

Restrictions on phonemes in affixes: A crosslinguistic test of a popular hypothesis

JOAN BYBEE

Abstract

The hypothesis that inflectional affixes use a restricted set of phonemes and that these are the less marked phonemes of the language is discussed and tested on the verbal affixes in a sample of twenty-three maximally unrelated languages. The results show that the tendency for languages to use only a smaller subset of their phonemes in verbal inflection than would be predicted by chance is only a weak trend and not by any means a universal of language. In addition, the tendency to use less marked or less complex segments in affixes is also only a trend and not a universal. However, some generalizations can be made about languages that have patterned exclusions from affixes. It is argued that no one explanation covers all the facts and that multiple diachronic trends, such as phonological reduction in grammaticization and the re-use of old affixes in creating new grammaticized affixes, produce the weak tendency evidenced by the data.

Keywords: affixes, grammaticization, markedness, phoneme complexity, phoneme inventories

1. The question

In 1965 Roman Jakobson, in discussing the many ways that linguistic expression fails to be completely arbitrary, observed that “affixes, particularly inflectional suffixes, in the languages where they exist, habitually differ from the other morphemes by a restricted and selected use of phonemes and their combinations” (Jakobson 1966 [1990: 414]). As examples, he cites English and Russian, where inflectional suffixes use a very limited set of consonants. The same idea, though restricted to flectional languages, can be inferred from Skalička’s typology. In particular, the statement that affixes in flectional languages tend

to be more uniform than in agglutinative languages implies such affixes would use the same set of phonemes (Dressler 1985, Plank 1998, Skalička 1979).

A further observation is that the phonemes excluded from affixes tend to be the more marked phonemes of a language, e.g., the absence of ejectives from affixes in Tojolabal, Slave, and Dakota. Other examples could be cited, and indeed, in the Optimality Theory framework Jakobson's observation is being proposed as a universal of language, expressed as an ordering priority of faithfulness in roots over faithfulness in affixes (McCarthy & Prince 1995: 116–117, Alderete 1999, Ussishkin 2000).

Hopper & Traugott (2003) discuss the proposed tendency in the context of grammaticization. They propose that as a result of qualitative reduction, the phonological segments in affixes “are drawn from a progressively shrinking set” (2003: 154), whose members are the universally unmarked segments. They note especially that affixes tend to contain “apical (tongue tip) consonants such as [n], [t], and [s], the glottal consonants [ʔ] and [h], and common vowels such as [a], [u], [i], and [ə]” (2003: 155). Two explanations are mentioned: one is the phonological reduction that accompanies the process of grammaticization results in unmarked segments; the other is the tendency for analogical spread of an allomorph to favor the allomorph with the highest text frequency which in turn favors unmarked segments.

The only systematic test of a related hypothesis that I know of appears in Willerman (1994), a study which examines the phonemes in pronouns versus the general phoneme inventory in thirty-two languages of a genetically stratified sample. The data from this study support the hypothesis as a tendency, though it is by no means applicable in all the languages of the sample. The method and results of the Willerman study will be compared to those of the current study in the sections below.

The present study seeks to test the hypothesis on verbal inflection in a twenty-three language stratified probability sample. More than testing the hypothesis, however, the goal is to find an explanation for the phenomenon to the extent that it can be verified in the languages of the world. Several possible explanations are available.

Willerman's explanation refers to the higher predictability of closed-class items, due both to their high frequency and their syntactic predictability. Presumably highly predictable items would not need to use as many phonemic contrasts as less predictable items. What is missing from this explanation is a mechanism by which the situation could come into being. Since pronouns and affixes derive diachronically from open class words (see Lehmann 1985 and other literature on grammaticization), which presumably utilize the full range of phonemes, a mechanism by which the number or phonetic quality of the phonemes in closed-class items is reduced due to predictability must also be proposed.

Explaining the nature of affixes requires reference to the process of grammaticization. As mentioned above, the tendency in question has an obvious and natural explanation within grammaticization theory, where a major phenomenon is the reduction of the phonological shape of grammaticizing words and phrases. In phonological reduction, secondary features, such as labialization, glottalization, etc. may be lost, producing what are regarded as less marked segments. It is important to note, however, that other common changes affecting consonants are voicing and sonorization, which do not necessarily produce consonants that are less marked in terms of the distribution of phonemes in the languages of the world, but would indeed restrict the number of phonemes occurring in affixes in a language. Thus the exact nature of the restrictions on phonemes in affixes will help to guide us to an appropriate statement of the hypothesis and a satisfactory explanation.

