
Natural morphology: the organization of paradigms and language acquisition*

Joan L. Bybee
University of New Mexico

1. Choosing a Word-Based Model

For centuries the traditional vehicle for teaching and learning inflectional morphology has been the paradigm laid out on the Latin model, with verb forms listed by tense, aspect or mood and nouns grouped into singulars and plurals. An example of a Spanish verbal paradigm arranged in the traditional way is given in (1).

(1) Spanish *cantar* "to sing"

IMPERFECTIVE			
Present			
Indicative		Subjunctive	
<i>canto</i>	<i>cantamos</i>	<i>cante</i>	<i>cantemos</i>
<i>cantas</i>	<i>cantáis</i>	<i>cantes</i>	<i>cantéis</i>
<i>canta</i>	<i>cantan</i>	<i>cante</i>	<i>canten</i>
Past			
Indicative		Subjective	
<i>cantaba</i>	<i>cantabamos</i>	<i>cantara</i>	<i>cantáramos</i>
<i>cantabas</i>	<i>cantabáis</i>	<i>cantaras</i>	<i>cantaráis</i>
<i>cantaba</i>	<i>cantaban</i>	<i>cantara</i>	<i>cantaran</i>
PERFECTIVE			
Indicative			
<i>canté</i>	<i>cantamos</i>		
<i>cantaste</i>	<i>cantasteis</i>		
<i>cantó</i>	<i>cantaron</i>		

We memorize the regular paradigm and then we know how to inflect any regular form; we memorize the irregular paradigms precisely because they are irregular. In this traditional model, mnemonic devices take one word of

a paradigm and form others by modifying it. For instance, the following rule of Spanish inflection works for all but a few highly irregular verbs.

To form the Present Subjunctive, take the 1st Sg. of the Present Indicative, and replace the *o* with *e* in First Conjugation, and with *a* in Second and Third. Add the person / number suffixes.

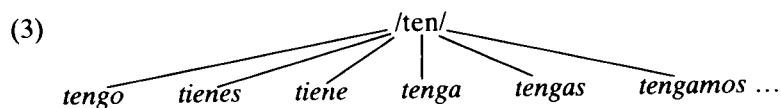
The verb *tener* 'to have' is an example of a case where this rule is useful:

(2)	<i>Present Indicative</i>		<i>Present Subjunctive</i>
	<i>tengo</i>	<i>tenemos</i>	<i>tenga</i> <i>tengamos</i>
	<i>tienes</i>	<i>tenéis</i>	<i>tengas</i> <i>tengáis</i>
	<i>tiene</i>	<i>tienen</i>	<i>tenga</i> <i>tengan</i>

In this model, morphological patterns emerge through the comparison of words within and across paradigms. That is, in this view, both the **word** and the **paradigm** to which it belongs are a part of the grammar.

In contrast, linguistic theory as practiced in this country since the 1920's has not considered paradigmatic organization to be a part of the grammar. Rather paradigms, and usually words as well, are regarded as derived by the grammar — they are artifacts of the grammar and not directly represented in any way. This view is the inevitable consequence of taking the morpheme as the basic unit of lexical representation, and treating word-information as the concatenation of morphemes.

To see why this is so, consider the usual generative treatment of words in a paradigm. Each word is derived separately from an underlying form, and inflectionally related forms are not directly related to one another, but rather each one is related to the underlying representation of the stem morpheme, as shown by the diagram in (3).



The problem with this model and variations on it have been pointed out many times (see for example, Matthews 1974). In fact, one could view most of the work in theoretical morphophonology over the last 60 years as directed toward trying to solve the problems that this model creates. I would suggest that most of these problems result from the choice of the morpheme as the basic unit. The difficulty with the morpheme is the following: while the morpheme has both **meaning** and **form**, morphemes are identified on the basis of constant meaning only. The fact is that morphological

form varies widely — both for a single morpheme and across morphemes. Thus to use the morpheme as the basic combinatory unit may work for the derivation of meaning, but to use this unit to derive **form** creates numerous problems, many of which have been studied intensely.¹

While considerable attention has been given to the acquisition of grammatical morphemes both in a first and second language, no one would seriously propose that the major task of language acquisition is to learn the morphemes of a language one by one. Rather it is clear that morphemes are acquired as parts of words and phrases, and not in isolation (MacWhinney 1978; Peters 1983). It is also clear that a large part of the task of acquiring grammatical morphemes consists of discovering their appropriate contexts of use, and this includes knowing words they may be a part of.

Thus because of problems in morphological segmentation and the way acquisition proceeds, the model proposed here will not be a morpheme-based model, but will rather take the word as the basic grammatical unit.² Further advantages for taking the word as basic will become apparent as we proceed, but for now, note one more argument for this proposal: the word is the unit in which form and meaning best correlate, which is to say that the word is the smallest unit which is complete both semantically and phonologically.

It would be unfair to give the impression that the word as a grammatical unit presents no problems of its own. In fact, there are segmentation problems associated with words just as there are with morphemes. Thus it is not always clear how much material is included in a word: contracted elements such as the English auxiliaries, and clitics such as the Spanish object pronoun clitics present problems. However, problems of segmentation are fewer at the level of the word than at the level of the morpheme.

The second problem usually raised in discussions of a word-based model is the problem of redundancy. If every word in a language has a lexical entry, then in highly inflected languages, tens of thousands of words, most of which are minimally different from other words of the same paradigm, and furthermore predictable by rule, will be listed in the lexicon. A third related problem is the potential loss of significant generalizations, such as those governing regular word-formation, which, according to standard belief, should have rule-like representation somewhere in the grammar.

The answer to these complaints has two parts. First, there is no reason to believe that neural organizations results in non-redundant storage of

information, nor that the brain's capacity would be strained by storage of all the words of a language or even several languages. Secondly, lexical storage should not be thought of as list-like in nature. Rather I would argue that we must conceive of the lexicon not as a dictionary which is limited to two dimensions, and in which every entry is approximately equal to every other entry, but rather as a highly structured and organized network in which entries are interconnected and form relationships of various sorts. Moreover, the mental lexicon differs from a dictionary in that it is not static and unchanging, but dynamic. Not all words have the same status: as words are used the strength of their lexical representation increases, as they go unused, their representations fade. Not all words have to be listed in a lexicon: morphologically complex words derivable from other words are not necessarily stored, particularly if they are infrequent.

