Lingua 52 (1980) 201–242

EXPLANATION IN MORPHOPHONEMICS: CHANGES IN PROVENÇAL AND SPANISH PRETERITE FORMS

Joan L. BYBEE and Mary Alexandra BREWER*

State University of New York at Buffalo, USA

Received May 1980

The degree of *autonomy* of a word is the extent to which a word is likely to have its own lexical representation. Autonomy is determined by semantic complexity, word frequency, and morphophonemic irregularity, such that the semantically simpler, more frequent, and more irregular words are more autonomous. In morphological systems, nonautonomous words are derived from autonomous words by rule. Dynamic data from psycholinguistics and diachrony are presented in support of this hypothesis. The diachronic discussion centers around the person forms of the preterite in Provençal and Spanish, where the third singular and the first singular are the most autonomous. There is considerable dialectal evidence that one or both of these forms can serve as the morphophonemic base(s) from which the other person forms of the preterite are derived.

Morphophonemic alternations (alternations with some morphosyntactic or lexical conditioning) would seem to be among the most arbitrary and conventional aspects of natural language. Their treatment in most theories characterizes them as synchronically arbitrary – Bloomfield (1933) and later Chomsky and Halle (1968) describe morphophonemics with historical reconstructions, while others are satisfied with long lists of alternants and statements about distribution. More recently, morphophonemic alternations have been excluded from phonology, as not being explainable by phonological principles, but belonging to some other domain (Andersen 1969; Vennemann 1972; Stampe 1973). The question now arises, what is this other domain,

* We are grateful to Henning Andersen, Edith Bavin, Suzanne Fleischman, Blair Rudes, Dan Slobin, Joe Stemberger, Peter Tiersma, Yakov Malkiel, David Zager and David Zubin for discussing this research with us at various stages of its development, and moreover, for asking many difficult questions. We alone are responsible for the way we have decided to answer these questions. This research was supported in part by a Social Science Research Council Post-Doctoral Training Fellowship to the first author.

and are morphophonemic alternations synchronically arbitrary, or are they somehow motivated by the communication process and its participants?

In this paper we take the point of view that morphophonemic alternations are not arbitrary, but rather tend to form diagrammatic relations as described in Andersen 1980. Andersen's proposal (based on some brief suggestions in Jakobson 1965) is that the relations among signantia (the phonological forms of morphemes) or the relations among variants of signantia, reflect parallel or diagram relations among signata (the semantic values of morphemes). In particular, we present evidence, as others have, that there is a basic–derived relationship among the forms of a stem paradigm that diagrams the basic–derived relationship in the semantic domain (section 1). More importantly, we investigate the question of what determines particular basic–derived relationships, and the question of why diagrams exist.

Our proposal is that words may be classified as to degree of autonomy – the extent to which a word is likely to be represented in the speaker's lexicon as a whole and independent unit (Zager 1978). The factors that determine the degree of autonomy of a word are discussed in section 4. The significance of autonomy is that highly autonomous words will be learned as whole units, even if they are (to the linguist) morphologically complex, and less autonomous words will be analyzed as derived from the autonomous words of the same paradigm. Thus autonomy is one of the determinants of the basic–derived relationship. The other determinants are the morphological categories and the hierarchical relations among those categories (section 1).

Because there may be more than one autonomous word in a stem paradigm, there may be more than one basic form for the paradigm. Furthermore, derived forms are not all related to basic forms in the same way. In sections 6–9 we discuss a series of morphophonemic changes in Provençal dialects, and a related but different series of changes in Spanish dialects, that help us understand the nature of the relations in a small area of verbal morphology – the person/number forms within a tense.

1. Diagrams

1.1.

One example of a diagram which Andersen (1979) cites is the morphophonemic reflection of the semantic basic–derived relationship. Vennemann (1972) discusses such diagrams saying, "Usually in natural languages, a

semantic derivation of secondary conceptual categories from primitive ones, tertiary from secondary ones, etc., is reflected by a parallel syntactic or morpho-phonological derivation" (p. 240). Consider Vennemann's representation of this principle.

	Primitive category	Secondary category
Semantic level	A	A + b
Level of overt manifestation	X	X + y

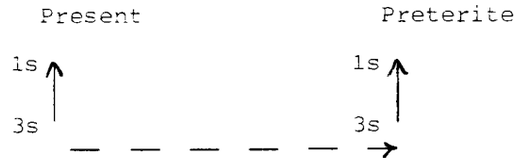
Here X is the signans corresponding to the signatum A, and y corresponds to the secondary category. The manifestation of y may be some modification of X, i.e., the application of a morphophonemic rule, or it may be the addition of an affix to X, or both.

Vennemann proposes that this diagram represents a language acquisition strategy, and that it explains the frequent observation that when morphophonemic alternations are eliminated (in so-called analogical leveling), it is the alternant found in the primitive or basic category that survives, e.g., the third person, the present, the indicative, the singular (Mańczak 1963). To take a brief example, we analyze the past tense of English *creep*, *crept*, as consisting of *creep* plus the past tense suffix, and a vowel change of [iy] to [ɛ], rather than analyzing the present as derived by a vowel change of [ɛ] to [iy]. The alternate past tense form *creeped* [kriypt] is described as arising by the elimination of the vowel change, which allows the basic form to emerge. In section 3 we will discuss more changes motivated by this type of diagram, and in section 4 we will try to explain why this type of diagram exists.

1.2.

Another diagrammatic relation is discussed in Hooper (1979), where it is shown that morphophonemic alternations in verb stems diagram the hierarchical relation in which person/number distinctions are subordinate to tense/aspect distinctions. The hypothesis was suggested by child language data in which it was evident that the children made verb stem substitutions of 3s for 1s within the same tense, while the relation of 3s stems in two

different tenses was more remote.[1] That is, the children under study, who were acquiring Brazilian Portuguese, appeared to learn the 3s preterite form separately from the 3s present form (Simões and Stoel-Gammon 1979). The arrows in the following chart represent the direction of substitution:



This pattern of substitution together with the loss of or failure to acquire morphophonemic alternations would lead to a situation in which verb stems within tenses would be more uniform than verb stems across tenses. The result is the following testable hypothesis:

> *Hypothesis*: Alternations in verb stems will correspond to tense/aspect distinctions across person distinctions, but not to person distinctions across tense/aspect lines.

For instance, there are many cases of special stem forms for present, past or future, but no reported cases of a special stem for, for example, a 1s that is used throughout the tenses. This hypothesis was tested on a stratified probability sample of fifty languages. Eighteen of these languages inflected for both person and tense/aspect, and in none of these were verb stem alternations found which corresponded to person distinctions across tense/aspect lines (Hooper 1979).

If morphophonemic alternations in verb stems coincide with tense/aspect distinctions, then they outline the major category distinctions in verbal paradigms. The greater importance of tense/aspect distinctions to verbs (over person/number distinctions) is evident in the fact that the presence of person/number inflections in a language implies the presence of inflections for tense, aspect or mode (Greenberg 1963: 112). Tense/aspect distinctions

[1] We employ the following abbreviations:

1s: first person singular	1p: first person plural
2s: second person singular	2p: second person plural
3s: third person singular	3p: third person plural

Person forms of verbs will be presented in the order given above.

are more important to verbs than person/number distinctions because of their meaning. Tense and aspect have meanings that refer directly to the event described by the verb, tense placing the action in time relation to the moment of speech, and aspect referring to the internal structure of the action. Person and number, however, refer to the arguments of the verb, and in their meaning have more to do with nominals than verbs. Thus, as regards the meaning of the *verb*, tense and aspect make larger meaning changes than person and number. The more direct relation of tense and aspect to the verb is reflected in the possibility of these categories conditioning morphophonemic modification of the verb stem itself. The peripheral status of person and number as verbal categories is reflected in their inability to affect verb stem morphophonemics across tense and aspect lines. The greater meaning change made by tense and aspect is reflected in a greater change in form. The result is that in both meaning and overt form, persons within a tense or aspect will show greater similarity and be more closely related to one another than to the same persons in another tense of aspect. In the discussion that follows we will be studying the basic–derived diagrams among the very closely related person forms within a tense.

It is important to note at this point that the morphemes whose function is to directly signal person/number form different patterns of relations than the verb stems of persons within a tense. Language acquisition data as well as historical data show many examples of person/number suffixes of one tense replacing those of another tense in the same person, as shown in the following diagram (Hooper 1979):

In our discussion we will have to ignore changes of this type.

Another type of morphophonemic change that we will not be discussing here is change in verb stems motivated from outside the paradigm. An example from Spanish is the loss of diphthongized stem variants, as in the verb *temblar* 'to shake', which has etymologically and in some dialects a diphthong when the stem is stressed, *tiémblo* (1s), *tiémblas* (2s), etc., but in other dialects has a simple vowel throughout. This type of change is a change in the class membership of the lexical item. An original diphthongizing verb moves into the class of regular verbs. This type of change is discussed in Bybee (1980).

2. Romance preterites

The verb forms we will be discussing are descended from Latin Perfect forms. The Perfect presents certain analytic problems in Latin, especially with regard to the person–number suffixes, for example, the occurrence of *-stī* in 2s, where other tenses have only *s*, the occurrence of *-runt* in 3p, where other tenses have only *-Vnt*, but there is a clearly separable perfect marker, at least in the weak perfects of formal Latin, in the form of *-v-*.

(1)

First conjugation[2]		*Fourth conjugation*	
cantāvī	cantāvimus	dormīvī	dormīvimus
cantāvistī	cantāvistis	dormīvistī	dormīvistis
cantāvit	cantāvērunt	dormīvit	dormīvērunt

The phonetically conditioned loss of the *-v-* ([w]) and at times the vowel following it, led to a situation in which no one segment could be identified as the perfect marker. The following forms are what Menéndez Pidal (1941) calls "latín popular" (p. 310):

(2)

cantái	cantámus	dormíi	dormímus
cantásti	cantástis	dormístī	dormístis
cantáut[3]	cantárunt	*dormíut	dormiérunt

Regular sound changes acting on these forms yielded the modern Castilian (Standard Spanish) forms:

(3)

canté	cantámos	dormí	dormímos
cantáste	cantásteis[4]	dormíste	dormísteis
cantó	cantáron	durmió	durmiéron

[2] Stress is indicated throughout by an acute accent and vowel length by an overbar.

