben found extensive evidence for a product-oriented treatment of this pattern. In forming plurals with this pattern, subjects added or omitted phonemes in the singular, or made use of the singular consonants and inserted the appropriate vowels, but always arrived at a plural that precisely fitted the schema.

It appears, then, that there is at least one language in which an obligatory inflectional category does not have a symbolic rule to fall back on for creating new plurals, but rather uses only the patterns that emerge from lexical associations. This is not meant to be a devastating argument against the dual-processing model as proposed by Pinker (1991), Marcus et al. (1992; 1993) and Prasada and Pinker (1993). Dual-processing could exist but simply not be used for this inflectional category of Hausa. This case does show, however, that there is no necessary correspondence between the most productive or default patterns and symbolic rules.

DIACHRONIC EXCURSUS

Despite the existence of languages such as Hausa, it is often the case that the productive pattern for a morphological category is a concatenative process, describable in source-oriented terms, while the irregular patterns tend to have more of an effect on the stem and are better describable in product-oriented terms. This cross-linguistic pattern in itself might be taken as evidence for the dual-processing proposal were it not for the fact that this pattern has a straightforward diachronic explanation. In many cases, the most productive pattern is the newer one, derived more recently from a periphrastic construction. In such cases, the diachronic source accounts for both the concatenative structure of the formation and its ability to apply to a wide variety of phonological types.

For instance, the English *-ed* derived from the past form of the ancestor of the modern verb *do*, in a periphrasis in which the past form of *do* stood after the main verb form. Such periphrases usually arise as strategies for forming verbs from nouns and adjectives and are extended to use with loan words, as a way of inflecting them. A strategy such as this is particularly welcome when the existing inflectional patterns requiring stem changes are not easily extendable to new lexical items, and at times when the language is experiencing an influx of borrowed words. These newer, more transparent, inflections can increase in type frequency and hence productivity to the extent that they replace the older means of inflection. Since they are originally periphrastic—that is, expressed as separate words—they impose

no phonological shape requirements on the main verb, which also helps them increase their productivity.

A verb meaning "do" or "be" is usually the auxiliary verb used in such cases (e.g. in Japanese or Basque), but other verbs are possible. A particularly interesting example occurs in the African language Kanuri (Hutchison, 1981; Lukas, 1937), where there are two inflectional classes for verbs. One class consists of approximately 150 verbs, entails complex phonological and morphological irregularities and is unproductive. The other more regular, productive class adds affixes to verb stems that are forms of the verb *ngin*, meaning "to say". Hutchison (1981, pp. 102–111) argues that the original members of this class must have been verbs derived from onomatopoeic words (hence the appropriateness of the verb "to say"), and that the formation spread from this specific use to use with derived verbs from other lexical items and loan words.

Since the major source of new morphology is grammaticisation of separate words, transparent affixes are more easily formed diachronically than patterns that affect a lexical stem (Bybee, Perkins, & Pagliuca, 1994; Heine, Claudi, & Hünnemeyer, 1991; Heine & Reh, 1984). In fact, stem changes are almost always derived from phonological changes in the stem conditioned by an affix (e.g. Germanic umlaut). This means that it typically takes a very long time for stem changes to develop. Since grammaticisation of new affixes is occurring at all times, older affixes may be replaced before they have a chance to fuse with the stem to an extent sufficient to create stem changes. Thus the greater frequency of affixes over stem changes is due to the way the two patterns arise diachronically and may have no bearing on the type of processing involved for the two patterns. While some (e.g. Dressler, 1985, pp. 316–319) would argue that affixes present an acquisition or processing advantage over stem changes, which would coincide with the use of symbolic rules, Bybee and Newman (1994) have shown experimentally that subjects can internalise and generalise stem changes as readily as affixes, suggesting that both are processed in the same manner.[6]

ENGLISH-SPEAKING CHILDREN'S OVER-REGULARISATION OF PAST TENSE

For decades, the most convincing evidence for the existence of productive morphological rules was to be found in children's over-generalisations, such as English-speaking children's regularisations of irregular past-tense forms. As traditionally understood, children's performance on English verbs