While grammaticization is the major force shaping affixes, two other factors may provide explanations for parts of the phenomenon. In discussing the absence of labio-velars from affixes in Indo-European, Dunkel (2001) proposes not that labio-velars were lost in affixes, but rather “that the labiovelars have COME INTO BEING AT ALL ONLY AMONG THE ROOTS” (Dunkel 2001: 10; emphasis in original). Thus a possibility to consider is that certain types of phonemes never develop in affixes. Another possible explanation for restrictions in affixes comes from language contact: phonemes that have entered the language through borrowed words are not likely to appear in affixes. For example, Inuit /p/, which is reported to occur only in borrowed words, also does not appear in affixes (Fortescue 1984). The materials used in the present study, however, did not provide evidence for the application of either of these factors other than the cases just cited.

Another source for restrictions on phonemes in affixes is proposed in the present study: it appears that languages with very restricted phoneme inventories in affixes create new affixes by combining other affixes, thus re-using the same phonemes in a number of affixes. We will see in Section 10.3 that this phenomenon accounts for restrictions in several languages of the sample used in the study.

Before exploring explanations, however, the tendency and its exact nature must be verified. It is important to bear in mind that some of the distributional facts evident in the languages of the world may be due to the statistical relation of affixes to lexical roots and unmarked to marked phonemes. First, the number of affixes in a language is generally very small in relation to the number of lexical roots. Second, affixes tend to be short (due to grammaticization) and therefore present fewer opportunities for the appearance of phonemes than roots, which in many languages include multisyllabic morphemes. Third, affixes may present limited opportunities for some phonemes to occur; for instance, if in a language unstressed syllables have only a subset of vowel con-

trasts and affixes are unstressed. Fourth, marked phonemes tend to have restricted distributions that make them less frequently used in the languages in which they occur (Greenberg 1966), decreasing the likelihood that they would show up among the small number of short morphemes constituting affixes. So to the extent that the tendency exists, it could very well be due simply to these factors. Thus it is important to test the universality of Jakobson's statement on a representative sample of the world's languages. As we shall see, the results of the statistical test on the languages of the sample show that the tendency is by no means universal. However, since a trend is discernible in some languages possible explanations should still be explored.

2. The method

2.1. Sampling

Often the evidence presented in the literature on the hypotheses to be tested here has been only anecdotal, in that researchers mention languages that have come to their attention that provide evidence in accord with the hypothesis, but they do not mention languages which do not conform to the hypothesis. Thus a more systematic test of the hypotheses is needed to ascertain the strength of the observed tendency or just how universal it is. As mentioned above, one such test of a related hypothesis has been undertaken. Willerman (1994) chose to test the hypothesis on the closed class of pronouns, presumably because almost all languages have pronouns and because this would constitute a small enough class of items to make the task manageable.

To test a crosslinguistic hypothesis, one must construct a systematic sample of the world's languages whose selection is relatively free of bias: bias from genetic relations, from areal relations, and potential bias introduced because languages are included that are known to support or not support the hypothesis. In choosing the number of languages to include in a sample, it is important to remember that we are testing a hypothesis and thus engaged in theory-building, not just looking for interesting facts (Perkins 1989). Thus we need a sample which is appropriate for doing statistics and such a sample must ensure the independence of cases. Thus each language in the sample must be independent of all others on the parameter examined. It is important to bear in mind that a sample could contain TOO MANY languages. That is, if two languages have a similar property because they are descended from a common ancestor, then they should not both be included in the sample. As Perkins (1989) has pointed out, the degree of independence of languages from their genetic sisters depends on what aspect of language is under examination.