[3] In 3s the loss of *i* rather than *v*, leads to the suffix *-aut* (or *avt*) which is found, along with *-ait* in inscriptions. Menéndez Pidal (1941: 309) says that *-aut* prevailed in Latin. Non-first conjugation forms such as *posiut* are also found in Spanish inscriptions.

[4] The *-steis* suffix of 2p for *-stes*, comes from the addition of 2p *-is*, which occurs elsewhere, to the preterite in place of the simple *-s* (Menéndez Pidal 1941: 280).

A slightly more complex history, which we will not discuss, gave the following Old Provençal forms (Anglade 1921: 272, 294):[5]

(4)

canta 'to sing'		*venre* 'to sell'	
cantéi	cantém	vendéi	vendém
cantést	cantétz	vendést	vendétz
cantét	cantéren	vendét	vendéron

(Also 3p *cantéron*, 1s *vendiei*, 2s *vendiést*.)

The paradigms in (3) and (4) present various analytic problems for the native speaker or learner of the language. As we mentioned above, it is difficult to isolate a single segment or group of segments that can be said to signal perfect (or preterite, as it becomes in the Romance languages), although there is a characteristic pattern of stressing the vowel following the stem. The speaker is faced with a number of questions: for example, in the Spanish data, is the *-é* of 1s a person marker or a preterite marker or both? Is it a special 1s marker for preterite, or a special preterite marker for 1s? And, in these languages, as earlier in Latin, speakers are faced with the problem of assigning a role to the *-r-* of 3p.

The problems in Old Provençal are even more complex.[6] There was a good deal of inconsistency with regard to the suffix vowels, both within and between conjugations, and a considerable amount of homophony between forms of the preterite and forms of other tenses. Even the characteristic stress pattern of the preterite was disturbed by such factors as variation between final and penultimate stress on the 1p and 2p forms (Ronjat 1937: 186, 192), and the existence of both weak and strong conjugations, as exemplified in (4) by *canta* and *venre*, and in (5) by *saber* and *prenre*, respectively (Anglade 1921: 300, 305).

[5] The 1st conjugation suffixes were replaced by those of third conjugation, so that the inflections in (4) do not derive from those in (2) at all.

[6] We use the term 'Old Provençal' to refer to the 'langue d'oc' and 'Provençal' to refer to its descendants, as distinct from the descendants of the 'langue d'oil' (French) to the north. Bec (1963) points out that Old Provençal was a koiné used for literary and administrative purposes, and does not represent a homogeneous regional dialect. This is of little concern since our present purpose is merely to point out the analytic difficulties presented by the preterite at this time. Our theoretical points will concern the developments in the modern dialects, which are well documented.

Fig. 1. Locations of dialects referred to in the text. Dialect areas indicated on map: points indicate cities or towns; dotted circles indicate regions of varying size.

(5) *saber* 'to know' *prenre* 'to take'

saup	saubém	pris	prezím
saubíst	saubétz	prezíst	prezítz
saup	sáupron	pres	préiron

The strong verbs (some infinitives in *-re* and *-er*, and all *-ér* infinitives) were distinguished from the weak verbs (all *-a(r)* and *-ir* infinitives, and some in *-re* and *-er*) by exhibiting stem stress in 1s, 3s and 3p, and by their lack of person/number inflection in 1s and 3s (Anglade 1921: 262). Finally, some verbs (including *prenre*) manifested as many as three stem allomorphs in the preterite.

The way speakers handle these complex analytical situations is revealed by the changes that take place in these paradigms in the various dialects of Spanish and Provençal. The location of the dialects surveyed is indicated on the map (see fig. 1). We have attempted to make our study exhaustive enough to provide the full picture of the range of possible changes that can result from the situations exemplified above.[7] An analysis of these changes provides evidence for some general principles governing diagrammatic relations. The first to be illustrated is the basic–derived relation described in the preceding section.

[7] The Provençal dialects shown on the map are scattered over the entire Provençal area and are representative of all the dialects to which we found any reference. The Spanish dialects are concentrated in the northern part of Spain because this is the only area where the changes to be discussed take place, with the possible exception of the change of 1p preterite *cantámos* to *cantémos*, which is quite widespread.

Provence
1. Toulouse
2. Lomagne
3. Foix
4. Couserans
5. Montignac
6. Limoges (Limousin)
7. Bas-Limousin
8. Rochechouart
9. St. Mathieu
10. Confolens
11. Montembœuf
12. Montbron
13. Aurillac
14. Aquitaine
15. Charente
16. Marche
17. Clermont-Ferrand
18. Ambert
19. Gilhoc
20. Loriol
21. Chabrillan
22. Châtillon-en-Diois
23. Gap
24. Nice
25. Forcalquier
26. Marseille
27. Nîmes
28. Alais (Alès)
29. Gévaudan

Spain
30. Babia, Laciana and Sisterna
31. Sanabria
32. Aliste
33. La Ribera
34. Lena
35. Cabrales
36. La Pas
37. Valle de Aragüés
38. Bielsa

3. Third singular as a basic form

3.1.

Watkins (1962) presents several clear examples of all the person forms of a tense being re-formed with the 3s taken as the base. Such re-formations are often preceded by a reanalysis of an original form composed of *stem* + *tense/aspect* + *person/number* as *stem* + *tense/aspect* + ∅. The following Provençal forms show just this sort of development:

(6) *Charente* (Meyer-Lubke 1923: 352)

cantí	cantétem
cantétei	cantétei
cantét	cantéten

(7) *Clermont-Ferrand* (Ronjat 1937: 193)

cantéte	cantétem
cantétes	cantétetz
canté	cantéton

The original 3s suffix *-t* has been extended to the other persons of the preterite (except 1s in Charente – a matter we return to in section 7). This change has two steps: the interpretation of *cantét* as containing no specific mark for person; the *-t* instead is analyzed as signaling preterite. Second, the choice of 3s as a base which can be used to create the other forms by the addition of person/number inflections drawn from other tenses. This second step is further confirmed by the fact that the rebuilt forms have stress on what was the final, stressed syllable of the 3s. It does not matter that the final *-t* of 3s has been deleted in some dialects (e.g. Clermont-Ferrand in (7) above).[8] In Old Provençal the *-t* still appeared before a vowel-initial word.

In some dialects, final *-t* has been replaced by a velar *-c* ([k]) in the 3s (Ronjat: 192; Anglade: 272; Meyer-Lübke: 353). In other dialects, such as that spoken in Foix, the velar appears throughout the preterite (voiced to [g] intervocalically) (Ronjat: 192).

[8] According to Ronjat, this is a quite recent phonetic development, not yet completed in some areas. In Marche and in some subdialects of Limousin, for example, the 3s *-t* is still pronounced, but only when followed by a pronominal subject.

(8)

cantégui	cantéguen
cantégues	cantéguets
cantéc	cantéguen

This development is apparently parallel to that of the preterites with *-t-*, with the 3s form in *-c* serving as the basis for a re-formation of the entire tense. Speakers may have taken the velar to be a marker of preterite because of the existence in the same dialects of a small class of highly frequent strong verbs with a velar stop characterizing the entire preterite. In these verbs, the velar arose because of phonetic developments involving the strengthening of Latin *-ui* to [g^{w}], and a concomitant absorption of the preceding stop.

(9) *voler* 'to wish, to want' (Grandgent: 144)

vǫlc, vuęlc, volguí	volguem
volguíst	volguetz
vǫlc	vǫlgron

This class included such verbs as *aber* 'to have', *poder* 'to be able', and *tener* 'to have, to hold'. Furthermore, it is significant that, while the velar began to spread through the weak paradigms as other forms were rebuilt off the 3s form, the 1s of the strong verbs often added the weak inflection *-i*, accompanied by final stress, thus making the 1s and 3s forms distinct: for example, 1s *degui*, 3s *dec*. Note that it is the 1s form which was re-made; the 3s form itself was under no circumstances re-made (Grandgent: 138).

As we noted earlier, the 3p form of the preterite was particularly problematic for speakers of Provençal, as the *-r-* which occurred there could not be assigned a specific semantic value. This *-r-* was restricted to 3p, and so could not be taken as a general preterite marker. It did not occur in any other 3p forms, and thus could not be a 3p marker. While there was an *-r-* in the future and conditional, it occurred in all persons and so could not be related to the *-r-* of 3p preterite. The problem the speakers faced, then, was that of determining the status of this *r*. As Blaylock noted (1975: 434), with regard to another verbal augment, the Latin *-sk-*, "where an empty [i.e., apparently meaningless] morph ... separates a neatly silhouetted base from an equally well-delineated inflectional ending, both grammarian and lexicologist can claim a legitimate interest, while the native speaker strives to clarify the status of the bothersome element...".

The changes in the various dialects show how speakers resolved the

question of this 'bothersome' element. In Lomagne, speakers apparently concluded that the *-r-* served no function whatever, and they dispensed with it entirely (Ronjat 1937: 272):

(10) cantéi cantém
cantés cantéts
cantéc cantén

The stress was regularized, so that it is now consistently final throughout the preterite. The problem of potential homophony with the present tense, which shares many of the preterite inflectional endings, did not arise, since in the present the stress was retracted in 1p and 2p, making it uniformly penultimate.

The other possibility for dealing with the *-r-* involved a segmentation of the 3p such that the *-r-* was interpreted as linking the stem and the 3p inflection, *parlé* + *r* + *on*. Once its role as a linking consonant had been established, the *-r-* could be used in the re-formation of other persons from a 3s base that ended in a vowel. In the examples in (11), it can be seen that the addition of a vowel-initial person/number inflection to a vowel-final stem, such as the 3s, would require either a vowel merger, or the 'interfixing' of a consonant to separate the two vowels. The latter solution yields a transparent sequence of morphemes and is adopted by the following dialects (Ronjat 1937: 187, 225, 226):[9]

(11)

Literary Provençal		*Nimes*	
parlére	parlérian	cantére	cantén
parléres	parlérias	cantéres	cantés
parlé	parléron	canté	cantéroun

Forcalquier

parlí	parlérian
parléres	parlérias
parlé	parléroun

For example, 2s is re-formed by using the 3s as a vase, *parlé*, adding the linking consonant *-r-* and a 2s suffix, *-es*: *parlé* + *r* + *es*. Subsequently,

[9] The 1s and 2s inflections were taken from the present tense, and the 1p and 2p inflections from the imperfect.

the *-r-* may be interpreted as a marker of preterite, as more forms are re-made in this way. Observe that 3p does not serve as a base, but only provides the *-r-* for use as a linking consonant, and that role depends entirely upon an analysis of 3p as consisting of the 3s as a base, plus *r*, plus the 3p suffix *-on*. Thus in all the Provençal preterite changes involving a consonant, whether it is *t*, *c* or *r*, it is clear that the 3s serves as the basic form. In some dialects, not all persons of the preterite were re-formed. In section 7, we will discuss the different patterns of re-formation.