[6]Dressler's discussion and examples also confound morphological transparency with regularity. In Bybee and Newman (1994), we try to disentangle these two variables.

appears to yield a U-shaped learning curve. In the first stage, children produce irregular past forms correctly, but after acquiring a number of regular past forms, they begin to produce regularised past forms, such as *goed* and *breaked* (the second stage). Only later do they sort out the irregular from the regular and correctly produce *went* and *broke* again. It is the second stage—the stage in which regularised pasts are produced—that provides the strongest evidence available for rule formation and application in children, because it is at this stage that children produce forms that they do not hear from adults.

Two recent developments now cast doubt on this phenomenon as evidence for symbolic rule formation in children. The connectionist modellers discussed on pp. 427–428 have shown that a parallel distributed processing model, which does not formulate symbolic rules, can nonetheless simulate the three acquisition stages described above, given input that reflects the type and token frequency of English regular and irregular verbs. In a second development, Marcus et al. (1992) examined regularisations in over 10,000 tokens of children's speech and found that the percentage of regularisations is unexpectedly low—only 2.5% of English irregular tokens are regularised. Even though individual children regularise at varying rates (from less than 1% to as much as 24%), Marcus et al (1992, p. 35) conclude: "the global data suggest that overregularization is a relatively rare phenomenon".

This result is surprising because regularisations are so salient in children's speech; apparently, correct irregulars go largely unnoticed by adult observers. Since a presumed high rate of regularisation has always been taken by linguists as evidence that children formulate rules, the actual rarity of regularisations casts doubt on the proposition that children internalise a symbolic rule for regular past tense. Low rates of regularisation demonstrate that the difference between regular and irregular pasts in English is much smaller than originally supposed, and more consistent with a uniform treatment of regular and irregular morphology with lexical representation and lexical schemas for all past forms.

Marcus et al. (1992) do not interpret their data this way. Even in the face of this evidence, they maintain the notion that children formulate a symbolic rule for regular past tense. They explain the low rate of application of this rule to irregulars as the result of a very strong blocking device that suppresses regularisations. Thus their model contains (i) lexical representations and schemas (for irregular forms), (ii) symbolic rules, AND (iii) a powerful blocking device to keep the symbolic rules under control. This can be compared to the more parsimonious model proposed here that requires only (i)—lexical representations and schemas of varying degrees of strength and generality. The network model does not require a blocking device to prevent the regular pattern from overapplying, since it is not a

categorical rule anyway, but simply derives its productivity from its type frequency.

In Rumelhart and McClelland's (1986) simulation, the onset of regularisation coincided with a sharp increase in the input of regular past-tense formations. The data examined by Marcus et al. did not reveal any chronological relationship between the number of regular verbs in the input (or the child's output) and the onset of regularisation. Instead, Marcus et al. argue, convincingly in my opinion, that the onset of regularisation is a result of the child's learning that the marking of tense is obligatory in English. Having established that tense-marking must occur, the child is forced into situations in which a past-tense form must be supplied, even if the appropriate form is not represented in the child's lexicon, or is perhaps only very weakly represented. In the dual-processing model, the symbolic rule takes over in these cases. In the network model, other lexical patterns are accessed, and the strongest of these is the one regarded as regular in English. The important point here is that it is not the formulation of the symbolic rule that triggers regularisation, but rather the child's formulation of the (correct) hypothesis that tense-marking is obligatory in English. The existence of regularisations is independent of the issue of whether or not children formulate a symbolic rule.

In the dual-processing model, which includes both lexical schemas and symbolic rules, the question arises as to how the symbolic rule is acquired. In order for the child to discover the rules, forms with and without the suffix (*walk* and *walked*) must be stored compared, and the pattern extracted from them. In fact, quite a number of such forms must be stored before the child can conclude that the pattern is not in fact appropriate for a lexical schema but must rather be represented in a separate component as a symbolic rule. Thus the child must switch from the lexical schema strategy for representing the regular pattern to the symbolic rule strategy.