Some properties change relatively quickly. For example, Perkins estimates that a good sample size for studying consonant types, to achieve genetic independence, would be between 84 and 91 languages. These are relatively high

numbers, reflecting the fact that consonant inventories can change relatively quickly. However, to avoid influence from areal contact, a smaller number of languages (between 59 and 84) should be used, since areal contact can affect consonant inventories. When it comes to basic word order, the numbers are much lower because word order changes more slowly. Also word order can be affected by areal contact. Dryer (1989) has demonstrated areas the size of continents affect word order distribution, which means that a sample of languages larger than five could possibly contain areal bias that makes the cases non-independent of one another. I mention these facts because of the popular assumption that more languages in a sample make the sample better. A larger number is better only if you are looking for interesting facts. To test a hypothesis, a sample must achieve a balance between being representative and containing only independent cases.

In the current case we have a set of hypotheses to test that involves the intersection of phonology (in particular phoneme inventories) and morphology (in particular inflectional affixes). As just mentioned, phoneme inventories (this is true of both vowels and consonants) can change relatively rapidly. However, the hypothesis also involves affixes, which are much more conservative. Since sister languages can have affixes from the same etymological source after several millennia of separation, care must be taken not to include languages in the sample that are related to such an extent that their affixes do not constitute independent cases.

The current study used a sample based on the GRAMCATS sample described in Bybee, Perkins, & Pagliuca (1994). This original sample contains 76 languages which were randomly selected within genetic groups. Thus while genetic relations were controlled for, there are still many related languages in the sample, for example, five Indo-European languages, six Sino-Tibetan languages, three Ural-Altaic languages, and so on.¹ The study for which this sample was constructed involved both periphrastic constructions and affixes. Periphrastic constructions can in general be assumed to be more recent in a language and thus less subject to inheritance in sister languages, but affixes would have a longer history and be more likely to be shared in sister languages. Thus there is a danger in using 76 languages in counting the same phenomenon more than once. For example, consider two Indo-European languages included in the original study, Baluchi from the Indo-Iranian branch and Modern Greek, occupying its own branch. The ancestors of these languages separated from one another in the neighborhood of 5,000 years ago (Gamkrelidze & Ivanov 1995) and yet an examination of their person/number suffixes shows that they have maintained some of the consonants present in their common ancestor as shown in Table 1.

1. Also, our post hoc analysis showed that areal contact did affect the results and that only forty-nine languages could be included in a sample that was relatively free of areal bias.

Table 1. *Person/number suffixes in Baluchi (Barker & Mengal 1969) and Modern Greek (Householder, Kazazis, & Koutsoudas 1964)*

	Baluchi	Modern Greek
1SG	-in	-ō
2SG	-əy	-s
3SG	-it	∅
1PL	-ən	-me
2PL	-yt	-te
3PL	-ənt	-oun

In particular, note the presence of a nasal consonant in 1PL, the /t/ in 2PL, and the /n/ in 3PL. We also know that the /t/ of 3SG as found in Baluchi is a common Indo-European trait, as is the /nt/ in 3PL. Thus despite 5,000 years of separation, these affixes do not constitute independent developments.² This strongly implies that only a small number of languages can be included in the sample to test a hypothesis about affixes, and that no more than one language from a group at the level of Indo-European should be included.

In the Voegelin & Voegelin classification of the languages of the world, the Indo-European family is considered a phylum. There are twenty-four phyla listed by Voegelin & Voegelin (1978) and this level of classification was important in the construction of the GRAMCATS sample. While the GRAMCATS sample contained 76 languages and thus several from each phylum, the current study, in order to avoid the bias just described, used one randomly chosen language from each phylum. In addition, two other languages were used. In the GRAMCATS sample an effort was made not to overrepresent language isolates and very small groups (groups with fewer than six members), so these were put together and two languages were selected from this group. Also, since Voegelin & Voegelin list pidgins and creoles under the affiliation of the lexifier language (e.g., Tok Pisin is listed as a Germanic language), all known pidgin and creole languages were put into a single group with one selection made from that group. These two groups plus the twenty-four phyla resulted in twenty-six groups. For the current study one language was selected from each group. Since the current hypothesis is relevant only for languages that have inflectional affixes, three languages were not used in the study: Cantonese and Palaung, which have no inflectional affixes, and Tok Pisin, which has only two (Mühlhäusler 1985, Woolford 1979). Thus this study is based on twenty-three

2. The finite verb forms of Baluchi appear to have been formed using the inflected forms of the copula. The person/number forms of the copula inflection thus appear on all verbs. Despite this recycling, the original source of the person/number forms of the copula appear to be descended from the person/number forms of Indo-European.

maximally unrelated languages, listed in Table 2 (below, Section 4). Access to the languages of the sample was through descriptive reference grammars.