In the following I will outline a model of the acquisition of morphology, based on Bybee (1985) and Bybee (1988), that focusses on the relations among words of a paradigm, relations which have been discovered by studying cross-linguistic, historical, experimental and acquisition data. This approach to the study of the acquisition of morphology differs from previous studies, especially of second language acquisition, which have concentrated on the order of acquisition of the grammatical morphemes of English (Dulay and Burt 1974; Bailey, Madden and Krashen 1974; Larsen-Freeman 1976) in that it is a broader empirical base: the data that led to its formulation come from languages with much richer morphological systems than English, and are not limited to acquisition data, but include other types as well.

This approach also differs from others in that it considers not just the **order** in which morphemes are acquired, but also the **way** in which they are acquired. Grammatical morphology, especially when it is inflectional (or bound), concerns relations among **words**: what has to be acquired are many different words and the ability to make more words. Among the factors that Brown (1973) considered in trying to explain the order of acquisition of English morphemes among children was the "grammatical complexity" of the morphemes: that is, he considered the syntagmatic relations the morphemes enter into. However, he did not consider the paradigmatic relations the grammatical morphemes create. In fact, paradigmatic relations have been largely ignored in the recent literature on second language acquisition, except for the observation by Van Patten (1984) that the morphemes that modify nouns should be considered separately from those that modify

verbs. Van Patten shows that when nominal morphology is separated from verbal morphology, the order of acquisition for first and second language learners, both children and adults, is virtually identical, reinforcing the idea that grammatical morphemes cannot be studied independently of the words they are part of.³

In arguing for this model, I will be arguing that both the cognitive strategies of the learner and the language input shape the learning process: in fact, the learner and the language input interact in such a way that neither of them is the same afterwards. The acquisition process and the use of language modify the language and determine its structure. That is why if we study cross-linguistic generalizations together with the acquisition process, historical change, psycho-linguistic experimentation, we have not only a theory of morphology but we also have a theory of the acquisition of morphology.

2. The Basic / Derived Relation

Very often among inflectionally related forms there is one that is distinguished by having **no** affixes even though it conveys an inflectional meaning. Since the meaning is present, a morpheme is thought to be present, but its formal representation contains no phonological matter and is thus considered to be a zero morpheme. It is a well-documented fact that the distribution of zeroes is not random but tends to be concentrated in certain members of categories, such as those listed in (4):

(4)	<i>unmarked</i>	<i>marked</i>
	singular	plural, dual, trial
	nominative	other cases
	1st and 3rd	other persons
	present	other tenses
	indicative	other moods
	cardinal	ordinal numbers

It is also well-known that the zero-marked forms are the semantically unmarked or basic members of their categories (in the sense of Jakobson 1957) and as Greenberg (1966) has shown, they are also the most frequently occurring inflected forms in texts.

As a consequence of the tendency for zero-marking to occur in the unmarked members of categories, in many cases a paradigm may be viewed

as constructed with the unmarked form serving as the base, while the other forms are constructed by adding to or modifying this base form. Consider again the verbal paradigm in (1): the 3rd Singular form *canta* has no tense, mood or person/number marker. Many of the other forms of the paradigm may be derived by adding a suffix to this form, for example, 2nd Singular *cantas* adds an *s*, 1st Plural *mos*, 3rd Plural adds *n*, and so on. For this reason, I refer to the relationship between the base or unmarked form and the other forms as the basic / derived relation.

Each paradigm does not necessarily have just one base form. Rather it appears to be the case that the domain of the basic / derived relation may be subgroups of forms within a paradigm. We will return to the determination of the boundaries of these subgroups in the next section, but for now note that in the Spanish paradigm, the 1st and 3rd Sg forms of Present Subjunctive may serve as the base for the other forms of the Present Subjunctive.

One may also note that there are parts of the paradigm where the relation I have been describing is not transparent, as for example in the Preterite Indicative, where the 1st and 3rd Sg have different stressed suffixed vowels than the other forms. Thus a transparent basic / derived relation is not a constant property of inflectional paradigms. The reason for this is that there are various factors involved in the historical evolution of paradigms that to some extent act against this principle, for example, phonological change. However, I would claim that whenever restructuring of the paradigm occurs, the nature and directionality of such restructuring is predicted by the basic / derived principle.⁴

This predictability is evident in the leveling of alternations. Consider a very straightforward example: the English verbs, *weep*, *leap*, *creep* and so on, have a vowel alternation in the Past Tense, e.g., *wept*, *leapt*, *crept*, and so on. In these verbs, there is a tendency for the alternation to be eliminated (although it is retained in *keep*, *sleep*, *leave*, etc.). To eliminate the alternation, either vowel could be chosen, that is, we could end up with a base form *wep* and a Past *wept*, or with a base *weep* and a Past *weeped*. Of course, it is the latter alternative which is realized, and the reason for this is that this type of leveling almost always favors the alternant found in the base form. This strong tendency has been documented for a large number of Romance and Germanic cases by Mańczak (1980), who formulates his principle in terms of the more frequent member of the category, claiming that this member has a greater tendency to remain unchanged or to serve as the basis of change in other members of the category.

It is important to outline clearly what is going on when leveling occurs, that is, when *weeped* begins to replace *wept*. It is not the case that the vowel of *wept* is changed, /e/ > /iy/. In fact, *wept* remains completely untouched. Rather, what happens is that a new Past Tense is constructed, working from the base form and adding the suffix /t/ to it. Then, of course, the two forms compete, both existing in the language, as *weeped* and *wept* do today.

The higher frequency of the base form has a lot to do with its role in leveling. Because the unmarked form is more frequent, it is more available in the input and has a stronger representation in the mental lexicon. Thus, if the Past form *wept* is not learned or not immediately accessible because of its lower frequency, the stronger form *weep* may be used to produce a regular formation.

In Bybee and Brewer (1980) and Bybee (1985), we presented historical and cross-linguistic evidence that the basic / derived relation is an important principle of morphological organization, and made only passing reference to acquisition evidence. Recent work in first language acquisition of a variety of languages (Slobin 1985) provides excellent evidence for this principle, as we shall see in the following paragraphs. Consequently, we would also expect that this principle applies to second language acquisition.