3.2.

Another example which shows the 3s form to be the basis of the other persons of a tense comes from a Spanish dialect in the valley of Bielsa in the Pyrenees. Consider the preterite forms in Bielsa (Badía Margarit 1950: 128–131):

(12)

1st conjugation		*2nd and 3rd conjugations*	
llevé	llevémos	metié	metiémos
llevóres	llevéz	metióres	metiéz
llevó	llevóren	metió	metióren

These paradigms exhibit several interesting peculiarities to which we will return below. In particular, note that the 2s has been remade on the basis of the 3s, using the *r* from the 3p, just as in the Provençal dialects discussed above.

What interests us for the moment, however, is the relation between the 1st cojugation suffixes and those of the 2nd and 3rd conjugations. It is immediately obvious from (12) that all three conjugations use the same suffixes: *-é*, *óres*, *ó*, *émos*, *éz*, *óren*, but that in the 2nd and 3rd conjugations, each suffix is preceded by *i*. This vowel is actually pronounced as the glide [y] when prevocalic. A comparison with the Castilian 2nd and 3rd conjugation forms reveals that the only unchanged form is the 3s:[10]

(13) metí metímos
metíste metísteis
metió metiéron

[10] The Castilian forms are probably not identical with the forms that preceded the current Bielsa forms, but they do show the etymological vowels following the stem.

Furthermore, if the forms in (13) are compared with the 1st conjugation forms in (12), it is clear that the only form that could have served as the basis for the remodeling of the Bielsa 2nd and 3rd conjugations is precisely the 3s form.

On the basis of the form *llevó*, speakers could quite correctly decide that *o* is the marker of preterite, 3s. If this analysis is applied to *metió* [metyó], speakers might further conclude that the stem for the preterite is *meti-* [mety-]. When the 1st conjugation suffixes are attached to this preterite stem, the current Bielsa forms result. Notice that in keeping with the hypotheses discussed in section 1, the stem for the preterite tense is used to build all the persons of the same tense, but the person markers come from the same person in another paradigm.

This change, or series of changes, probably took place over an extended period of time, with some forms changing before others, and with a period in which both conservative and innovating forms competed. The point we would like to stress here is that only the 3s form could be the source of this restructuring, for it is only in the 3s that the correspondence between the suffixes (*-o*) in the three conjugations could lead to the reanalysis of the stem as ending in a glide in the 2nd and 3rd conjugations.

These examples, then, demonstrate the existence of a basic–derived relationship among the forms of a paradigm, with the 3s often taken as the most basic.

4. Autonomy

4.1.

Two factors influence the basic–derived relationships that speakers construct. (1) Morphological categories such as tense, person and number establish the axes upon which such relationships stand. We will adopt a rather traditional view about such categories and not say more about them here. (2) The directionality of the basic–derived relationship depends on the degree of autonomy: the more autonomous form serves as the basic form and the less autonomous forms are derived.

The degree of autonomy is a property of individual words, not of morphemes, nor of members of categories as whole sets. It is the extent to which a word is likely to be represented in the speaker's lexicon as a whole and separate unit. Our hypothesis is formulated as a matter of degree

for two reasons: to allow for the possibility that individual speakers may make slightly different decisions concerning lexical storage, and because there are (at least) three determinants of autonomy in morphologically related forms, all of which may be present in a given word in varying proportions. A hypothesis formulated in such a way makes predictions of statistical tendencies in diachronic change, language acquisition and psycholinguistic experimentation. It cannot, nor it is intended to, generate a unique grammar for a body of linguistic data.

Since a fully autonomous word is a word with its own lexical representation, any word which is not morphologically related by productive rule to another word will be autonomous. That is, a word is autonomous if it is not semantically and phonologically predictable from some other word by general rule (Hooper 1976a). Thus English *archive*, *sibling* and *flagrant* (to take some random examples) are all autonomous, but a plural such as *siblings* need not be, for it can be derived by a productive rule. This position on lexical storage is not controversial and will not be discussed further. We turn now to a discussion of autonomy in words related by productive morphological rules. The discussion is phrased to apply primarily to inflectionally related words, but it is also applicable to words related by productive derivational rules.

4.2.

Among the inflectionally related words of a paradigm, some will be more autonomous than others; that is, some will be more amenable to lexical storage, and some to derivation by rule. We will consider the following three factors as determinants of autonomy: semantic simplicity, word frequency in discourse, and morphophonemic irregularity. We begin with a consideration of relative semantic simplicity.

Our hypothesis is that words with fewer semantic components to segment can more readily be stored as whole units, while words with more semantic parts are more likely to be broken down into separate morphological units. That is, more units of meaning look for more units of expression. This notion of semantic complexity depends upon the theory of markedness, in which one member of a category is 'unmarked' semantically and thus simpler. For example, Spanish 3s present indicative *cánta* is simpler than 1p present indicative *cantámos* (even though it takes the same number of traditional terms to describe them – third, singular, present and indicative for *cánta*, first, plural, present and indicative for *cantámos*) because the 3s

is unmarked for person and number. The 3s does not contain any information concerning person and number, unless one infers that the lack of person/number marking signals 3s. With regard to our hypothesis, it should be clear that there is less segmenting to do for *cánta* than for *cantámos*, and less for *cantámos* than for *cantábamos* (1p imperfect indicative). Of these three *cánta* would be the most autonomous.[11]

Words that are highly frequent in discourse are more likely to have separate lexical entries than less frequent words. It is well known that frequency is an extremely important determinant of rote learning. A highly frequent model does not require decomposition by rule. Numerous diachronic changes attest to the nonsegmentation of highly frequent sequences, as when idioms become fixed, or verbal phrases such as *going to* and *got to* become *gonna* and *gotta*. Less frequent words and phrases maintain the separate identities of their parts, and thus seem not to be autonomous.

Word frequency is an important diagnostic for establishing the degree of autonomy, because it is a property of words that can be measured with more confidence than some other properties which determine autonomy. It must be remembered, however, that word frequency is a result of the meaning of the word, the real world situation it refers to, and the view of the world taken by speakers.

Jakobson's (1957) definition of unmarked and marked clearly indicates a wider range of usage and thus a higher frequency for the unmarked member of a category:

> The general meaning of a marked category states the presence of a certain (whether positive or negative) property A; the general meaning of a corresponding unmarked category states nothing about the presence of A and is used chiefly but not exclusively to indicate the absence of A. (Jakobson 1957: 136)

[11] There are several reasons why we have chosen not to phrase the discussion entirely in terms of markedness. First, we do not consider markedness to constitute explanation in itself; rather, we propose to break the notion of markedness into its constituent parts to determine the causal relations among them. We find a variety of causes for autonomy, and these may be somewhat obscured by adopting a unitary notion of markedness. Further, we do not believe that markedness relations are the same type in phonology and morphology, and we suspect that even within the realm of morphology there may be a variety of types of relations, which might remain uninvestigated if all relations were assumed to be of a single type. Finally, we feel that it is more useful, in the sphere of morphophonemics, to develop the notion of autonomy, which, unlike markedness, is a property of *words* that is independent of morphological relations.

Furthermore, Greenberg observes that frequency can be based on "the situation in the world with which the users of the language must deal" (1966: 66). Thus in English *author* means writer of either sex, or male writer, but, because of the real world situation in which nurses are usually females, *nurse* refers to nurses in general or to female nurses. Similarly, the most frequent case of a noun may vary according to the referent – the word for hammer may be most frequent in the instrumental, place names in the locative (Mańczak 1958), etc. In addition to the semantic range of the word and the situations in the real world, frequency may also be conditioned by the psychological processes that determine how we perceive reality and how we use language to refer to it. For instance, the more frequent use of singulars may be due to a tendency to focus attention on single referents, and frequency of 1s forms may be attributed to an egocentric bias. (Zubin 1979 applies these psychological principles to the German case system.)

It is uncontroversial that a certain amount of morphophonemic irregularity or opacity forces autonomy. (For present purposes, irregularity will refer to the presence of a morphophonemic alternation that is restricted to certain morphemes, and opacity will refer to portmanteau expression or the absence of a transparent segmentation of morphemes.) In cases of total suppletion, such as English *go*, *went*, separate lexical entries are clearly needed. With less striking cases of opacity or irregularity, such as the English strong verbs, it is not so easy to decide on the basis of internal evidence alone whether a derivation by rules, which would be quite restricted and complex, is preferable to separate lexical storage. We hypothesize, however, that any amount of irregularity or opacity can produce a tendency towards autonomy, and that a more precise picture of the relation between these properties and autonomy can be established experimentally. (Cf. section 4.5.)

4.3.

The autonomy hypothesis is, then, that words will have their own lexical representation if they have one or more of the following properties: low semantic complexity, high frequency and morphophonemic irregularity or opacity. There is considerable diachronic evidence for this hypothesis. It has long been observed, and more recently documented by Mańczak (1958, 1963, 1980) that certain forms in paradigms are more *resistant* to morphophonemic change than others. Mańczak reports that these are the more frequent words of the paradigms, usually singulars, third persons, present

tense, and indicative mood. Since the morphophonemic changes in question result from the replacement of old forms with new forms which result from the application of morphological and morphophonemic rules, the autonomous forms, being registered lexically and not produced by rule, are less likely to change, and the nonautonomous forms, those produced by rule, are most likely to change.