2.2. *Coding*

The GRAMCATS sample was used because the selection of languages was quite independent of the current hypothesis, the selection was principled, and the materials had already been assembled. In addition, the coding of affixes and phoneme inventories had already been completed in the context of the earlier study, so a very time-consuming part of the task had already been completed. For the current study, it was necessary to examine all the affixes and tabulate the phonemes used in them. As this study is based on the same data used in Bybee et al. (1994), it includes only inflectional affixes that occur on verbs. These data provide a useful testing ground for the hypothesis because verbs have more inflection than nouns or adjectives and indeed, many languages have inflection on verbs, but none on nouns. The use of only verbal affixes does bias the test towards confirmation of the hypothesis. A language that appears to conform to the hypothesis might not if noun affixes were also considered, since these could contain phonemes not found in the verbs. Thus including nouns could only weaken the fit of the data to the hypothesis; it could not strengthen it. As we shall see, the data from the sample show that patterning as predicted by the hypothesis is only a tendency and not a universal, and this conclusion would not be changed by the inclusion of nominal affixes in the data.

Verbal affixes had been previously coded according to the criteria used in Bybee et al. (1994: 37–40). Affixes formed by reduplication were not included in the study as they contain no fixed phonemic representation. Also, the inclusion of reduplication might greatly expand the inventory of phonemes used in affixes. Presumably the hypothesis was not intended to apply to reduplication. The decision of whether to consider a grammatical morpheme as an affix rather than an auxiliary or particle is based largely on the descriptive tradition of the language as reported in the reference grammar used. An attempt was made to restrict the study to inflectional affixes by coding only those affixes that are lexically general; however, as the line between inflection and derivation is a rather fuzzy one, some highly general derivational affixes may be included in the study. Affixes that form part of a periphrastic construction were also included. Constructions formed with auxiliary verbs often require a certain affix on the main verb, as in the English Progressive or Present Perfect. Thus Baluchi *-əg*, an infinitive marker, is included because it is used in several periphrastic constructions with the verb. However, despite occurring in various constructions, it was only counted once in the current study in the tally for number of affixes in the language. On the other hand, affixes with the same phonological shape but different meanings, such as the /z/ plural and possessive on nouns in

English, were counted as different affixes. (A list of the affixes coded is found in Appendix 2.)

For each morpheme coded in Bybee et al. (1994), a principal allomorph was selected and it is the phonemic composition of this allomorph that is used in this study to test the hypothesis. The principal allomorph was chosen by applying the following criteria: first, the longest allomorph was chosen; however, if the longest allomorph was clearly irregular or marginal, or if all the allomorphs were the same length, then Bybee et al. (1994) coded the allomorph that the author indicated was the most basic, the most frequent, or the most important in some way. If none of the criteria yielded a unique choice, then the allomorph that the author mentioned first was coded. This practical procedure affects the hypothesis only in cases of suppletion, where a suppletive allomorph might have a different set of phonemes from the principal allomorph. However, suppletive expression was extremely rare in the corpus studied and is thus not likely to have an effect on the results reported here.

The phonemes for each language were determined on the basis of the reference grammars used, all of which contain a list of phonemes for the language with their phonetic descriptions. (The phonemes for each language are listed in Appendix 1.)

3. The hypotheses

The hypothesis as stated by Jakobson could be interpreted in various ways. In order to be testable, however, a hypothesis must be quite rigorously stated. In this section I present four possible formulations of the hypothesis. The phenomenon under investigation actually involves two questions which are independent of one another. The first concerns the number of phonemes used in affixes and whether or not this number is smaller than the number that would be expected by chance, given the number of phonemes and the number of affixes in a language and their structure. The hypothesis can be stated as follows:

Hypothesis 1: The number of distinct phonemes used in the inflectional verbal affixes of a language is smaller than the number that could be expected to occur by chance.

The test of this hypothesis will be presented in Section 4.