Is there any evidence for such a switch in processing strategies? When and how does this switch take place? Since Marcus et al. (1992) did not find any dramatic spurt of regularisations in the data they examined, the point of restructuring from lexical schema to symbolic rule cannot be identified in this way. In fact, for most children, the rate of regularisation, once begun, remains fairly stable. The emergence of regularisation itself cannot be used to identify the point of change in processing type, since this is argued to be due to the emergence of tense-marking as obligatory. Implicit in Marcus and co-workers' argument is the supposition that the symbolic rule is already formed when regularisations begin. While Marcus et al. list a number of possible ways in which the child could determine which past-tense formation process is the productive or default one, they do not address the question of exactly how the rule is extracted and deposited in a separate module. Thus despite the fact that the relevant section in their paper is named "How Might

a Regular Rule Be Learned?", the question of how the restructuring from lexical schema to symbolic rule takes place is actually not addressed.[7]

EVIDENCE FOR THE REPRESENTATION OF HIGH-FREQUENCY REGULARS

The most innovative aspect of the model proposed in Bybee (1985; 1988) is the hypothesis that high-frequency regularly inflected words are stored in the lexicon, while low-frequency regulars are derived in the lexicon by applying the strongest schema to base forms. This follows from the lexical strength proposal—for regular forms just as for irregular forms, lexical strength varies according to frequency of use. Because of the extremely open nature of this schema, experiments have turned up no evidence for effects of lexical distribution such as that found in the case of more restricted schemas (Prasada & Pinker, 1993; Stemberger & MacWhinney, 1988). An approach taken by Losiewicz (1992), however, has resulted in evidence in support of differential representation for high- and low-frequency regular forms.

In a series of experiments, Walsh and Parker (1983) found that English /s/ in word-final position is longer in acoustic duration if it is the plural morpheme (*laps*) than if it is part of a monomorphemic word (*lapse*). Losiewicz (1992) found this same distinction applies to morphemic /d/ or /t/ (the past tense) as in *rapped* versus non-morphemic final /d/ or /t/ as in *rapt*. Losiewicz further reasoned that if the difference in length is due to the non-morphemic segment being part of a lexical representation, while the morphemic one is added to the stem in processing, then the same difference in length should appear in low- versus high-frequency words with morphemic /d/, since the low-frequency words would be formed by using a schema and the high-frequency words would be accessed directly from the lexicon.

Losiewicz asked subjects to read sentences containing English past-tense forms which constituted rhyming pairs of high- and low-frequency verbs (*covered*, *hovered*; *needed*, *kneaded*). For all subjects and all pairs of verbs, the final past-tense morpheme was longer in the low-frequency verb of the pair. The average difference in duration was 7 msec, a difference that was highly significant and not due to overall differences in word length. It is not proposed that such a length difference is either perceptible or learnable from

[7]Clahsen et al. (1992) argue that German children with "specific language impairment" distinguish regular from irregular inflection. Goad (in press), however, suggests that the network model makes the correct prediction for subjects with specific language impairment. It seems that such subjects are able to memorise highly frequent inflected forms, but are not able to form the relations among sets of forms that allow the formation of schemas.

input, but rather that it reflects a difference in processing type. These results, then, can be taken to support the hypothesis that high-frequency inflected verbs are stored in the lexicon, whereas low-frequency inflected forms are produced by applying the strongest schema to base forms. It is conceivable that further investigation of an experimental nature will yield more evidence of differences between high- and low-frequency regular forms.[8]

Faced with the idea that regular morphological forms can be stored in the lexicon just as irregulars are, one could ask why Prasada and Pinker (1993) found no effect of existing lexical verbs in their nonce probe tasks which used the English regular past tense. The answer is that the regular past tense has a totally open schema together with extremely high type frequency, which means it can apply to virtually any string of sounds and thus shows no effect of lexical distribution.[9]