In the earliest stages a child acquires individual words apparently without being aware that they belong to paradigms. The first form acquired is usually the basic form — the nominative singular of nouns or a 1st or 3rd singular present indicative for verbs. At first the child uses this one form in place of the other forms of the paradigm. So a child acquiring Spanish will use *canta* for all persons and tenses, and a child acquiring English will use the form *sleep* for all tenses. Then when the child begins to acquire the functions that go with the other forms, he or she builds the other forms using the first one as the base. This is evident in the errors the child makes. Consider the following examples:

In Hebrew the regular feminine plural suffix is *-ot*, and as can be seen in the examples in (5), some nouns take this suffix in place of singular *-a* with no other changes, while others require a change in the vowel pattern and sometimes the consonant as well (Berman 1985):

(5)	Singular	Plural	Child's plurals	
	<i>tikra</i>	<i>tikrot</i>	<i>tikrot</i>	'ceiling'
	<i>simla</i>	<i>smalot</i>	<i>simlot</i>	'dress'
	<i>ricpa</i>	<i>rcafot</i>	<i>ricpot</i>	'floor'

Berman reports that children learn the *-ot* suffix and apply it across the board to feminine nouns, and that younger children simply add the suffix to the singular noun, replacing *-a*, but without making the changes in the stem vowels and consonants. A similar phenomenon may be observed in the acquisition of masculine nouns.

In verbs, the singular and plural are in a similar relationship. German has some stem alternations in auxiliary verbs between the Third Singular and Plural, as shown in (6). The child reported on in Mills (1985) produced Third Plural forms that used the Singular forms as a base:

(6)	3rd Singular	3rd Plural	Child's 3rd Plural	gloss
	<i>kann</i>	<i>können</i>	<i>kann'n</i>	'be able'
	<i>ist</i>	<i>sind</i>	<i>is'n</i>	'be'
	<i>will</i>	<i>wollen</i>	<i>will'n</i>	'want'

This example is especially instructive because in the case of *können* and *wollen*, the Third Plural form is the same as the infinitive. If one supposes, as traditional grammar does, that the infinitive is the basic form of the paradigm, then it would be difficult to explain these formations. If one makes the assumption, which I think is more reasonable, that a finite form is the basis of a verbal paradigm, in particular a singular form of the present, then these formations are predictable.

Where there is differentiation of 1st, 2nd and 3rd person in verb forms, evidence from historical change and child language point to the 3rd Sg form as basic, although there are also cases where it appears that 1st Sg, or the Imperative form plays this role (Bybee and Brewer 1980).

In languages where the verb agrees with the subject, there are reports of children in early stages using the 1st Sg pronoun with a 3rd Sg verb form, e.g. in these Brazilian Portuguese examples where the 1st Sg pronoun *eu* is paired with the 2nd/3rd Sg verb forms, Present Tense *acha* and Preterite *tirou* (Simões and Stoel-Gammon 1979):

- (7) *Eu acha o rabo dele.* 'I finds his tail.'
Eu tirou! Eu tirou! 'I took (it) off!'

A further stage of development finds the child constructing a 1st Sg form using the 3rd Sg as a base, that is, simply adding the 1st Sg affix to the stem used in the 3rd Sg. Examples of this phenomenon are found in Brazilian Portuguese (Simões and Stoel-Gammon 1979), and also in Polish, where Smoczyńska (1985) reports that a child confronted with the three

conjugation classes of Polish verbs, as shown in (8), consistently produced 1st Sg forms that were made up of the 3rd Sg plus a suffix *-m*. Note that this strategy also led to the leveling of stem alternations in favor of the alternant in the 3rd Sg as well as the introduction of a variety of vowels preceding the *-m*.

(8) Polish (Smoczyńska 1985)

Present forms in the Singular

	Conjugation 1		Conjugation 2		Conjugation 3
	'write'	'take'	'do'	'see'	'read'
1st	<i>pisz-ę</i>	<i>bior-ę</i>	<i>rob-i-ę</i>	<i>widz-ę</i>	<i>czyt-a-m</i>
2nd	<i>pisz-e-sz</i>	<i>bierz-e-sz</i>	<i>rob-i-sz</i>	<i>widz'-i-sz</i>	<i>czyt-a-sz</i>
3rd	<i>pisz-e</i>	<i>bierz-e</i>	<i>rob-i</i>	<i>widz'-i</i>	<i>czyt-a</i>

Child's 1st Singular forms

1st	<i>piszem</i>	<i>bierzem</i>	<i>robim</i>	<i>widz'im</i>	<i>czytam</i>
-----	---------------	----------------	--------------	----------------	---------------

In addition to these cases, there are some reports of 1st Sg being substituted for the 3rd Sg, for example, in Spanish (Clark 1985). This is possibly the explanation for a case in German, where the 3rd Sg form does not follow the adult model, but appears to be constructed out of another form of the stem, possibly the 1st Sg form. In this case, an unlauded vowel appears in the 2nd and 3rd Sg forms only, all others have a back vowel or diphthong:

- (9) Present Indicative Child's form 'run'
- | | | |
|----|---------------|--------------|
| 1s | <i>laufe</i> | |
| 2s | <i>läufst</i> | |
| 3s | <i>läuft</i> | <i>lauft</i> |

The child produced *lauft* for the 3rd Sg form. A standard interpretation for this form is that the infinitive is being used as the base. We have already seen a case above where this interpretation is not possible. Another possibility is that the type frequency of the back diphthong vs. the unlauded one in the input leads to the selection of back diphthong over the other. If so, this is a different principle and how it interacts with the basic-derived principle would have to be determined.

Finally, there are also some reports of an Imperative form substituting for others, e.g., in Japanese (Clancy 1985). Schieffelin (1985) reports that one Kaluli child she studied produced 1st Sg Present forms on the basis of a Present Imperative stem:

(10) Adult forms:

Present Imperative	1st Sg Present
<i>daguma</i>	<i>dudɔl</i> 'peel/shell'
<i>gadama</i>	<i>gidɔl</i> 'put together'
	Child's 1st Sg Present
	<i>dagulo</i>
	<i>gadolo</i>

At least two factors determine the selection of basic forms. One is the semantic criterion of markedness or semantic simplicity, and the other is frequency of use. Since the two factors coincide in most cases, it is difficult to determine if one or the other is primary. Important evidence for the role of frequency is the phenomenon of "local markedness" discussed by Tiersma (1982). Tiersma discusses a number of interesting examples in which the semantically marked form serves as the basis of morphophonological regularization. For example, vowel alternations between singular / plural pairs in Frisian usually regularize with the use of the vowel of the singular for both forms, but a small set of nouns shows the opposite directionality — the vowel of the plural comes to be used in the singular. However, as Tiersma argues, these are all nouns in which the plural is more frequent than the singular, nouns that refer to objects that ordinarily appear in pairs or groups, e.g. *arm*, *goose*, *horn* (of an animal), *stocking*, *tooth*, *splinter*, *thorn*, *tear*. For these nouns the plural is stronger than the singular. (See also Bybee 1985: 74-77). Other cases of local markedness pointed out by Mańczak (1980) are the locative case of place names and the instrumental case of nouns designating tools.