Several hypotheses could be formulated concerning the nature of the phonemes which are excluded from affixes. Note that these hypotheses are independent from Hypothesis 1 because even if a language were to use more phonemes in its affixes than would be predicted by chance, it could still exclude certain phonemes and these could form a systematic set of some kind. The prediction made in the literature, particularly within Optimality Theory is that the exclusions are systematic and that it will be the more marked consonants that are

excluded from affixes. However, it might be the case that markedness as formulated by Jakobson or in generative phonology (e.g., Chomsky & Halle 1968) is not exactly the right concept and a notion of complexity such as that formulated by Willerman might provide a better fit the data. Thus in the discussion in Sections 6 through 8, we will compare the predictions of Hypothesis 2 and 3:

Hypothesis 2: The phonemes that occur in affixes tend to be the less marked segments of the phoneme inventory.

Hypothesis 3: The phonemes that occur in affixes tend to be the less complex segments of the phoneme inventory.

The prediction made by grammaticization theory would be that the phonemes constituting affixes would be more reduced phonologically and thus less complex, but not necessarily less marked.³ While reduction of secondary articulations, loss of non-pulmonic airstream mechanisms, and extreme articulations might produce unmarked sounds, voicing, sonorization, assimilation, reduction to glottal stop and [h], and vowel centralization, which also occur in reduction might produce more marked sounds. Thus we might predict that the phonemes in affixes are in general the products of well-known reduction processes. A complication in applying this theory is that certain segment types, such as low vowels or nasal consonants, might be more stable in the face of the overall reduction that occurs in grammaticization. These segment types might occur more in affixes because they are less likely to reduce. Thus it is not possible to state a single hypothesis from grammaticization theory, though Hypothesis 3 would be one prediction made by grammaticization.

A final, weaker, hypothesis might provide a better fit with the data. This hypothesis makes no claims about markedness or reduction and has no particular explanation associated with it, but merely looks for systematicity in the set of phonemes that is avoided in affixes.

Hypothesis 4: The phonemes absent from affixes form systematic sets.

If this hypothesis turned out to be confirmed by the data, then an explanation for the particular excluded sets would need to be formulated on a case-by-case basis.

These hypotheses are discussed and tested against the sample languages in Sections 4–9.

3. The use of the term “grammaticization theory” here is appropriate for while grammaticization is a well-documented process, the postulation that all affixes derive from grammaticization by phonological reduction of lexical forms is a theory.

4. The results of the test: Number of phonemes

In order to find out if the number of phonemes occurring in affixes is the result of chance rather than the result of systematic processes of some kind, it was necessary to have a way of calculating the chance factor. This calculation was applied to both consonant phonemes and vowel phonemes (see Section 9), but prosodic features were not considered at all. First we address consonants. The question is the following: given the number of consonant phonemes a language has, and the number of positions occupied by consonants in affixes, what is the most likely number of consonant phonemes to appear in affixes? A formula was developed to calculate the probabilities of each number of consonants appearing in affixes for a language and the number with the highest probability would be the number we would expect the language to have if no other factors intervened. It is important to note that the test is based on the assumption that all the consonants in a language have an equal probability of occurring in any consonant slot. Of course, this is not true: phonotactic constraints will limit consonant occurrence and as mentioned above, more marked or complex consonants being less frequent than simpler or unmarked consonants, would be less likely to occur in affixes. Unfortunately there was no systematic way to incorporate these variables into a general formula equally applicable to all languages.

The following presents the formula used to determine the number of consonant phonemes to appear in affixes in a language based on chance, given the number of consonant phonemes in the language and the number of positions in affixes using consonants. The following abbreviations will be used:

N = the number of consonant phonemes in the language

S = the number of positions in affixes using consonants in the language

x = the number of distinct consonants actually used in the affixes

S was determined by counting the number of consonants used in each affix in the language. Then the probability of observing x distinct consonants (assuming each consonant is equally likely for each of the affixes) is:

$$(1) \quad \frac{C(N, x) \cdot C(S - 1, x - 1)}{C(N + S - 1, N - 1)}$$

$C(N, x)$ means “ N choose x ” where for any positive integers N and x with $N \geq x$,

$$(2) \quad C(N, x) = \frac{N!}{x!(N - x)!}$$

So taking an example where ($N = 38$, $S = 17$, $x = 9$):

$$(3) \quad \frac{C(38, 9) \cdot C(16, 8)}{C(54, 37)} = .044492$$

The highest probability occurs at $x = 12$, with a probability of .2508. Thus the example language has three fewer phonemes in affixes than would be expected by chance. This deviation is significant at the level of $p < .05$.