CONCLUSION

To summarise the arguments given, we have first seen that contrary to the claims of Clahsen and Rothweiler, the case of German past-participles is precisely parallel to the case of English past tense in the sense that the productive affix is the one with the highest type frequency. In the case of German plural formation, there is no one pattern with overwhelmingly high type frequency (over 80%, as in the case of English past tense and German participles) and, as predicted, there is no one productive pattern. One pattern of very low type frequency (the *-s* plural) has some productivity, and it is the pattern with the most open phonological definition. The application of this pattern does not exhibit full productivity and does not seem rule-like, since its application to nonce forms is affected by its distribution in the lexicon.

A second property that distinguishes symbolic rules from lexical schemas is source- versus product-orientation. Symbolic rules are necessarily source-oriented and schemas may be either source- or product-oriented. As Hausa plural formation shows, there is no requirement that a language should have a source-oriented means for forming obligatory inflectional categories. All of the means of forming plurals in Hausa are product-oriented and some of these means show productivity. Thus Hausa does not make use of two processing avenues, but uses only one—lexical schemas. Even the evidence

[8]Bybee (1995) studied *t/d* deletion in a corpus of American English and found that regular past and participial /t/ and /d/ are more likely to be deleted in high-frequency verbs than in low-frequency verbs. This finding suggests that high-frequency regulars are listed in the lexicon and low-frequency regulars are produced from a schema, in which in the /t/ and /d/ are less reduced because the schema is influenced by its use with low-frequency verbs.

[9]Daugherty and Seidenberg (1994) show for a connectionist model that a sufficiently high type frequency and lack of phonological coherence in a pattern leads the model to behave eventually as though the pattern were independent of its lexical distribution.

for a symbolic rule in the case of English past-tense formation is considerably weakened by the finding that the actual rate of regularisation of irregular past-tense forms in English-speaking children is quite low. Finally, as to the controversial claim that high-frequency regular past-tense forms are listed in the lexicon, whereas low-frequency forms may not be, Losiewicz (1992) has found a consistent measurable difference in the phonetic shape of the suffix in low- versus high-frequency inflected forms.

The issues discussed here suggest that some systematic distinctions should be made among the terms "productive", "default" and "regular". I propose the following, which are based on the way these terms appear to be used in the relevant literature: *Productive* means that a pattern may be applied to new items. Degrees of productivity are highly correlated with the type frequency of a pattern within a language. *Default* refers to the pattern that applies when all else fails; default patterns have the most open phonological and lexical definitions (where "lexical" refers to features such as loan word and proper noun). The *regular* pattern is the one with the least allomorphy in affix and stem. This term seems also to be used when type frequency and the most open phonological schema converge on the same pattern, as they do in English. It is important to note, however, that productivity, default application and regularity are independent and may be properties of distinct patterns in some languages (see Bybee, in press).

The data cited here, and the arguments given, support a functionalist conception of the lexicon, as a store of words and phrases that is highly affected by actual language use (cf. Bates & MacWhinney, 1982; 1987). The strength of lexical representations of individual items is in part a reflection of token frequency, while the strength (and one determinant of productivity) of lexical associations or schemas is built up by type frequency. The distinction between storage of complex forms versus formation by schema application are based on the availability of stored items, which is determined by frequency of use, not by structural distinctions, such as the classification into regular and irregular patterns. The maintenance of irregularity in high-frequency forms is just one piece of evidence that use overrides structure in determining representation.

The alternative proposal—that structure determines representation—is not supported by the data. The one structural criterion—the openness of the phonological definition of the schema—only gives rise to full productivity in combination with high type frequency. Other aspects of structure, such as agglutinative structure or apparent source-oriented application, are epiphenomenal and due to diachronic patterns. The synchronic reality of varying degrees of productivity is derivable from the way language is used.