Local markedness phenomena have also been reported in the child language literature. Berman (1985) reports for Hebrew the formation of singular nouns from plural ones in the case of the nouns for 'sea-shells', 'tears' and 'bones'. As shown in (13), the child's singular form uses the stem of the plural.

(11) Adult forms	Child's singular		
singular	plural		
<i>cédef</i>	<i>cdafim</i>	'sea-shell'	<i>cdaf</i>
<i>dim'a</i>	<i>dma'ot</i>	'tear'	<i>dma'a</i>
<i>écem</i>	<i>acamot</i>	'bones'	<i>acama</i>

Schieffelin (1985) also observes that in Kaluli the verb 'give' *dimina* occurs almost exclusively in the Present Imperative in children's speech,

and almost never in the 1st person or in the Negative Imperative. In contrast, the verb 'take' *dima* occurs in the 1st person Present and in the Negative Imperative and never in the Imperative or Future.

Local markedness, then, provides the clue to the processes behind the more usual cases where the semantically unmarked form serves as the base. Local markedness suggests that as speakers our understanding of morphological relations has a great deal to do with the way that related forms occur naturally in discourse. It suggests that the notion of a basic form is not a strictly grammatical notion, but rather one based on how we perceive the world and our social relations and how we want to talk about them. Two distinct elements of meaning can be combined if they occur together often enough. Thus for most nouns "singular" is an inherent part of the meaning of the noun. Consider the concept evoked with the word 'dog' or 'book'. The concept is singular, that is, it contains only one exemplar of the category. That is why for these words singularity is an inherent part of the meaning. Their monomorphemic status (due to the zero mark for singular) is appropriate — they are not conceived of as morphologically complex. Similarly, entities that more naturally occur in numbers greater than one may have the concept of plurality as part of their inherent meaning. That is why words like *dice* and *lice* for many English speakers appear to have no singular form: plurality is an inherent part of their meaning, not an additive part, as in *dogs* or *books*. If it is true that the most commonly occurring concepts or categories tend to be treated as unitary or non-complex in our naming conventions, then, the notions of "semantically unmarked" and "relatively frequent" are one and the same notion.

In the most usual case, then, the base form represents a simple, unitary concept. The other forms of a paradigm semantically contain this form in addition to extra grammatical meaning. It is for this reason that speakers treat the non-basic forms as though they consisted of the form for the base plus added elements. In cases of local markedness, a form that is usually not "basic" is strongly represented and acts as though it were basic.

Perhaps the most fascinating consequence of local markedness is that not all paradigms in a language have the same organization, even in cases where they appear to involve the same formal properties. Thus the Hebrew singular / plural pairs mentioned earlier might yield two different sorts of relations, as shown in (12), where the word written in larger print is the stronger word of the pair, and thus the more basic word. The lines drawn between the words are connections representing the features shared by the

two words. In this case these relations are morphological relations, since they consist of parallel semantic and phonological connections. (It is also possible to have connections that are only semantic or only phonological.)



Similarly, the English pairs *mouse*, *mice* and *louse*, *lice* have different relations, as shown in (13):



The learning strategy involved in constructing a paradigm, then, is the principle that new input is analyzed and stored in terms of existing structures. The first form of a paradigm to be acquired will be the most frequent and the semantically most coherent. Once this form is represented other less frequent and more complex forms may be analyzed and stored by reference to the existing form. While this principle has been illustrated with examples from first language acquisition, it is reasonable to expect that such a general learning strategy would also apply to second language acquisition.

3. Lexical Strength

In the explanation of the basic / derived relation, I have argued that the token frequency of individual words plays an important role, such that words of high frequency have a greater **lexical strength** than low frequency words. It is not just the basic / derived relation that provides evidence for lexical strength, but two other phenomena as well. One is the fact that frequent words are accessed faster in lexical access experiments. That is, when subjects are asked to say if a string of letters or sounds represents a word of English, they respond faster if it is a frequent word. The second is that morphological irregularity most commonly occurs in the most frequent paradigms of the language. The reason for this is that the leveling of these irregularities is resisted by those forms that are more frequent and have stronger, more accessible mental representation. The stronger representa-

tion of high frequency words is also apparent in the fact that both L1 and L2 learners of English acquire the high frequency irregular Past Tense verb forms before the regular Past forms. Note that what is important here is the token frequency of individual words, not the frequency of grammatical morphemes (which is what Brown (1973) and Larsen-Freeman (1976) counted). I will return to a discussion of the role that morpheme frequency (or type frequency) plays in acquisition in section 6.

4. Degree of Relatedness

Each word of a paradigm is either a base form, or is connected to a base form by connection lines such as those shown in (12) and (13). However, since there may be two or more base forms for a particular paradigm, the question arises as to which forms are related to which base. That is, what is the domain of the basic / derived relation? In this section, I will present evidence for a semantically-based hierarchical organization of paradigmatically related forms, which yields an arrangement of forms into closely related clusters which have a common base form.

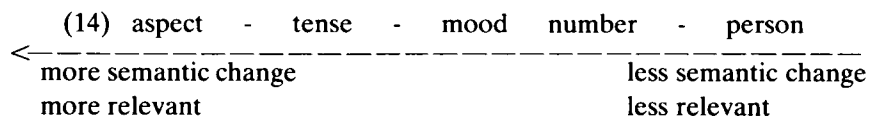
If we consider just words related in inflectional paradigms, we can identify varying degrees of semantic relatedness depending on the meaning of the affix category. In Bybee (1985), I proposed that morphological categories (and in particular those of the verb) differ semantically in the degree of relevance that they exhibit toward the meaning of the stem that they modify. Since the semantic content of verbs deals mainly with events — actions, processes and states — the most relevant categories are those they modify the described event. They may do so by modifying the argument structure of the verb (as valence and voice do), or by modifying the temporal structure (as aspect does). The more relevant categories have a greater effect on the meaning of the verb, so that a verb modified by a highly relevant category will be less closely related to its base than a verb modified by a less relevant category.

As an example, consider the verbal category of aspect. Aspect is directly relevant to a verb and affects its meaning, since the meaning of all verbs has some inherent temporal structure, and aspect modifies the “internal temporal constituency” of the event or state described by the verb (Comrie 1976). Moreover, a change in aspect can produce quite a difference in the event described by the verb. Consider the example of the

Spanish Preterite / Imperfect distinction, which is an aspectual distinction. The verb meaning 'sleep' in the Preterite, *durmió* 's/he slept', describes a completed event, which includes falling asleep and sleeping. The Imperfect *dormía* translates approximately as 's/he was sleeping' implying a state s/he was in when something else occurred. For some verbs this aspectual distinction is large enough to produce distinctions that are expressed by separate verbs in another language. For instance the Preterite of the Spanish verb *saber* 'to know' translates into English as 'found out'.