Table 2 gives the numbers used in the calculations and the number of consonants expected to occur by chance in affixes. The first column lists the languages and the references used for the phonological and morphological information. The second column shows the number of consonant phonemes reported for the language. The third column shows the count of the number of slots within affixes that are occupied by consonants. For example, an affix with the shape CV has one consonant slot, while an affix with the shape CVC has two, and so on. The fourth column shows the results of the application of the formula: it lists the number of consonant phonemes that would be most expected by chance, given the number of consonant phonemes and the number of consonant slots. The final column lists the actual number of consonant phonemes that appear in affixes. The comparison of the last two columns constitutes the test of the hypothesis. These results are shown in Table 3, along with probabilities as provided by the formula. The phoneme inventories of all the languages and the list of affixes used in the study are given in the appendices.

The pattern displayed in Table 3 shows that seventeen of the twenty-three languages have the number of consonants in affixes that is not significantly different from the number predicted by the formula to occur by chance. Six languages (marked by an asterisk) have numbers of consonants in affixes significantly different from the number predicted by chance, five of these are in the predicted direction and one is in the opposite direction. Thus only five languages out of twenty-three fit the predictions of the hypothesis. However, there are more languages with fewer consonants in affixes than predicted than there are languages with more consonants in affixes than predicted by chance. We can also note that the probabilities of the number of consonants in affixes in Slave and Inuit come close to significance. Thus the phenomenon in question may be said to represent a tendency, albeit a rather weak one.

These results confirm that the putative tendency for languages to use a smaller number of phonemes in affixes than in stems could be the result of mere chance distribution and thus not a universal of language. Indeed, further analysis of the actual consonants used confirms a chance distribution for all but six of the languages, further eroding the possibility of a universal of language behind the data. However, in some of the languages of the sample and in some additional languages to be discussed below, some patterns of exclusion are discernible. Moreover, we know that the determination of which phonemes are used in affixes is not entirely a matter of chance since affixes are the result of a long process of evolution from lexical items and not created by sprinkling phonemes across affixes in the way that the formula suggests. Thus a closer examination of the patterns found in particular languages will help to complete the picture.

Table 2. *Numbers used in calculation and expected numbers of affixes, by language. N = number of consonant phonemes; S = number of consonant slots.*

Language name	<i>N</i>	<i>S</i>	Number of consonants where $Pr(X = x)$ is greatest	Number of consonants used in Affixes
Inuit (Fortescue 1980, 1984)	16	221	15	13
Margi (Hoffmann 1963)	38	24	15	14
Cheyenne (Leman 1980)	11	135	11	9
Tucano (Sorensen 1969, West 1980)	28	28	14, 15	14
Gugu-Yalanji (Patz 1982, Hershberger 1964)	13	25	9	12
Pangasinan (Benton 1971)	15	29	10	6
Tohono O'odham (Hale 1959; Mason 1950; Mathiot 1973, 1978; Zepeda 1983)	19	34	12	12
Abkhaz (Hewitt 1979)	64	125	39, 40	42
Guaymí (Alphonse 1956; Kopesec 1975; Young 1986, 1987)	15	8	5, 6	6
Kui (Subrahmanyam 1971, Winfield 1928)	19	18	9, 10	8
Abipon (Najlis 1966)	15	39	12	11
Karok (Bright 1957)	16	60	14	13
Baluchi (Barker & Mengal 1969)	25	18	10	5
Yagaría (Renck 1975, 1977)	15	53	12	11
!Kung (Snyman 1970)	97	12	12	9
Slave (Rice 1989)	34	34	17	14
Engenni (Thomas 1978)	19	3	3	3
Kanuri (Hutchison 1981, Lukas 1937)	25	26	13	9
Palantla Chinantec (Merrifield 1963, 1968)	19	24	11	11
Shuswap (Kuipers 1974)	36	42	19	14
Tojolabal (Furbee-Losee 1976)	23	27	11, 12	9
Dakota (Boas & Deloria 1941; Buechel 1939, 1983; University of Colorado Lakshota Project 1974, 1976a, b, c)	27	19	11	11
Buriat (Poppe 1960)	17	64	13	15

We turn now to a discussion of the random and patterned exclusions.