Only one "module" is necessary for morphology: a highly structured lexicon. Morphological structure and organisation emerge from the connections made among related stored items. Even forms produced by

combination are produced in the lexicon by accessing a stem and utilising a schema that is available in the lexicon. Given that schemas have the same properties as other types of human categorisations, it is distinctly possible that storage and access of morphologically complex words may require very little cognitive ability that is specific to language.

Manuscript received May 1994
Revised manuscript received May 1995

REFERENCES

Baayen, R., & Lieber, R. (1991). Productivity and English derivation: A corpus-based study. *Linguistics*, *29*, 801–843.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds), *Language acquisition: The state of the art*. New York: Cambridge University Press.

Bates, E., & MacWhinney, B. (1987). Competition, variation and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Bybee, J.L. (1985). *Morphology: A study of the relation between meaning and form*. Philadelphia, PA: John Benjamins.

Bybee, J.L. (1988). Morphology as lexical organization. In M. Hammond & M. Noonan (Eds), *Theoretical morphology*. San Diego, CA: Academic Press.

Bybee, J.L. (1995). Phonology of the lexicon: Evidence from lexical diffusion. Paper presented at the *Rice University Symposium on Usage-based Models of Language*, Houston, TX, March.

Bybee, J.L. (in press). Productivity, regularity and fusion: How language use affects the lexicon. In R. Singh (Ed.), *Morphophonology: Contemporary responses*. Berlin: Mouton-DeGruyter.

Bybee, J.L., & Moder, C.L. (1983). Morphological classes as natural categories. *Language*, *59*, 251–270.

Bybee, J., & Newman, J. (1994). Are stem changes as natural as affixes? *Linguistics*, *32*, 33–34.

Bybee, J.L., & Slobin, D.I. (1982). Rules and schemas in the development and use of English past tense. *Language*, *59*, 251–270.

Bybee, J.L., Perkins, R.D., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago, IL: University of Chicago Press.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Clashen, H., & Rothweiler, M. (1992). Inflectional rules in children's grammars: Evidence from the development of participles in German. *Yearbook of Morphology 1992*, pp. 1–34.

Clahsen, H., Rothweiler, M., & Woest, A. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition*, *45*, 225–255.

Cottrell, G., & Plunkett, K. (1991). Using a recurrent net to learn the past tense. In *Proceedings of the 13th Annual Meeting of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Daugherty, K.G., & Seidenberg, M.S. (1994). Beyond rules and exceptions: A connectionist modeling approach to inflectional morphology. In S.D. Lima, R.L. Corrigan, & G.K. Iverson (Eds), *The reality of linguistic rules*, pp. 353–388. Amsterdam: John Benjamins.