It has also been observed at least since Paul (1890), and documented recently in Hooper (1976b) for English strong verbs, that very frequent paradigms retain their morphophonemic irregularity much longer than infrequent paradigms. This fact is also accounted for by the autonomy hypothesis, and furthermore, reveals the interplay between morphophonemic irregularity and frequency in determining autonomy. Specifically, in order for a word to maintain a morphophonemic irregularity diachronically, it must be frequent enough to be learned and registered as an autonomous form. If it is not, it will be replaced by a form derived by rule. The acquisition of autonomous and nonautonomous forms is discussed in the next section.

The phenomenon of split paradigms also presents evidence for the autonomy hypothesis. Split paradigms are paradigms which are based on more than one stem, and each of these stems has an etymologically different source. English examples are *go/went* and *be/am, is, are/was, were* (the latter having three etymologically separate stems). In order to create a split paradigm, such as Modern English *go*, the past forms of *wend* had to become independent of the present tense forms, and gradually replace the past forms of *go*. This process requires a split in both the paradigms of *go* and of *wend*. These splits could only occur if the past forms had separate lexical entries from the present forms. Such splits would be impossible if the past forms were derived by rule from the present forms. In the case of *go*, the split of the present from the past is aided by the fact that the prior past forms of *go* were also from a different stem (Old English *gā* (present) and *ēode* (past)), and would have been autonomous for this reason. A connection between split paradigms and morphophonemic irregularity is argued for by Bolozky (1979), who observes that a large number of split paradigms in Modern Hebrew have arisen in paradigms that previously displayed other morphophonemic irregularities. However, morphophonemic irregularity is not a necessary motivation for paradigm split. Paradigms do not split simply because their forms are so different as to appear unrelated, for while *went* was not a totally regular form, its morphological connection to *wend* could hardly be considered opaque. Rather, in this

case, we must look to frequency of use, and realize that a highly frequent word may be autonomous, even if it is transparently derivable from some semantically simpler related word. In this regard, it should also be mentioned that split paradigms occur most often in the very frequent lexical items in any language. (Rudes 1980 discusses many such examples.)

Paradigm splitting and resistance to morphophonemic change both support the autonomy hypothesis, and furthermore give evidence for the roles of frequency and morphophonemic irregularity in determining autonomy. We turn now to an examination of the language acquisition process with respect to our hypothesis.

4.4.

The acquisition of morphology and morphophonemics takes place through a combination of rote learning and rule formation. MacWhinney (1978) gives a detailed account of the process, which is usually thought of as comprising at least three stages. In the earliest stage, morphologically complex words are learned by rote, as amalgams, but little segmentation or analysis is applied to them. At this point, many irregular forms are produced correctly, but children are not able to apply rules to nonce forms. In the second stage, the child begins to formulate rules, and in this stage many overgeneralizations occur. Sometimes irregular forms that were earlier produced correctly are now overgeneralized. Thus a child, who made the past tense of *break* as *broke* earlier, may now begin to produce *breaked*. The child, then, has two ways of producing a past tense form: one is to use a stored irregular form; the other is to apply a regular rule. The child's task is to learn which method works for which verb. In this stage, the child seems to prefer rule application. However, irregular and regular forms of the same verb are often used side by side, and some verbs do not seem to be so prone to overgeneralization as others. Slobin (1971) studied overgeneralizations of English past tense based on a sample of the speech of 24 children between the ages of 1½ to 4 years, and found many more correct past tense formations than regularized ones. Furthermore, there are some verbs that are never regularized in this sample. Slobin also computed the input frequency of these verbs based on 64 hours of speech to children in this age range by mothers and investigators in natural dialog. Z. Greenberg (1975) found a positive correlation between correct forms and input frequency in Slobin's data. At this point in the acquisition process, according to MacWhinney, the child is producing forms

and monitoring his or her productions: comparing the forms produced by rule with rote forms. The success of rote forms against rule generated forms depends upon what MacWhinney refers to as their "lexical strength" (1978: 9). Lexical strength apparently depends upon frequency in input – availability of the model, and frequency in production by the child, since the more often a word is produced, whether correct or incorrect by adult standards, the more right it will seem to the child. In the final stage, the child must sort out the irregular forms and learn them. This process seems to proceed one verb at a time, and resembles rote learning more than minor rule acquisition.

It is very difficult to account for the perseverance of irregular forms in a language, such as the English strong verbs, without according an important role to word frequency, especially considering that only the most frequently used irregular forms remain over periods of time. But the importance of *token* frequency suggests rote learning, since rule formation requires high *type* frequency (MacWhinney 1978).

Our autonomy hypothesis would further predict that there is a much higher degree of portmanteau expression in highly frequent forms, just because they can be learned unanalyzed and do not need to be transparently segmentable. For the same reasons, highly frequent *lexical* items will often maintain portmanteau expression. There is evidence that these properties are indeed the result of the way the forms are acquired. Simões and Stoel-Gammon (1979), reporting on the acquisition of verbal forms for four children acquiring Brazilian Portuguese, observe that several verbs in the children's speech were not subject to the regularizing processes that affected other verbs, i.e. 3s was not used for 1s, but rather 1s was used correctly from the beginning. Simões and Stoel-Gammon point out that these verbs have two characteristics in common: they were very frequent in the children's speech, and they were monosyllabic. The majority of Portuguese verbs are polysyllabic, with the stem consisting of one syllable or more, and the inflections adding extra syllables, e.g. *falo* 'speak (1s)', *fala* (3s). The verbs that were acquired as exceptions, e.g., *vou* 'go (1s)', *va* (3s), *dou* 'give (1s)', *da* (3s), have only a single consonant as a stem, and are thus difficult to segment. It seems that both their frequency and their opacity contribute to the likelihood of their being learned by rote.

4.5.

Psycholinguistic research has concentrated a good deal of attention on studying the rule-formation process in developing children, but has devoted considerably less energy to determining the interplay between rote learning and rule formation. The major problem here is that if a child correctly produces a word, and also shows evidence of having a productive rule that would generate the word, we still do not know if the word was produced by rote or by rule. For example, Berko (1958) found that children aged 4 to 7 years could for the most part produce a correct plural for the word *glass*, but only about a third of them supplied the plural of the nonce form *tass* as *tasses* (with the *əz* allomorph). This would indicate that they do not yet have the rule that produces *-əz*, and that *glasses* is a rote form. At a later stage, when there is evidence of a rule producing *-əz*, that is, when *tass* is pluralized 'correctly', is *glasses* still produced by rote, or is it now produced by rule? This, we do not know.

One attempt to approach this question uses a reaction time experiment. MacKay (1976) asked subjects to produce the past tense forms of English verbs and measured the time it took the subject to begin the response. Differences were found, for example, between verbs such as *talk/talked* and *teach/taught*, but it is not clear how this data should be interpreted. MacKay assumes that a longer reaction time will be required as more rules have to be applied to produce the past tense form. But there is no evidence that this is so. It may be that a longer time is required to look up an irregular past tense in the lexicon.

Perhaps a more promising experimental paradigm is used in Stanners et al. (1979). In these experiments, subjects were asked to indicate whether a sequence of letters presented visually constituted an English word or not, and their reaction time was measured. If the same word is presented twice in the same test, the decision time for the second presentation is considerably reduced, even if several minutes intervene between the first and second presentation. Stanners et al. found that approximately the same results obtained if the first presentation was a form of a verb inflected with *-s*, *-ed*, or *-ing*, and the second presentation was the base verb (e.g. *lends*, *lend*, *pouring*, *pour*). However, if the first presentation was an irregular past tense, for example, *dug* or *shook*, and the second the base verb, *dig* or *shake*, the priming effect still occurred, but was not nearly so great. Stanners et al. suggest that regularly inflected forms do not have separate lexical storage, so that activating any one of them activates the base verb, but that irregular past tenses have their own representations. These representations are connected to that of the base verb, however, which accounts for the existence of the (smaller) priming effect of irregulars. While these

results are quite consistent with our hypothesis, it should be remembered that the stimuli in these experiments were orthographic, so that the relevance to spoken language remains to be established. This does illustrate, however, one of the ways our hypothesis might be tested rather directly.

4.6.

Our second hypothesis with regard to autonomy is that autonomous forms are the basic members in the basic–derived relationship. Given two words formed from the same stem, the more autonomous one is more likely to serve as the basis from which the less autonomous one is derived. This follows from our first hypothesis, because in order for a word to be nonautonomous, it must be derived from some word with a lexical representation.

We are proposing that the larger phenomenon of autonomy underlies Vennemann's basic–derived diagram discussed in section 1.1. Notice that our criterion of semantic simplicity for autonomy in a sense restates Vennemann's hypothesis. We have added a consideration of word frequency, however, as an important basis for establishing lexical representations, and as an important diagnostic for autonomy. The interaction of word frequency with semantic simplicity predicts that all paradigms need not have the same structure, a situation not countenanced in Vennemann's hypothesis. The consideration of morphophonemic irregularity as a determinant of autonomy ensures the same result. Let us consider briefly how morphophonemic irregularity affects the basic–derived diagram.

Many Spanish verbs have an unchanging stem throughout the paradigm, for example, *cantár* 'to sing' has *cant-* in every one of its forms. It may be sufficient to set up only one lexical representation for such a verb (although the discussion in sections 7 and 8 should be taken into consideration). Some verbs, however, use a modified stem for the preterite, for example, *ponér* 'to put or place' has *pon-* for present indicative, imperfect, future and conditional, but *pus-* for preterite and past subjunctive. This paradigm, then, will consist of two autonomous forms, and two sets of basic–derived relations. In one, a form of the present indicative, possibly the 3s, *póne*, will serve as the basis for forming present indicative, imperfect, etc., while a preterite indicative, such as 3s *púso*, might serve as the basis for other preterites and the past subjunctive. Actually, this paradigm may even have a third base, the 1s present indicative, *póngo*, which could serve as a base for the present subjunctive forms (*ponga*, *pongas*, etc.), if the velar is analyzed as belonging to the stem rather than to the suffix.

As we mentioned earlier, the major evidence for the basic–derived relation is morphophonemic change in which the allomorph used to express the semantically basic member replaces the allomorph of the derived member. The basic members are usually third person, singular, present tense, and indicative (Mańczak 1958, 1963, 1979). This list has the same membership as the list of forms that are the most resistant to change, which argues for a close relation of the basic–derived diagram with autonomy. Notice further that with the autonomy hypothesis we can predict which *lexemes* are more likely to change with the same hypothesis that predicts the *direction* of morphophonemic change.