5. The results of the test: Patterned exclusions

A patterned exclusion would be any natural class of sounds that is totally excluded from affixes. This class could be defined by manner of articulation, place of articulation, voicing, or airstream mechanism. Of the languages listed above the majority do not have full natural classes excluded from affixes. Let

Table 3. Languages listed according to the number of phonemes used in affixes compared to the number expected by chance: Under “0” are the languages with the same number of phonemes as predicted by chance; “+1” indicates one phoneme more than predicted by chance and “-1” etc. indicates one fewer than predicted by chance.

-6	Baluchi .0004*, Shuswap .01*
-5	
-4	Pangasinan .006*, Kanuri .02*, Tojolabal .03*
-3	Slave .06
-2	Inuit .06, Cheyenne .14, !Kung .09
-1	Yagaria .22, Margi .19, Abkhaz .13, Kui .16
0	Abipon .27, Tohono O’odham .23, Guaymí .33, Palantla Chinantec .24, Karok .28, Engenni .73, Dakota .23, Tucano .21
+1	Buriat .19
+2	
+3	Gugu-Yalanji .03*

us consider the languages in the following groupings:

- (i) Engenni has only three affixes and uses three different consonants in them. This language is probably neutral with respect to the hypothesis.
- (ii) Gugu-Yalanji uses all but one of its consonants in affixes; however, the one excluded phoneme /j/ occurs as an allomorph of an important verb class marker. Four others use all but a few consonants in affixes: Inuit excludes /ʃ/ and /t/ which occur in only one dialect in any case, /h/, which occurs only in loan words and injections, and /p/, though geminate /p/ does occur in affixes. Karok uses all but three consonants, /s/, /x/, and /j/. One cannot conclude that Karok excludes voiceless fricatives, because /f/, /θ/, and /ʃ/ are used in affixes. Buriat uses all consonants except /z/ and /w/ and Cheyenne uses all consonants except /p/ and /k/.
- (iii) Twelve languages exclude more consonants, but the exclusions appear random. The languages are Abipon, Abkhaz, Baluchi, Guaymí, Kanuri, Karok, Kui, !Kung, Palantla Chinantec, Pangasinan, Tohono O’odham, and Yagaria.

As to (iii), consider for example, the consonant inventories of Pangasinan and Tohono O’odham as listed in Tables 4 and 5 respectively, with the consonants not used in affixes in parentheses. (See Appendix 1 for a full listing of consonant and vowel phonemes used and not used in affixes in the sample languages.) These two examples show what is meant by a random distribution: no class of consonants based on a particular manner or place of articulation or voicing is excluded unless one were to regard the absence of liquids and glides in Pangasinan as a patterned exclusion.

Table 4. *Consonant inventory of Pangasinan, with consonants not used in affixes in parentheses*

p	(t̚)	k
(b)	(d̚)	g
s		(h)
m	n	(ŋ)
(w)	(j)	
	(r), (l)	

Table 5. *Consonant inventory of Tohono O'odham, with consonants not used in affixes in parentheses*

(p)	t		k	ʔ
b	d	ɖ	(g)	
(s)	ʃ		h	
	(tʃ)			
	(ʒ)			
m	n	ɲ		
(w)	r (l)			

Another example that suggest randomness is !Kung, which has eight affixes that met the coding criteria and these use nine consonants, four of which are clicks.

Taking groups (ii) and (iii) listed above together and excluding Engenni, sixteen of the twenty-two languages have random exclusions from affixes. Group (iv) contrasts with these.

- (iv) Six languages have exclusions that constitute natural classes of consonants:
- Margi has no voiceless stops, no fricatives, no implosives, and no palatalized velars in affixes;
 - Tucano has no “voiced aspirates” (which are probably ejectives) in affixes;
 - Shuswap has no ejectives in affixes;
 - Tojolabal has no voiced stops, ejectives, or liquids in affixes;
 - Slave has no labialized consonants in affixes and only one out of five ejectives;
 - Dakota has no coronal stops, no fricatives, and only one out of four ejectives.