Now compare aspect to person agreement. The function of agreement is to index the participants in the state or event described by the verb and has nothing to do with the inherent meaning of the verb. Two verb forms that differ only by person are much the same semantically if their tense, aspect and mood are the same. The semantic differences produced by person markers are not the type that would be expressed by entirely different verbs. Thus such forms are more closely related than forms that differ in aspect.

Using this same kind of reasoning we can order aspect, tense, mood, number and person as shown in (14) and say that two verb forms that differ only by person are the most closely related; two forms that differ by aspect are the least closely related. (See Bybee 1985, Chapter 2 for further motivation of this hierarchy.)



If we organize paradigmatically related forms on the page according to the degree of their relatedness, the arrangement is similar to that of a traditional paradigm. In fact, traditional paradigms capture this notion of degree of relatedness to a large extent. For instance, we never see paradigms arranged by person/number forms, that is, by listing all the 1st Sg forms together, then all 2nd Sg forms, and so on. The difference between the current model and the traditional one is that in the current one all forms are not equally represented, since some are stronger than others, and in addition, the citation form is a finite, rather than a non-finite form.

It is a significant cross-linguistic fact that the degree of semantic relatedness is paralleled by morphophonemic alternations: the more closely related two forms are semantically, the more likely they are to be similar morpho-phonemically. This means, for example, that stem changes in verbs

are more likely to distinguish aspects than to distinguish person forms especially across aspects or tenses. Thus in Spanish there is a set of irregular verbs that have stem changes for the Preterite aspect. The verb *saber* 'to know' has the 3s Preterite form *supo*, as shown (15); similarly *tener* 'to have' has the 3s Preterite form *tuvo*; *querer* 'to want' has the 3s Preterite form *quiso*, and so on.

(15) Spanish 'to know'

Present		Imperfect		Preterite	
<i>sé</i>	<i>sabemos</i>	<i>sabía</i>	<i>sabíamos</i>	<i>supe</i>	<i>supimos</i>
<i>sabes</i>	<i>sabéis</i>	<i>sabías</i>	<i>sabíais</i>	<i>supiste</i>	<i>supisteis</i>
<i>sabe</i>	<i>saben</i>	<i>sabía</i>	<i>sabían</i>	<i>supo</i>	<i>supieron</i>

These irregular stems occur throughout the Preterite person forms, and thus set off the Preterite (perfective aspect) from the Present and Imperfect (both imperfective aspect). On the other hand, there are no stem changes in Spanish that set off, for example, all 1st person forms in all aspects and tenses from all other person forms. Indeed, the hypothesis is that such a situation would be very rare. This hypothesis has been tested on a sample of 50 unrelated languages, and was not disconfirmed. On the contrary, it was found that stem alternations of consonants or vowels are extremely frequent where aspectual distinctions are concerned and extremely rare where person distinctions are concerned (Hooper 1979; Bybee 1985).⁵

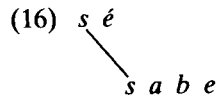
Thus I would argue that the degree of relatedness among words is primarily determined by the number and type of semantic features shared. The degree of phonological similarity will often parallel the degree of semantic relatedness.

The hierarchy in (14) arranges categories according to how important they are semantically **to the verb**. It makes certain predictions concerning the order of acquisition of these categories as verbal markers. To a large extent, these predictions are upheld by the evidence from first language acquisition, and we would thus expect them to also be upheld by data from second language acquisition. For instance, the examples cited earlier from Brazilian Portuguese show the child using the distinction between Present and Preterite, which is an aspectual distinction (*acha* 'find' vs. *tirou* 'took'), but not yet marking the distinction between 1st and 3rd person on the verb (although it is present in the pronouns). It is also argued for several languages that aspect is acquired before tense (Antinucci and Miller 1976; Weist et al. 1984).

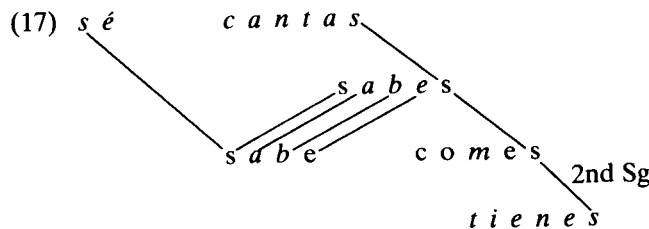
5. Illustration

In this section I would like to be more specific about how paradigmatic structures are built up in the acquisition process: The illustration will be in terms of the verb 'to know' in Spanish as given in (15). The sequence of developments is purely speculative and not based on any hard data about the order of development of forms of this verb. Note in particular that I think it is more plausible for this particular high frequency stative verb that person distinctions in the Present are acquired before the Preterite (another case of local markedness).

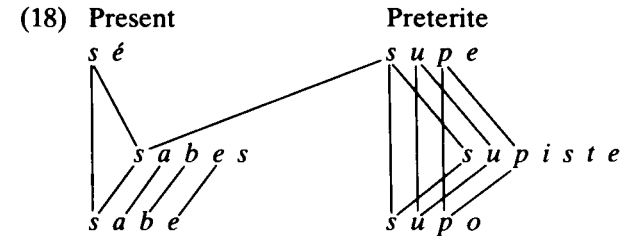
1. First, the most frequent individual words are acquired. When a word is acquired of course, its representation contains both semantic and phonological material. *Sé* 'I know' and *sabe* 's/he knows' are acquired early, and the two are related: they have a strong relation semantically, but their phonological relation is minimal. They are both very frequent and can have individual, though related representations. The line drawn between them represents a morphological relation; that is, shared semantic features paralleled by shared phonological features.



2. The form *sabes* enters the picture; it maps both semantically and phonologically onto *sabe*, with only an *-s* left over. When other such pairs are acquired, the *-s*'s of 2nd singular are mapped onto one another.



3. The Preterite forms *supo* and *supo* come in; the semantic and phonological match with *sabe* is imperfect. A new cluster of forms is established but with connections to the *sabe* group, due again to shared semantic and phonological features.