Dressler, W.U. (1985). *Morphophonology: The dynamics of derivation*. Ann Arbor, MI: Karoma Press.
Francis, W.N., & Kučera, H. (1982). *Frequency analysis of English usage*. Boston, MA: Houghton Mifflin.
Goad, H. (in press). Remarks on the network model: Productivity, derivational morphology and atypical populations. In R. Singh (Ed.), *Morphophonology: Contemporary responses*. Berlin: Mouton-DeGruyter.
Guillaume, P. (1927/1973). The development of formal elements in the child's speech. In C.A. Ferguson & D.I. Slobin (Eds), *Studies of child language development*, pp. 240–251. New York: Holt, Rinehart and Winston.
Hare, M., & Elman, J. (1995). Learning and morphological change. *Cognition, 56*, 61–98.
Haspelmath, M. (1989). Schemas in Hausa plural formation: Product-orientation and motivation *vs.* source-orientation and generation. *Buffalo Working Papers in Linguistics, 89–01*, 32–74.
Heine, B., & Reh, M. (1984). *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske Verlag.
Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. Chicago, IL: University of Chicago Press.
Hutchison, J. (1981). *The Kanuri language: A reference grammar*. Madison, WI: University of Wisconsin, African Studies Program.
Janda, R.D. (1990). Frequency, markedness, & morphological change: On predicting the spread of noun-plural *-s* in Modern High German and West Germanic. In *ESCOL '90*, pp. 136–153.
Köpcke, K.-M. (1988). Schemas in German plural formation. *Lingua, 74*, 303–335.
Köpcke, K.-M. (1993). *Schemata bei der Pluralbildung im Deutschen. Versuch einer kognitiven Morphologie*. Tübingen: Narr.
Kraft, C.H., & Kirk-Greene, A.H.M. (1973). *Teach yourself Hausa*. London: Hodder and Stoughton.
Langacker, R. (1987). *Foundations of cognitive grammar, Vol 1. Theoretical prerequisites*. Stanford, CA: Stanford University Press.
Langacker, R. (1988). A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics*, pp. 127–161. Amsterdam: John Benjamins.
Lobben, M. (1991). *Pluralization of Hausa nouns, viewed from psycholinguistic experiments and child language data*. MPhil thesis, University of Oslo.
Losiewicz, B.L. (1992). *The effect of frequency on linguistic morphology*. PhD thesis, University of Texas, Austin, TX.
Lukas, J. (1937). *A study of the Kanuri language: Grammar and vocabulary*. Oxford: Oxford University Press.
MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition, 29*, 121–157.
Marcus, G.F., Pinker, S., Ullman, M., Hollander, M., Rosen, T.J., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57*(4).
Marcus, G.F., Brinkmann, U., Clahsen, H., Wiese, R., Woest, A., & Pinker, S. (1993). *German inflection: The exception that proves the rule*. Occasional Paper No. 47. Center for Cognitive Science, MIT, Cambridge, MA.
McCarthy, J., & Prince, A. (1990) Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory, 8*, 209–283.
Mills, A.E. (1985). The acquisition of German. In D.I. Slobin (Ed.), *The cross-linguistic study of language acquisition, Vol. 1. The data*, pp. 141–254. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Moder, C.L. (1992). *Productivity and categorization in morphological classes*. PhD thesis, State University of New York at Buffalo, NY.
Omar, M. (1973). *The acquisition of Egyptian Arabic as a native language*. The Hague: Mouton.
Park, T.-Z. (1978). Plurals in child speech. *Journal of Child Language, 5*, 237–250.
Pfeffer, J. (1964). *Grunddeutsch: Basic (spoken) German word list: Grundstufe*. Engelwood Cliffs, NJ: Prentice-Hall.
Pinker, S. (1991). Rules of language. *Science, 253*, 530–535.
Pinker, S., & Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S.D. Lima, R.L. Corrigan, & G.K. Iverson (Eds), *The reality of linguistic rules*, pp. 353–388. Amsterdam: John Benjamins.
Plunkett, K., & Marchman, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition, 38*, 43–102.
Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes, 8*, 1–56.
Rumelhart, D., & McClelland, J. (1986). On learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In J. McClelland, D. Rumelhart, & the PDP Research Group (Eds), *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.
Ruoff, A. (1990). *Häufigkeitswörterbuch gesprochener Sprache*. Tübingen: Niemeyer.
Stemberger, J.P., & MacWhinney, B. (1988). Are inflected forms stored in the lexicon? In M. Hammond & M. Noonan (Eds), *Theoretical morphology*, pp. 101–116. San Diego, CA: Academic Press.
Vennemann, T. (1947). Words and syllables in natural generative phonology. In *Natural Phonology Parasession*, pp. 346–374. Chicago, IL: Chicago Linguistics Society.
Walsh, T., & Parker, F. (1993). The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics, 11*, 201–206.
Wang, H.S., & Derwing, B.L. (1994). Some vowel schemas in three English morphological classes: Experimental evidence. In M.Y. Chen & O.C.L. Tzeng (Eds), *In honor of Professor William S.-Y. Wang: Interdisciplinary studies on language and language change*, pp. 561–575. Taipei: Pyramid Press
Zager, D. (1980). *A real-time process model of morphological change*. PhD thesis, State University of New York at Buffalo, NY.