5. Relative autonomy in person forms

Before turning to a more detailed discussion of changes in Provençal and Spanish preterite forms, it is necessary to discuss briefly the expected relative autonomy of person/number forms. We will refer only to very general principles that apply to many languages, but these principles are quite sufficient for the data to be discussed.

A general observation about person forms is that third person differentiates itself from first and second by being a non-person, the form used with all nouns (Benveniste 1946; Kuryłowicz 1968). First and second persons are of course extremely restricted in the sense that they refer only to the speaker and hearer respectively. Considering also that singulars are unmarked with respect to plurals, the 3s form emerges as Jakobson's (1957) unmarked form and Kuryłowicz's 'base' form.

The various word frequency counts mentioned in Greenberg (1966) show third person and singular verbal forms to be the most frequent. Juilland and Chang-Rodríguez (1964) is a frequency count of Spanish words that distinguishes tense, mood and person forms of verbs. Unfortunately, no totals are given by morphological category membership, so it was necessary to take a small sample of verbs to obtain totals. For this purpose 49 highly frequent verbs were chosen, and the frequency of the person forms in the present and preterite were tabulated.[12] Juilland and Chang-Rodríguez used

[12] We chose the same verbs that Gili Gaya used to obtain the frequencies reported in fn. 14, with the exception of *ser*, which was excluded because it constituted nearly half of the tokens. With the other count, the *Recuento*, only 33 verbs could be used, because the others had noun homonyms for some crucial form, which distorted the frequency.

texts written in peninsular Spanish between the years 1920 and 1940, and distinguished five different styles. Written discourse would not be relevant for the study of person inflections because of the lowered frequency of first and second person reference as compared to spoken discourse. For this reason we have used only the figures for the category 'Plays', in which the frequency of person forms should be comparable to spontaneous spoken discourse.

Another frequency count (Rodríguez Bou 1952) has some advantages over this one: it records actual spontaneous conversations of children in grades 1 through 6, and it is based on considerably more tokens. One disadvantage for present purposes is that it is based on Puerto Rican Spanish, where the 2p is always expressed with the 3p form. Another disadvantage is that homonyms are not distinguished. Thus the 2s familiar imperative is included in the 3s present indicative, and the relation of 1p present and 1p preterite had to be established on the basis of verbs that use different forms for the two, since many verbs use the same form. Despite these problems, the two frequency counts show very similar results:

(14) *Juilland and Chang-Rodriguez*

Present (n = 3570)		Preterite (n = 405)	
1s 23%	1p 7%	1s 31%	1p 4%
2s 16%	2p 1%	2s 7%	2p 0%
3s 44%	3p 9%	3s 47%	3p 10%

Rodriguez Bou

Present (n = 14,332)		Preterite (n = 10,414)	
1s 24%	1p 4%	1s 22%	1p 4%
2s 11%	2p –	2s 4%	2p –
3s 41%	3p 20%	3s 51%	3p 19%

These figures show 3s to be the most frequent, as predicted by its status as least marked form. The competition offered 3s by the 1s form is striking, especially in the preterite, and clearly sets off 1s and 3s from the other forms. The figures suggest relative autonomy for 1s forms, and this suggestion is supported by other facts as well. J. Greenberg (1966: 44) mentions the occurrence of zero marking for 1s in Dutch and German, and suggests that between first and second person, first is the least marked. (This is contrary to Jakobson's (1957) claim that first person is marked with respect to second.) It is easy enough to speculate on why 1s forms are so frequent

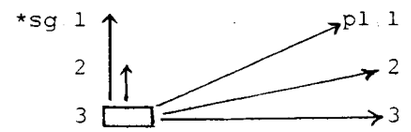
even though their reference is so restricted, and even though it is difficult to make a case for the semantic simplicity of 1s. The high frequency of 1s reflects the ego-centric bias of language users, which is reflected in language in many other ways (Zubin 1979).[13]

Examining now the other forms, we find a frequency ranking after 3s and 1s, of 1p and 3p (although the order is reversed in the preterite), followed by 1p, then 2p. We will use this frequency ranking as a rough guide to autonomy in our discussion of the relations among person forms. This frequency count, although made on Spanish, is applicable to Provençal, as long as we bear in mind that Spanish uses the third person verbal forms for second person in formal contexts, thereby detracting from the frequency of second person in favor of third, while in Provençal, the 2p verb form is used for 2s in formal contexts. Thus for Provençal dialects, the frequency of 2p will be greater, and the frequency of third person will be somewhat less. This will not affect our main points, however, since we are confident that even in Provençal, 3s is by far the most frequent verb form, and 1s is next in order.

6. Paradigm organization

The formal organization of a paradigm diagrams the semantic organization so that forms that are more similar semantically will be more similar in morphophonemic shape. Thus, two forms sharing many semantic properties will be related in that the less autonomous of these will be derived from the more autonomous base form. Given that the six persons within the preterite (of Spanish and Provençal) are all closely related, and 3s is the most autonomous of these, we might expect a set of basic–derived relations such as those indicated by the arrows in (15). However, we have found no evidence for paradigms organized as in (15).

(15)



[13] Note that children acquiring languages with person number inflection use 1s and 3s verbal forms before any of the others (Mikeš 1967; Pačesova 1976; Simões and Stoel-Gammon 1979).

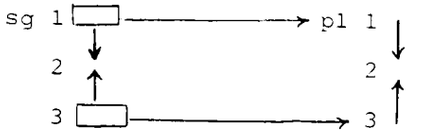
Rather, the person forms seem to be linked to 3s via some semantically related form with intermediate autonomy. Thus we have found evidence for a set of relations as shown in (16):

(16)

In (16) 3s is the basis of the paradigm, and 1s, 2s and 3p are derived directly from it. But 1p and 2p are derived from 3p, rather than directly from 3s. In this pattern the number distinction cuts across the person distinctions and generalizations apply within the singular and plural.

A second organizational type we have found evidence for is built around two base forms, 3s and 1s. Here the relative autonomy of 1s not only exempts it from being derived from 3s, but also allows it to serve as the basis of other closely related forms. Evidence will be presented below for the relations diagrammed in (17).

(17)



This diagram represents the possibility that 1p may be derived directly from 1s. The relatively nonautonomous 2s and 2p are in an ambiguous situation and may be derived from either of the more autonomous forms.

One logical possibility for which we have not found any dynamic evidence is the derivation of 2p from what should be the more autonomous 2s. This should be possible because both forms are second persons. The reason that this alternative is not attested in our data is that the 2s form, while more autonomous than 2p, is not autonomous enough to be formally independent of 1s and 3s.

Given the two organizational possibilities shown in (16) and (17), or combinations of them, we are able to explain a wide array of morphophonemic changes that have taken place in preterite forms in Provençal and Spanish. It should be noted that our account of these changes assumes that the preterite forms (especially 3s) are autonomous vis-à-vis the present

tense forms. There are several reasons for regarding this as highly plausible. First, preterite forms in these languages are the most used of past forms, and are second only to present tense forms in frequency.[14] Second, there is a certain amount of morphophonemic opacity in these preterite forms. In order to derive a 3s preterite from a 3s present, it is necessary to go through two steps: a change in the final vowel, and a change in the placement of stress. The result is a form that cannot be clearly segmented into parts corresponding to the components of meaning, as we mentioned earlier. Finally, there is evidence from children learning Brazilian Portuguese, whose verbal morphology and morphophonemics are quite similar, that 3s preterite forms are learned as autonomous. Portuguese distinguishes three conjugation classes, as illustrated by the following forms:

		falar 'to speak'	*bater* 'to hit'	*abrir* 'to open'
Infinitive:				
Present:	1s	fálo	báto	ábro
	3s	fála	báte	ábre
Preterite:	1s	faléi	batí	abrí
	3s	falóu	batéu	abríu

These are the first four forms used by children. The 3s present is used first, then the 3s preterite is added. Note that the three conjugation classes are fully distinguished only in 3s preterite, and only two classes are distinguished in 3s present. Simões and Stoel-Gammon (1979) report, however, that all four children they studied used the 3s preterite forms with correct suffixes from the very beginning. However, for 1s preterite forms the first conjugation suffix *-ei* was used with all verbs, including 2nd and 3rd conjugation verbs. This suggests an important difference between the way 3s and 1s preterite forms are acquired. In particular, it seems that 3s forms are learned as whole units, while 1s forms are produced by rule in this case.

[14] Gili Gaya (1972) chose 50 of the most frequent verbs used by school children, and tabulated the frequency, according to the *Recuento*, of these verbs in various verbal forms. The five most frequent were:

Present indicative	64,372
Preterite indicative	36,247
Infinitive	29,585
Imperfect indicative	12,359
Imperative	10,423

The relative prominence of the preterite is quite apparent here.

7. Provençal preterites

Earlier we discussed the Provençal strategy for creating new preterite forms using the 3s, a linking consonant, and a person/number suffix. In some dialects all persons except 3s are re-formed. In other dialects, only some of the persons are re-formed. The particular changes attested in the dialects form interesting patterns that can be explained by referring to degree of autonomy and morphological category membership.

All the types of preterite found in Provençal are shown in (18); the subsequent corresponding table in (19) shows the form patterns that are found in Provençal dialects. The 'x' marks the forms with linking consonants. In the case of *r*, this consonant is etymological in 3p.

(18) A. *Bas-Limousin* (Ronjat: 252)[15]

parlí	parlén
parléra	parlá
parlé	parléron

B. *Alais* (Ronjat: 263)[16]

parlére	parlén
parléras	parlés
parlé	parléron

C. *Forcalquier* (Ronjat: 226)

parlí	parlérian
parléres	parlérias
parlé	parléroun

Charente (Meyer-Lübke: 352)[17]

cantí	cantétem
cantétei	cantétei
cantét	cantéten

D. *Provençal litt.* (Ronjat: 224)

parlére	parlérian
parléres	parlérias
parlé	parléron

Clermont-Ferrand (Ronjat: 193)[18]

cantéte	cantétem
cantétes	cantétetz
canté	cantéton

(19) A.

sg 1	pl 1
2 x	2
3	3 x

B.

sg 1 x	pl 1
2 x	2
3	3 x

C.

sg 1	pl 1 x
2 x	2 x
3	3 x

D.

sg 1 x	pl 1 x
2 x	2 x
3	3 x

[15] Bas-Limousin is the only dialect we have found which displays this particular pattern.