It should be noted that segmentation into morphemes is a consequence of the mapping function. Words do not have to be literally pulled apart when their relations with other words are established, rather the matching parts are mapped onto one another. If there are only partial similarities (as with *sabe* and *supo* or other cases of allomorphy) or left over bits (such as empty morphs), it does not interfere with the mapping of the parts that do match. Thus the model identifies morphemes in the same way that real speakers and even linguists do — by comparing words that have similar semantics and similar phonological shape.

The notion of lexical strength that I mentioned earlier means that some words are more strongly represented than others. Infrequent words that are derivable from other words are not necessarily directly represented. For instance, the 2nd Singular Present Indicative of a Spanish verb is quite predictable: it consists of 3 Sg form plus a suffix *-s*. For frequent verbs, the 2 Sg form may be present but for infrequent verbs it may not be. However, the base is present and the *-s* which has the meaning 2nd Sg is present, as we shall see in section 6.

6. Interparadigmatic Relations

So far in our discussion we have dealt primarily with relations in the same paradigm — that is, relations among words that share the same lexical stem. However, in section 5, the *-s* of 2nd Sg was discussed in terms of its occurrence on verbs of different paradigms. Inflectional affixes are acquired as parts of whole words, and the learner begins to form a notion that a meaningful word ending exists as s/he experiences it on a variety of different stems. Connection lines develop among the various occurrences of 2nd Sg *-s* because of the shared phonological and semantic features. Since all verbs use *-s* for 2nd Sg and since it occurs in all tenses and moods, the *-s* to 2nd Sg association is very strong.

Some models of morphology would assign rule status to the relation of 2nd Sg and the suffix *-s* to the verb stem. However, in the current dynamic lexical model, this same phenomenon is treated as a very frequent and general lexical pattern, rather than as a separable morpheme or rule. I offer the following justification for this proposal:

The expression of morphological categories is highly varied. There are affixes, zeroes, stem changes and reduplication. Moreover, some specific means of expressing morphological categories are very general (in that they apply to all or almost all of a lexical category, such as verbs) and regular, like the Spanish 2nd Sg suffix we have been discussing, while others are less general and less regular. For example, the means of expressing the Past Tense in English is a fairly regular suffix, *-ed*, but there are about two hundred English verbs that use a stem change instead of or in addition to the suffix, making it somewhat less regular and general than the Spanish 2nd Sg suffix. Among these irregular verbs, there are some classes that have very similar Past forms and exhibit some degree of productivity (in the sense that the class may gain new members). An example is the class that includes the Past forms *strung*, *stuck*, *dug* and so on (Bybee and Slobin 1982; Bybee and Moder 1983). Then there are some smaller, completely unproductive former classes, such as the one including the Past Tenses *brought*, *taught*, and *sought*. The end point of the irregularity scale are cases of total suppletion, in which the Past form is not even historically related to the base form, e.g. *go* and *went*.

Most theories agree that the last case, *go* and *went*, must be treated in the lexicon, and cannot be described by a rule. (A rule that covers only one item is only a rule in a trivial sense.) Similarly, despite the fact that there are several verbs such as *brought*, there is no evidence that speakers unite them as a class in any way, and the "rules" that would treat them are so specific and complex, that a simple lexical listing of both base and Past forms is not controversial.

On the other hand, a somewhat larger and semi-productive class such as the *strung* class in (19) is more problematic. Even though speakers can assign nonce words to this class, this productivity cannot be accounted for by a rule that changes features of the base form to derive the Past form. In Bybee and Slobin (1982) and Bybee and Moder (1983), we argue that the generalization is over the Past forms, not the base forms, because the base forms are not uniform in vowel quality while the Past forms are; e.g., many of the verbs have the vowel [I] in the base, but [ay] occurs in *strike*, [æ] in

hang, and [iy] in *sneak*. Further, in the nonce probe task reported on in Bybee and Moder (1983), we found some subjects who tried to put nonce words into this class even though the input base form did not exactly fit in its consonant structure. The result was that the subjects changed the consonants, too, producing for example, the Past of *spriv* as *sprung* or *sprug*. This suggests that the Past tense forms themselves contribute to a pattern or schema associated with the meaning 'past tense' and that in forming a new Past, speakers modify a form to fit the schema. If the generalization is a pattern of similarity among Past forms, then that means they must be represented in the mental lexicon.

(19) An English Strong Verb Class

/n/	<i>spin</i>	<i>spun</i>	/ŋk/	<i>slink</i>	<i>slunk</i>
	<i>win</i>	<i>won</i>			
/ŋ/	<i>cling</i>	<i>clung</i>	/k/	<i>stick</i>	<i>stuck*</i>
	<i>fling</i>	<i>flung*</i>		<i>strike</i>	<i>struck*</i>
	<i>sling</i>	<i>slung*</i>		<i>sneak</i>	<i>snuck**</i>
	<i>sting</i>	<i>stung*</i>		<i>shake</i>	<i>shuck**</i>
	<i>string</i>	<i>strung*</i>			
	<i>swing</i>	<i>swung</i>	/g/	<i>dig</i>	<i>dug*</i>
	<i>wring</i>	<i>wrung</i>		<i>drag</i>	<i>drug**</i>
	<i>hang</i>	<i>hung*</i>			
	<i>bring</i>	<i>brung**</i>			

* Indicates an addition to this class since the Old English period.

** Indicates a non-standard addition to the class.

Furthermore, this class is defined by the phonological features of the stem, primarily the final consonant or cluster, but to a lesser extent the initial consonant or cluster, which often includes /s/. However, there is no strict requirement that a verb must have all of the relevant phonological features to belong to this class, or to be attracted into it. Rather, the class has a prototype or family resemblance structure, such that each member shares some features with a prototype, but more marginal members will not necessarily share features with one another. That is why both *won* and *dug* may belong to this class although one ends in an alveolar nasal and the other in a non-nasal velar. The prototype contains the feature nasal and velar, and each of the verbs has at least one of these features. The non-categorical nature of the generalization suggests a mechanism other than a generative rule. What it suggests is a pattern of lexical connections associating both the phonological and semantic features of the Past forms.