[16] Other dialects with this pattern are Gévaudan (Ronjat) and Nîmes (Ronjat).

[17] This pattern is also found in Couserans Limousin, Bassiere-Badil, Rochechouart, Saint-Mathieu, Ambert (Ronjat); in Confolens, Montbron, Montembœuf (Chabaneau); in Gilhoc (Meyer-Lübke). This is also the pattern of regular preterite formation in Catalan, with *r* occurring in all persons, except 1s and 3s.

[18] This is the most common pattern for the preterite, and is found in Forcalquier, Marseille, Gap, Montignac, Aurillac, Clermont-Ferrand, Toulouse, Chabrillan, Loriol, Châtillon-en-Diois, and throughout Aquitaine (Ronjat); in Niçois (Compan); in Rhodanien (the modern 'standard' Provençal, as reported by Ford); and in Provençal *littéraire* (Ford).

We have already observed that the 3s form does not change, and we have claimed that this form serves as the basis for the other changes. The other striking regularity in the paradigm is that 2s is *always* changed, even if no other changes have taken place, as in A. This is because 2s is the least autonomous of the singular forms. We hypothesize that the 1s and 2s forms are the most likely to be re-formed because they are more closely related to 3s than the plurals. Between 1s and 2s, the 2s form is more likely to change than 1s because 2s is always less autonomous than 1s. These two factors together, the degree of autonomy and the morphological category membership, explain why 2s is most readily re-formed.

The 1s also undergoes change in some dialects, as we might expect, since it is a singular. However, 1s is the only form (other than 3s) which can still resist a change, even when all other forms have undergone it. Here, then, we have evidence for the relatively high autonomy of 1s.

In evaluating the changes in plural forms, it is important to remember that the *-r-* originates in the 3p. With this in mind, it can be seen that either all the plurals take the linking consonant or none of them do. This would depend upon whether the consonant is taken to be a general marker of preterite. If it is, none of the plural forms is autonomous enough to resist re-formation, so they all adopt the consonant and the new set of suffixes. We do not interpret types A and B as having 1p and 2p forms that resist change, but rather it appears that the re-formation has been restricted to the singular. Note in this regard that preterites using *-t-* or other consonants show a uniform re-formation of the plural.

8. Spanish preterites

Changes in Spanish preterites involve the vowel following the verb root, which is usually referred to as the theme vowel. This vowel was originally uniform throughout the tense, but due to a number of regular sound changes, alternations in this vowel were introduced. The standard Castilian preterites are repeated here as examples of forms unaffected by the type of change we will discuss:

(20)

1st conjugation		*2nd and 3rd conjugations*	
canté	cantámos	dormí	dormímos
cantáste	cantásteis	dormíste	dormísteis
cantó	cantáron	durmió	durmiéron

In many dialects, the 1st conjugation forms have *e* or *o* rather than *a* as the theme vowel in the 2s and the plural, giving evidence for re-formation of these person forms on the basis of 1s and 3s. The exact patterns of re-formation will be our concern in this section.

The examples in (21) illustrate all the attested patterns for theme vowels in first and second person of first conjugation. We will discuss the third person changes separately below.

(21)

A. *La Ribera*[19]		B. *Valle de Aragüés*[20]	
canté	cantémos	canté	cantémos
cantátes	cantátis	cantástes	cantéis
cantó	cantáron	cantó	cantóron
C. *Asturias*		D. *Lena*[21]	
canté	cantámus	canté	cantémos
cantéste	cantásti	cantéste	cantésteis
cantóu	cantánu	cantó	cantáron

The italicized vowels are the changed theme vowels that we will be discussing. There are also a number of changes in the person/number inflections, but we will have to ignore them, and concentrate on the theme vowels.

The paradigm in A illustrates a single change – the re-formation of the 1p form using 1s as a base. A form such as original *cantámos* is replaced by *cantémos*, which is made up of 1s *canté* plus the 1p marker *-mos*, which is the unique and constant marker of 1p. This change is apparently the

[19] Preterites with *e* in 1s and 1p only are documented in Asturias (García 1946: 168), the two provinces of Castille (Menéndez-Pidal 1941: 311–312), and in the Pasiego dialect (Penny 1969), as well as in La Ribera (Llorente Maldonado 1947).

[20] Preterites with *e* in 1s, 1p and 2p are found in the valley of Aragüés (González-Guzman 1953), and in Babia and Laciana (Alvárez 1949), and in some of the Sanabria dialects studied by Krüger 1954.

[21] Preterites with *e* in 1s, 1p, 2s and 2p are found throughout Asturias, León and Aragón (Zamora Vicente 1970) specifically in Sisterna (Álvarez Fernández 1960), Cabrales (Cañedo 1963), Lena (Neira 1955), Aliste (Baz 1967) and La Ribera (Llorente Maldonado 1947).

first one to occur, with ample documentation in León and other areas in the Middle Ages (Krüger 1954: 72). It is also the most widespread geographically. This we attribute to the high autonomy of 1s, and the close relation of 1p to 1s.

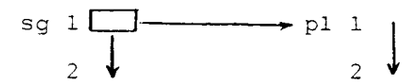
The B-paradigm type has undergone the same change as A, and one further change – the 2p form has also been re-made using the *e* vowel. The change of 2p is in some way dependent on the prior change of 1p. The change of 2p does not, however, depend on a change in 2s, as seen in B. This suggests that 2p is re-formed because of its relationship to 1p.

Only one case of the C-type paradigm was found, and this occurs where neighboring dialects usually have the D-type, which is extremely common. Both types show that 2s may be re-formed on the basis of 1s. If all three changes occur, as in D, the result is a system in which all first and second persons have the same theme vowel, distinguishing them from third person. This would create a nice opposition of 'participant in the speech event' versus 'nonparticipant' if it were not for the fact that the third person verb form is used with *usted(es)* for second person formal. Thus the distinction between participant and nonparticipant applies only to informal speech situations. It may be that this constitutes the real generalization, but we are hesitant to make any claims in this regard. In any case, it is clear that 1s is serving as the morphophonemic base for both first and second persons.

It should be further noted that these changes across persons demonstrate that the primary function of the *é* of 1s is not to mark 1s, but rather to mark preterite. If it were primarily a person marker, it could not serve as the basis for the re-formation of the second persons. This interpretation of the *é* as a marker of the more important verbal categories, tense/aspect, rather than person/number, is another reflection of the hierarchy discussed in section 1.2.

The patterns of theme vowel distribution illustrated in (21) represent the full range of attested patterns. The changes that led to these patterns give evidence for relations among first and second persons as illustrated in (22).

(22)



These examples clearly indicate that 1s can serve as the basis for other morphologically related forms.

In some dialects, another parallel series of changes overlays those we have just discussed. These changes occur in the dialects of León which have a diphthong in 1s, i.e., *ei* (< *ai* < *avi*). Consider first Aliste, which has the diphthongs and has undergone the changes we have just discussed (Baz 1967).

(23) *Aliste*

sg 1 -éi	pl 1 -émos
2 -éstes	2 -éstes
3 -óu	3 -ón

Some dialects, such as Sanabria and Sisterna, farther to the north, have undergone a second re-formation of participant forms, using the 1s diphthong as the basis.[22]

(24) *Sanabria* (Krüger 1954: 54)

sg 1 -éi	pl 1 -éimos
2 -éste	2 -éstes
3 -óu	3 -ónen

Sisterna (Fernández 1960)

sg 1 -éi	pl 1 -éimus
2 -éiste	2 -ésti, -éisti
3 -óu	3 -ánun

As in the cases discussed above, the most commonly attested change is that affecting the 1p form. The difference between this change and the preceding one is that 2s seems to be affected more commonly than 2p (Krüger 1954). This can be seen in the forms in Sisterna, where 2s has changed completely, but 2p still has two competing forms. More evidence of this sort would lead us to postulate a relation between 2s and 2p.

With the 1s serving as the basis for first and second persons, the 3s, in most dialects, has only 3p within its domain. Here, quite predictably, we find the *ó* of 3s showing up in 3p. There are a variety of 3p forms attested. The simplest occurs in Aliste (see (23) above) where the stressed *ó* of 3s is simply followed by the *-n* that marks 3p in all other tenses and moods. Most dialects, however, maintain the bisyllabic nature of the 3p suffix, altering it from *-áron* to one of the following: *-óron*, *-ónon*, *-órun*, *-óren*, *-órin* (Krüger 1954: 76). Some of the variation in the unstressed vowel is phonological. The point that interests us, however, is the *ó* in the stressed syllable, which shows that 3p has been re-formed from 3s.[23]

To complete the picture of theme vowel changes, it should be mentioned that there are also regions in Aragón where the preterite theme vowel is consistently *ó* (Alvar 1953: 236–237).

(25) cantó	cantómos
cantós	cantóz, cantóis
cantó	cantóron

In this case all the person forms are rebuilt on the 3s base. There are also paradigms such as the following (Alvar 1953: 236):

(26) canté	cantémos
cantés	cantéz
canté	cantéron

This paradigm appears to have built on the 1s form, which would seem to be a possibility, given what we have learned about the role of 1s. However, Menéndez Pidal (1941), Alvar (1953), and Zamora Vicente (1970) all subscribe to the idea that the *e* comes from the Latin *-avit*, which, instead of reducing to *-avt* > *-awt* > *-o*, reduced to *-avit* > *-ait* > *-e*. Malkiel (1976: 449) points out on the other hand that this "pattern of contraction" is unproven, and that the morphological explanation is quite plausible, given the cases where 3s *o* serves as preterite marker. Further, we may add, given the fact that 1s *e* serves as the base for first and second persons in so many dialects, we are not surprised to find a few cases where it serves as the basis of the entire tense paradigm.