In Bybee and Slobin (1982) we assumed that the more productive suffixation of *-ed* to form the Past tense in English was best handled by a rule that applied in a morphological component, and that schemas were different from rules. This would mean that two different mechanisms are involved in the acquisition and use of the Past forms of different verbs. Such an assumption leads to a number of mechanical problems in getting the model to work, the chief among them being how to block the suffixation for a verb whose Past form is listed in the lexicon. Another alternative, and the one that I now favor, is that the difference between phenomena handled by schemas and phenomena handled by rules is mainly a quantitative difference: because *-ed* applies to so many verbs it is extremely powerful, which makes it appear to be categorical. However, it is not categorical: people still make mistakes with it, spontaneously and in experiments, in some cases leaving the suffix off, in others making novel vowel changes instead of using suffixation.⁶

Until recently, productivity was considered indicative of rule-governed behavior: if a new form can undergo some modification, that must mean that there is rule to handle this modification. However, MacWhinney (1978) recognized that besides derivation by "combination" (i.e. the application of regular concatenation rules), new forms could be produced by "analogy", that is, by accessing similar forms in the lexicon and producing a new form that follows the existing pattern. Thus *brung* might be produced in analogy with *strung*. But note that MacWhinney (1978), proposes a difference between "combination" and "analogy" — that is, he proposes that there are two distinct mechanisms for handling morphological patterns.

In contrast, Rumelhart and McClelland (1986) have proposed a model based on a computer architecture called "parallel distributed processing" which simulates the learning process of English Past Tense verbs, and can handle new forms in a fairly human manner, but without ever formulating an explicit rule. All the model does is recognize patterns in existing verb forms. This suggests that a unitary treatment of morphology as lexical patterning is possible and what we thought of before as rules are instead very strong and general schemas.

What I am proposing, then, is that there is no discrete cutoff point between "regular" and "irregular", between "rules" and "schemas". Rather the same learning and representational principles apply in both cases and the differences are due to the number of items involved in the pattern and its scope or generality. It is true that in English we see a rather striking dif-

ference between the "regular" and "irregular" but if this represented some basic processing difference, then we would expect other languages to manifest the same difference. However, there are many cases of morphological systems in which no one pattern represents the regular pattern. For instance, German Plural formation interacting with the gender system offers several strong patterns and not just one rule. For this reason, Köpcke (1988) argues that the situation is best handled by schemas. Similarly, Hausa Plural formation has a large number of Plural patterns, many of which involve changes to the stem that are best handled by schemas, rather than one 'regular' pattern (Haspelmath 1988).

Regular affixes, such as English Past Tense *-ed* and Spanish 2nd Sg *-s* can also be treated as schemas, that is, as patterns of lexical connections associating the phonological features at word ends with semantic features of the word. Such a schema can be stated as follows:

(20) [[verb] -s] second singular

Here "verb" represents the base word (the 3rd Sg). While the schema is just a generalization over existing words, it does represent a sound — meaning pairing that can be called into service to form new words, or to supply the needed form of an infrequent word if it is not stored.

All types of morphological patterns can be acquired by the same process — the storage of items, the creation of connections among them, and the formation of patterns that range over sets of connections. The differences among them are due largely to the number of distinct lexical items involved — a big class is more productive and forms a stronger schema than a small class. A large class has a high type frequency, that is, the number of different words containing the affix or the pattern is high. This does not mean that the affix itself has a high frequency, which is the reason that correlations of morpheme acquisition order and frequency in the input do not always work (Brown 1973). It is possible for an affix to have an extremely high frequency because it occurs with a few high frequency stems, and yet be unproductive because it does not occur in enough distinct combinations to be recognized as an affix.

Consider the conjugation classes for French verbs. The First Conjugation is the productive one for both children and adults, and this is related to the fact that it applies to the largest number of verbs. Guillaume (1927) studied the rate of occurrence of verbs of each conjugation in the speech of nursery school children talking among themselves. The following table

shows the number of occurrences of each conjugation class and the number of verbs used from each:

(21) Conjugation class	number of uses	number of verbs
First (<i>chanter</i>)	1,060 36.2%	124 76.0%
Second (<i>finir</i>)	173 6.0%	10 6.1%
Third (<i>vendre</i>)	1,706 57.8%	29 17.9%

The figures show that Third Conjugation verbs occurred more frequently in the children's speech, but only a small number of distinct verbs were from this class. The productive First Conjugation occurred less often, but a very large number of verbs were from this class. With First Conjugation affixes occurring on a large number of different verbs, there is more chance to compare forms and extract a pattern. The Third Conjugation verbs, which are very frequent, could be learned and stored as independent items, each with a high lexical strength, and the pattern inherent in these verbs may not be exploited. In this way, lexical strength, (based on token frequency) and schema formation (based on type frequency) interact to influence the acquisition of morphology. I would suggest that frequency in the input is extremely important for our understanding of acquisition, but rather than counting the grammatical morphemes, we must look to word frequency and study the interaction of type and token frequency.

7. Conclusion

The model I have described here accounts for a range of cross-linguistic, historical and acquisition phenomena that have no explanation whatever in generative models or other morpheme-based accounts. These include the order in which children acquire forms of a paradigm and the strategies they use in innovating forms. The cross-linguistic facts accounted for are the distribution of zeroes, as well as the distribution of stem alternations and suppletion. The model also predicts the general direction of leveling of alternations in historical change.

In addition, the model makes certain predictions about second language acquisition:

(i) That the most frequent and semantically basic words of a paradigm will be acquired and used first; that related words of a paradigm are analyzed and reformed using the first-acquired words as a base.

(ii) That the more "relevant" distinctions (for verbs, aspect, tense) are acquired before the less relevant ones.

(iii) Affixes are acquired only after a number of words containing the affix have been acquired and compared; similarly, subclasses or less regular patterns are acquired in the same way — with the acquisition of a number of words containing the pattern followed by the construction of associations among similar words.

The learning mechanisms invoked in this model are of a very general nature and not restricted to language. They are that the most often repeated experiences (in production or perception) have the strongest representation and that new experiences are analyzed and stored in terms of existing representations. Given these two very general principles, we only need to understand the nature of the input to predict the course of language acquisition. For this reason, the point I would stress about this model is that it allows us to establish a relation between the way language is used — that is, how forms actually appear in natural discourse — and the way language is represented mentally. Not only is this important for the understanding of the acquisition and learning processes, but it also helps us understand and predict the ways in which language use molds language structure.

Notes

*I am grateful to the editors and to Fred Marshall for useful suggestions on both the content and form of this paper. The expression "natural morphology" is not intended to indicate the Natural Morphology of Dressler et al. (1987), although there are some similarities in our approaches.