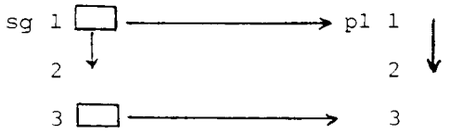
All of these changes show that the theme vowel in the preterite is coming to be a signal of preterite which shows variation corresponding to the person distinctions. In the Provençal data we saw that 3s was the basis of all the re-formations, and that the autonomy of 1s was apparent in the resistance that form exhibited to changes on the basis of 3s. In the Spanish data we see 1s not just resisting change, but serving as the basis for the changes in

[22] Krüger (1954) argues convincingly that the *-ei-* diphthongs in 2s, 1p and 2p are not etymologically derived from *-avi-* as the 1s diphthong is. The argument is that this assumption leaves unexplained the *-e-* that occurs in 2p in Sisterna and in other forms in related dialects. Since in these dialects *ei* has quite consistently remained without monophthongization, the older *e* in 1p, 2s and 2p must be analogical, and therefore the newer *ei* must also be analogical.

[23] A much rarer form is *-eron* (Krüger 1954: 52). This is usually attributed to the second and third conjugation 3p suffix *-ieron*.

other forms. The relations evidenced in the forms in (21) can be diagrammed as follows:

(27)



The precise directions of change in (25) and (26) cannot be discerned. It is clear from (25) that 3s may serve as the base for 1s, as in some dialects of Provençal, and (26) opens the possibility of 1s serving as the base for 3s.

The fact that 1s resists change in Provençal, but actually serves as the basis for change in Spanish dialects is an important argument for the autonomy hypothesis. One could argue that 1s is not serving as a basis for remodeling 1p, 2s and 2p, but rather that a rule rewriting the preterite marker as /e/ in 1s has been generalized to apply to all participant categories (Long 1979). However, the resistance of 1s to change in Provençal demonstrates that a 1s preterite form can have its own lexical representation, since, as we argued above, this is the most straightforward way to account for resistance to change. If a form that resists change has its own lexical representation, then a form that not only resists change but also serves as the basis of change must also have its own lexical representation.

The competing autonomy of 1s and 3s opens up several possibilities for divisions of the six persons. In particular the second persons are in a somewhat ambivalent position – they may pattern with the first persons as participants and in opposition to nonparticipating third person, or they may be derived from the third person, creating a division between ego and non-ego forms. Thus in Provençal, 2s is derived from 3s, while 2p seems to always follow 1p.

Although we do not completely understand the factors that influence the way speakers organize and perceive the system, for example, whether second person relates to first or to third, we have been able to narrow considerably the range of possible organizations of the six person/number forms by employing the notions of autonomy and morphological relatedness. The following three generalizations hold for the many dialects we have considered:

(1) 3s can serve as the basis of all other persons.
(2) 1s can serve as the basis of first person and second person forms (and possibly all persons).
(3) Singulars can serve as the basis for plurals.

As a final illustration of the organizational possibilities of person and number, we turn to a brief discussion of the verbal paradigms of Bielsa, where changes according to person have taken place not only in the preterite, but also in the imperfect and conditional.

9. The organization of persons in Bielsa

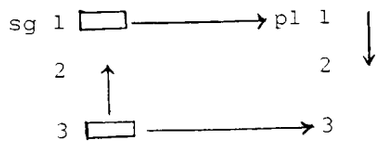
In section 3.2 we discussed the preterites in Bielsa to show that the 2nd and 3rd conjugation suffixes were taken directly from 1st conjugation via the 3s form. The preterite forms are repeated here:

(28)

1st conjugation		*2nd and 3rd conjugations*	
llevé	llevémos	metié	metiémos
llevóres	llevéz	metióres	metiéz
llevó	llevóren	metió	metióren

This is the only Spanish dialect we have encountered where we find *r* used as a linking consonant in the 2s, as it is in Provençal. Neighboring Catalan, however, has a regular preterite with *r* in the 2s and throughout the plural. As in the many cases discussed above, it is apparent that the 2s is formed directly off of 3s with the *r* or *re*, plus the 2s suffix *-s*. The 3p may also be analyzed as derived from 3s. In fact the re-formation of 2s depends upon the analysis of 3p as *llevó* + *r* + *en*. On the other hand, 1p and 2p have been re-formed from 1s. (The *-z* of 2p is the unique and constant marker of 2p throughout the verbal paradigm.) The lines drawn among the forms in (28) show the division of the paradigm due to the relations illustrated in (29).

(29)



What is interesting about this dialect is that several other tenses exhibit similar relations among the persons. Note first that this set of relations is apparent in the future tense of standard dialects, as well as in Bielsa.

(30) *Future*

Bielsa		*Castilian*	
llevaré	llevarémos	llevaré	llevarémos
llevarás	llevaréz	llevarás	llevaréis
llevará	llevarán	llevará	llevarán

The only difference between Bielsa and Castilian is in the 2p marker. The conditional forms in Bielsa exhibit the same pattern, but this is due to changes undergone in Bielsa.

(31) *Conditional*

Bielsa		*Castilian*	
llevarí	llevarímos	llevaría	llevaríamos
llevarías	llevaríz	llevarías	llevaríais
llevaría	llevarían	llevaría	llevarían

In Castilian *ía* always follows the base *llevar* to mark conditional, and the person/number suffixes follow *ía*. In Bielsa only *í* is present in 1s, 1p, and 2p, distinguishing these forms from 2s, 3s and 3p. We do not know the mechanism for this change, but the result suggests that once 1s is changed, 1p and 2p were rebuilt with 1s as a base.

The imperfect indicative has also undergone a number of changes. Observe, however, that the consistent *b* throughout the Bielsa paradigms is a conservative feature.

(32) *Imperfect indicative*

Bielsa: 1st conjugation		*Castilian: 1st conjugation*	
llevábe	llevános[24]	llevába	llevábamos
llevábas	llevábez	llevábas	llevábais
llevába	llevában	llevába	llevában

[24] A form similar to the Castillian 1p, *llevábamos*, probably preceded this Bielsa form *llevános*. The *b* has been deleted, possibly because of the phonetic environment: the 1p form is the only one is which two syllables follow the *b*. The change of *-mos* to *-nos* probably occurred before the deletion of the *b*. We suspect this because the appearance of *nos* rather than *-mos* is a widespread dialectal phenomenon in 1p forms with antepenultimate stress, e.g. *llevábanos*,

2nd conjugation		*2nd and 3rd conjugations*	
metébe	meténos	metía	metíamos
metébas	metébez	metías	metíais
metéba	metéban	metía	metían

3rd conjugation	
partíbe	partínos
partíbas	partíbez
partíba	partíban

In Castillian, 1s and 3s are both unmarked for person, but in Bielsa, the 1s marker is *e*. The 2p form also has this *e* plus the 2p marker *-z*. The 3s, 2s and 3p all contain *ba*. The latter two can therefore be considered to be formed from the 3s. The 1p form has peculiarities all its own, which are discussed briefly in note 24. The imperfect indicative, then, divides up in much the same way as the preterite, future and conditional.

These forms illustrate quite nicely the ambivalent position of the least autonomous forms, the second persons, and the nonnecessity of a symmetrical system. The pairing of 1p and 2p has precedent in the verbal system of Romance, in present indicative and subjunctive, where stress on the suffix vowel sets these forms apart from the others, which have stress on the stem. This stress difference gives rise to vowel alternations in both French and Spanish.

10. Conclusion

With these forms from Bielsa, we can illustrate the hypothesis concerning lexical storage that emerges from the autonomy hypothesis. The data considered here suggest that there could be as many as two lexically stored

and present subjunctive forms in dialects with stem stress, *llévenos*. This develops because speakers consider the verbal suffix *-mos* and the clitic pronoun *nos* to be allomorphs of a single 1p morpheme. The clitic *-nos* occurs attached to the end of the verb when the verb is an infinitive (*hablárnos* 'to speak to us'), a gerund (*hablándonos* 'speaking to us'), or imperative reflexive (*sentémonos* 'let's sit down', in which the *s* of *mos* is deleted). In the latter two cases, the addition of *-nos* produces antepenultimate stress. The form *mos* occurs primarily with penultimate stress (except in the imperfect indicative and subjunctive). Apparently, speakers form the abductive generalization that *-nos* appears preverbally and postverbally with antepenultimate stress, while *-mos* occurs only postverbally, with penultimate stress. Deductive application of this rule led to these changes in the imperfect forms.

words for some tenses for the verbs of Bielsa, since there appear to be two autonomous words for each tense. Thus the preterite in Bielsa could consist of the lexical forms *llevé* (1s) and *llevó* (3s), plus a set of rules that add suffixes to derive the 1p and 2p from the 1s, and 2s and 3p from the 3s. Similarly, the future will have the forms *llevaré* and *llevará*, the conditional *llevarí* and *llevaría*, the imperfect *llevábe* and *llevába*.[25]

We have argued that the autonomy hypothesis helps to explain why 1s can resist change in some Provençal dialects, and serve as the basis of change in some Spanish dialects. We have similarly presented the autonomy hypothesis as the best explanation for the role of 3s in the re-formation of the preterite in Provençal. An alternative account would claim, perhaps, that the 3s form is not being used as the base for other person forms, but rather that the rules that produce 3s are spreading to other persons. The difficulty with this account lies in explaining why a rule that spells out 3s as *t* or some other consonant, would ever apply to a person other than 3s. Similarly, if one claimed that the rule which supplies *r* in 3p spreads to other persons, then one would have to explain why it always spreads to 2s, and only sometimes to the rest of the plural and 1s. On the other hand, the notions of autonomy and degree of relatedness provide good explanations for these changes.

While we have argued from several examples that a speaker's analysis consists of an autonomous word or a paradigm plus rules to derive other closely related words, we have not shown that speakers *never* analyze a paradigm into a stem plus affixes, where the stem does not appear in the language as an autonomous word. In fact, it is conceivable that in more transparently agglutinative languages, this type of analysis might be possible. Moreover, in the same languages we have been discussing here, there are cases that appear to show segmentation of stem and suffix. In section 3.2 we discussed the 2nd and 3rd conjugation preterite in Bielsa, saying that the creation of a new 1s preterite form *metié* [metyé] depended on the analysis of 3s *metió* as [mety + ó]. We have observed also that the existence of autonomous preterite forms does not mean that speakers do not know how to make a preterite given some other form of the paradigm. That is,

[25] This model is similar in some respects to the model of suppletive representation proposed by Hudson (1975) and applied in Hooper (1976a). What is better captured in the current model is the subordinate or derived status of certain forms, which is dependent upon morphological category membership. The arguments given in Hooper (1976a) in support of Hudson's model, i.e., that historical changes show some paradigms to have more than one underlying form, also apply to the current model.

there must be rules that link autonomous forms to one another, at least in morphophonemically regular cases. The question, then, is what is the nature of these rules, and do they presuppose the segmentation and storage of the stem as a separate unit? This would certainly be one analysis, but another possible analysis maintains the notion of an autonomous word. Under this analysis, an autonomous base form may be modified, not just by affixation, but also by changing features or segments. Thus *metió* serves as the base for *metié* in the sense that *metié* can be derived from *metió* by changing one feature of the final vowel, but without ever extracting and storing a separate stem. We would like to leave these and other questions concerning specific analyses open until further debate and more data can clarify the issues.