- Here are a few examples of the problems that arise from taking the morpheme as the basic unit of form:
 - There are difficulties with segmentation into morphemes in cases such as the English Past Tense, where *walked* is segmentable enough, but *sang*, *brought* and *went* are problematic.
 - There are "meanings" that appear to have no form, such as the English noun singular (a "zero morpheme"); there are formal bits that do not seem to have any meaning (3rd Plural Preterite, *ro* in Spanish).
 - There are cranberry morphs — bits left over when other meaningful elements have been segmented. For more examples see Bybee (1988).
- For arguments in favor of taking the word as the basic lexical unit in generative grammar see Jackendoff (1975) and Arnoff (1976).
- In studies of first language acquisition, this point has been made by Antinucci and Miller (1976) and Bloom et al. (1980), who argue that the lexical meaning of a verb influences which aspects and tenses it will occur with first.

4. For instance, in case of the Spanish Preterite Indicative and the related forms in Provençal, Bybee and Brewer (1980) show that a whole group of "analogical" changes serve to re-establish the basic / derived relation.
5. Further evidence for a hierarchical ranking among grammatical categories is the distribution of forms in suppletive or split paradigms, such as the paradigm for *go* which has a Past form *went*, which is historically from a completely different verb stem. When inflectional paradigms split and realign, forming suppletive paradigms, the splits occur more often among forms that are **less** closely related semantically, than among forms that are **more** closely related. Rudes (1980) has studied suppletive verbal paradigms in a large number of languages, and found that splits occur in general along aspect or tense lines, as with *go* and *went*, and along person agreement lines only in the present tense, the most frequent tense.
6. Haber (1975) describes a nonce-probe experiment with English plural formation, which is even more regular than suffixation for Past Tense. Her results show that speakers do not apply Plural suffixation in a regular and uniform way.

References

- Antinucci, F. and R. Miller. 1976. "How Children Talk about What Happened." *Journal of Child Language* 3.167-189.
- Aronoff, M. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.
- Bailey, N., C. Madden and S. Krashen. 1974. "Is There a 'Natural Sequence' in Adult Second Language Learning?" *Language Learning* 24.235-243.
- Berman, R.A. 1985. "The Acquisition of Hebrew." In Slobin 1985.255-371.
- Bloom, L.M., K. Lifter and J. Hafitz. 1980. "Semantics of Verbs and the Development of Verb Inflections in Child Language." *Language* 56.386-412.
- Brown, R. 1973. *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.
- Bybee, J.L. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.
- Bybee, J.L. 1988. "Morphology as Lexical Organization." *Theoretical Morphology* ed. by M. Hammond and M. Noonan, 119-41. Orlando: Academic Press.
- Bybee, J.L. and M.A. Brewer. 1980. "Explanation in Morphophonemics: Changes in Provençal and Spanish Preterite Forms." *Lingua* 52.271-312.
- Bybee, J.L. and D.I. Slobin. 1982. "Rules and Schemas in the Development and Use of English Past Tense." *Language* 58.265-89.
- Bybee, J.L. and C.L. Moder. 1983. "Morphological Classes as Natural Categories." *Language* 59.251-70.
- Clancy, P.M. 1985. "The Acquisition of Japanese." In Slobin 1985.373-524.
- Clark, E.V. 1985. "The Acquisition of Romance, with Special Reference to French." In Slobin 1985.687-782.
- Dressler, W., W. Mayerthaler, O. Panagl, and W. Wurzel. 1987. *Leitmotifs in Natural Phonology*. Amsterdam: John Benjamins.
- Dulay, H.C. and M.K. Burt. 1974. "Natural Sequences in Child Second Language Acquisition." *Language Learning* 24.37-53.
- Greenberg, J.H. 1966. *Language Universals*. The Hague: Mouton.
- Guillaume, P. 1927. [1973]. "The Development of Formal Elements in the Child's Speech." *Studies in Child Language Development* ed. by C. Ferguson and D. Slobin, 522-541. New York: Holt, Rinehart and Winston.
- Haber, L.R. 1975. "The Muzzy Theory." *CLS* 11.240-256.
- Haspelmath, M. 1988. "Schemas in Hausa Plural Formation." *Buffalo Working Papers in Linguistics*.
- Hooper, J.B. 1979. "Child Morphology and Morpho-phonemic Change." *Linguistics* 17.21-50.
- Jackendoff, R. 1975. "Morphological and Semantic Regularities in the Lexicon." *Language* 51.639-671.
- Jakobson, R. 1957. "Shifters, Verbal Categories and the Russian Verb." *Roman Jakobson, Selected Writings II*. ed. by , 130-47. The Hague: Mouton.
- Köpcke, K.M. 1988. "Schemas in German Plural Formation." To appear in *Lingua*.
- Larsen-Freeman, D.E. 1976. "An Explanation for the Morpheme Acquisition Order of Second Language Learners." *Language Learning* 26.126-134.
- MacWhinney, B. 1978. *The Acquisition of Morphophonology*. Chicago: Monographs of the Society for Research in Child Development, vol. 43.
- Mańczak, W. 1980. "Laws of Analogy." *Historical Morphology* ed. by J. Fisiak, 283-8. The Hague: Mouton.
- Matthews, P.H. 1974. *Morphology*. Cambridge: Cambridge University Press.
- Mills, A.E. 1985. "The Acquisition of German." Slobin 1985. 141-254.
- Peters, A.M. 1983. *The Units of Language Acquisition*. Cambridge: Cambridge University Press.
- Rudes, B.A. 1980. "On the Nature of Verbal Suppletion." *Linguistics* 18.655-676.
- Rumelhart, D.E. and McClelland, J.L. 1986. "Learning the Past Tenses of English Verbs: Implicit Rules or Parallel Distributed Processing." *Mechanisms of Language Acquisition* ed. by B. MacWhinney, 1985. Hillsdale, NJ: Erlbaum.
- Schieffelin, B.B. 1985. "The Acquisition of Kaluli." In Slobin 1985.525-594.
- Simões, M.C.P. and C. Stoel-Gammon. 1979. "The Acquisition of Inflections in Portuguese: A Study of the Development of Person Markers on Verbs." *Journal of Child Language* 6.53-67.
- Slobin, D.I. (ed.) 1985. *The Cross-linguistic Study of Language Acquisition*. Hillsdale, NJ: Erlbaum.
- Smoczyńska, M. 1985. "The Acquisition of Polish." In Slobin 1985. 58.832-49.
- Van Patten, B. 1984. "Processing Strategies and Morpheme Acquisition." *Universals of Second Language Acquisition* ed. by F.R. Eckman et al., 88-98. Rowley, MA: Newbury House.
- Weist, R.M., H. Wysocka, K. Witkowska-Stadnik, E. Buczowska and E. Konieczna. 1984. "The Defective Tense Hypothesis: On the Emergence of Tense and Aspect in Child Polish." *Journal of Child Language* 11.347-374.