While the notions of autonomy and degree of relatedness do not in themselves settle all problems that might arise in trying to write a grammar, they do make some very specific predictions about the likelihood of forms to change, and the direction of change, based on facts about frequency, morphophonemic irregularity, semantic simplicity and semantic relatedness. Furthermore, these hypotheses make predictions that can be tested cross-linguistically, concerning the distribution of stem alternants, and the location of morphophonemic irregularity and portmanteau expression in a morphological system.

We hope to have shown, furthermore, that morphological analysis does not need to depend entirely on considerations of economy, symmetry and distributional facts. Indeed, these might be of very little importance in a speaker's analysis. Instead, if we look at dynamic data, we find that facts about child language and historical change support the hypothesis that the speaker's analysis is guided by certain psychological principles of acquisition and storage. These are the considerations that should be input to the development of morphological and morphophonemic theory.

The theory we have been developing here is a natural theory in the sense used most recently by Donegan and Stampe (1979), who describe natural phonology as follows:

> This is a natural theory ... in that it presents language ... as a natural reflection of the needs, capacities and world of its users, rather than as a merely *conventional* institution. It is a natural theory also in the sense that it is intended to *explain* its subject matter, to show that it follows naturally from the nature of things; it is not a conventional theory ... in that it is not intended to *describe* its subject matter exhaustively and exclusively (p. 127).

Morphology and morphophonemics have been subject to a great deal of

description – from traditional grammars on the Latin model to *The sound pattern of English*. Morphological systems have not in general been subject to explanation, nor have they been treated as natural objects whose properties follow from their functions, whose outward form diagrams their content. We hope to have shown here that it is useful to approach morphology and morphophonemics in just this way.

References

Alvar, M., 1953. El dialecto aragonés. Madrid: Editorial Gredos.

Alvarez, G., 1949. El habla de Babia y Laciana. Revisto de filología española (Madrid) 49.

Álvarez Fernández, J., 1960. El habla de Sisterna. Revista de filología española 74.

Andersen, H., 1969. A study in diachronic morphophonemics: The Ukrainian prefixes. Language 45, 807–830.

Andersen, H., 1980. Morphological change: Towards a typology. In: J. Fisiak (ed.), Historical morphology. The Hague: Mouton.

Anglade, J., 1921. Grammaire de l'ancien provençal. Paris: Klincksieck.

Badía Margarit, A., 1950. El habla del valle de Bielsa. Barcelona: Consejo superior de estudios científicos (Instituto de estudios pirenáicos).

Baz, J. M., 1967. El habla de la tierra de Aliste. Revista de filología española (Madrid) 82.

Bec, P., 1963. La langue occitan. Paris: Presses universitaires de France.

Benveniste, E., 1946. Structure des relations de personne dans le verbe. Bulletin de la société linguistique de Paris 43, fasc. 1, no. 126, 1–12.

Berko, J., 1958. The child's learning of English morphology. Word 14, 150–177.

Blaylock, C., 1975. The Romance development of the Latin verbal augment *-sk-*. Romance Philology 28, 434–444.

Bloomfield, L., 1933. Language. New York: Holt, Rinehart and Winston.

Bolozky, S., 1979 (MS). On the autonomy of the tense paradigm: The Modern Hebrew evidence. Amherst: Univ. of Massachusetts.

Bybee, *see* Hooper.

Bybee, J., 1980. Morphophonemic change from inside and outside the paradigm. Lingua 50, 45–59.

Cañedo, J. A. F., 1963. El habla y la cultura popular de Cabrales. Revista de filología española (Madrid) 76.

Chabaneau, C., 1874–5. Grammaire limousine. Revue des langues romanes 6, 171–205; 7, 145–178.

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

Compan, A., 1965. Grammaire niçoise. Nice: Éditions Tiranty.

Donegan, P., D. Stampe, 1979. The study of natural phonology. In: D. Dinnsen (ed.), Current approaches to phonological theory, 126–173. Bloomington: Indiana UP.

Ford, H. E., 1921 [1966]. Modern Provençal phonology and morphology. New York: AMS Press.

Garcia de Diego, V., 1946. Manual de dialectología española. Madrid: Instituto de cultura hispánica.

Gili Gaya, S., 1960. Funciones gramaticales en el habla enfantil. Publicaciones pedagógicas, serie II, no. 24. Río Piedras: Univ. de Puerto Rico.

González Guzmán, P., 1953. El habla viva de Valle de Aragüés. Zaragosa: Monografías del instituto de estudios pirenáicos.

Grandgent, C. H., 1905. An outline of the phonology and morphology of Old Provençal. New York: D. C. Heath.

Greenberg, J., 1963. Some universals of grammar with particular reference to the order of meaningful elements. In: J. Greenberg (ed.), Universals of language, 73–113. Cambridge, Mass.: MIT Press.

Greenberg, J., 1966. Language universals. The Hague: Mouton.

Greenberg, Z., 1975 (MS). Overgeneralization of past tense inflection. Univ. of California at Berkeley.

Hooper, J. B., 1976a. Introduction to natural generative phonology. New York: Academic Press.

Hooper, J. B., 1976b. Word frequency in lexical diffusion and the source of morphophonological change. In: Wm. Christie (ed.), Current progress in historical linguistics, 95–105. Amsterdam: North-Holland.

Hooper, J. Bybee, 1979. Child morphology and morphophonemic change. Linguistics 17, 21–50. (Also in: J. Fisiak (ed.), Historical morphology, 157–187. The Hague: Mouton.)

Hudson, G., 1975. Suppletion in the representation of alternations. Univ. of California at Los Angeles (Doct. diss.).

Jakobson, R., 1957. Shifters, verbal categories and the Russian verb. In: Roman Jakobson, Selected writings, II, 130–147. The Hague: Mouton.

Jakobson, R., 1965. Quest for the essence of language. In: Roman Jakobson, Selected writings, II, 345–359. The Hague: Mouton.

Juilland, A., E. Chang-Rodríguez, 1964. Frequency dictionary of Spanish words. The Hague: Mouton.

Llorente Maldonado de Guevara, A., 1947. Estudio sobre el habla de la Ribera. Salamanca: Tesis y estudios salamantinos.

Long, M., 1979. On Romance preterites and the nature of phonological change. In: F. H. Nuessel, Jr. (ed.), Essays in contemporary Romance linguistics, 230–246. Rowley, Mass.: Newbury House.

Krüger, F., 1954. El perfecto do los verbos *-ar* en los dialectos de Sanabria y sus zonas colindantes. Revista de filología española 38, 45–82.

MacKay, D. G., 1976. On the retrieval and lexical structure of verbs. Journal of Verbal Learning and Verbal Behavior 15, 169–182.

MacWhinney, B., 1978. The acquisition of morphophonology. Monographs of the Society for Research in Child Development, no. 174, vol. 43.

Malkiel, Y., 1976. From falling to rising diphthongs: The case of Old Spanish *ió* < **éu*. Romance Philology 29, 435–500.

Mańczak, W., 1958. Tendances générales des changements analogiques. Lingua 7, 298–325, 387–420.

Mańczak, W., 1963. Tendances générales du développement morphologique. Lingua 12, 19–38.

Mańczak, W., 1980. Laws of analogy. In: J. Fisiak (ed.), Historical morphology. The Hague: Mouton.

Menéndez Pidal, R., 1941 [1968]. Manual de gramática histórica española. Madrid: Espasa-Calpe.

Meyer-Lübke, W., 1890/1902 [1925]. Grammaire des langues romanes. Vol. 2. New York: Stechert.

Mikeš, M., 1967. Acquisition des catégories grammaticales dans le langage de l'enfant. Enfance 20, 289–298.

Neira Martínez, J., 1955. El habla de Lena. Oviedo: Instituto de estudios asturianos.

Pačesova, J., 1976. Some notes on developmental universals in Czech-speaking children. Neurolinguistics 5, 192–198.

Paul, H., 1880 [1970]. Principles of the history of language. Trans. by H. A. Strong. College Park: McGrath Publ. Co.

Penny, R. J., 1969. El habla Pasiega. London: Tamesis Books.

Rodríguez Bou, I., 1952. Recuento de vocabulario español. Vol. II. Río Peidras: Univ. de Puerto Rico.

Rodríguez-Castellano, L., 1954. Aspectos del bable occidental. Oviedo: Instituto de estudios asturianos.

Ronjat, J., 1937. Grammaire historique des parlers provençaux modernes. Vol. 3. Montpellier: Société des langues romanes.

Rudes, B. A., 1980 (MS). On verbal suppletion. Washington, D.C.

Simões, M. C. P., 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., 1971. On the learning of morphological rules: A reply to Palermo and Eberhart. In: D. I. Slobin (ed.), The ontogenesis of language. New York: Academic Press.

Stampe, D., 1973. A dissertation on natural phonology. Univ. of Chicago (Doct. diss.).

Stanner, R. F., J. J. Neiser, W. P. Hernon, R. Hall, 1979. Memory representation for morphologically related words. Journal of Verbal Learning and Verbal Behavior 18, 399–412.

Vennemann, T., 1972. Rule inversion. Lingua 29, 209–242.

Watkins, C., 1962. The Indo-European origins of the Celtic verb. Dublin: Dublin Institute for Advanced Studies.

Zager, D., 1978 (MS). A grey area of morphology. State Univ. of New York at Buffalo.

Zamora Vicente, A., 1970. Dialectología española. Madrid: Editorial Gredos.

Zubin, D. A., 1979. Discourse function of morphology: The focus system in German. In: T. Givón (ed.), Syntax and semantics, 12: Discourse and syntax, 469–504. New York: Academic Press.