It can be seen that consonants with obvious complexity such as glottalization figure heavily in this list, with implosives and ejectives being more frequently excluded than any other class, and labialized and palatalized consonants also

being excluded. Further generalizations are difficult to make since Margi excludes voiceless stops, but Tojolabal excludes voiced ones. Both Margi and Dakota exclude fricatives.

The most striking generalization seems to concern ejectives but lest it be concluded that the exclusion of ejectives from affixes is a universal, we must note the case of Abkhaz in which seven out of fourteen ejective stops occur in affixes.

A mention of Indo-European is relevant here since the hypothesis about restrictions of phonemes in affixes has been mentioned frequently with respect to Indo-European. Dunkel (2001) addresses precisely this issue, concluding that Proto-Indo-European lacked labio-velars and voiced stops in affixes. The lack of velars with secondary labialization in affixes fits with the types of restrictions listed above, and while the absence of voiced stops occurs in Tojolabal, it is not a commonly observed restriction. However, note that Dunkel presupposes the traditional reconstruction of the Proto-Indo-European obstruent system as consisting of a series of voiceless stops, voiced stops, and voiced aspirated stops. An alternate reconstruction proposed in Gamkrelidze & Ivanov (1973) and Hopper (1973) has voiceless stops, glottalized stops, and voiced stops. In this reconstruction, the series absent from affixes would be the glottalized stops, a more common restriction crosslinguistically, as we have just seen.

Thus only a small number of languages have patterned exclusions (six out of twenty-three) and among these there are no universal exclusions. The expected favoring of apical or coronal consonants in affixes failed to emerge in this sample (see Section 8).

6. Marked or complex consonants

The popular consensus on the phonemes excluded from affixes is that they are the more marked or more phonetically complex phonemes existing in the language. Of course the difficulty with testing this hypothesis is that markedness and complexity are often not well defined in the literature. The only serious attempt to define the notion of complexity with respect to this phenomenon is found in Willerman (1994). She uses a biomechanical model to construct a scale of articulatory complexity for the variables used in producing both consonants and vowels. In her scheme, some variables are considered basic while others are penalized for creating a more difficult articulation. A brief review of her categorization is given here, with mention only of those features that figured in the present study. For further details and motivation, the reader should consult Willerman's work directly.

Before proceeding, however, let me make a brief remark about the appropriateness of the articulatory view of complexity that Willerman uses and that

I shall also adopt here. Of course there is perceptual complexity that could be taken into account, but the reason for choosing articulatory variables is that affixes are formed by the reduction of phonological material in the process of grammaticization. An important factor in this reduction is the increase in frequency of use that accompanies grammaticization (Bybee 2003). The fact that frequency plays a role in phonetically reductive processes points strongly to articulatory or production mechanisms in change (Bybee 2001). Frequently repeated neuromotor sequences are prone to reorganization into single units and thus to overlap and reduction of movement such as is characteristic of the reduction in grammaticization. Frequent perceptual stimuli in contrast should be less prone to change, as they are frequently present in the input and should be easier to learn and more likely to be maintained unchanged. For this reason, I take articulatory measures of complexity to be appropriate for the present study.

In Willerman's tabulation of complexity certain articulatory states and gestures were taken to be basic, while others were taken to add complexity, and thus in her terminology "penalized". The following places of articulation were considered basic: bilabial, labio-dental, dental, alveolar, palatal, velar, and glottal. Penalized variables are: palato-alveolar, retroflex, uvular, pharyngeal, and labio-velar. Secondary articulations of labialization, palatalization, etc. were also penalized.

While voiced and voiceless glottal sources were considered basic, aspiration, laryngealization, breathy, etc. are penalized as creating more complexity.

The basic manners of articulation were: plosive, fricative, r-sound, tap, flap, approximant, and nasal. The penalized ones were: implosive, ejective, click, affricate and trill. Other penalized manner features were nasalization, nasal release, lateral release, and length.

Willerman then categorized phonemes as "basic" if they used no penalized variables (e.g., a voiceless bilabial stop), "elaborated" if they used one penalized feature (a bilabial ejective), and "complex" if they used more than one (a uvular ejective). Willerman found that when all the phonemes from the thirty-two languages of her study were pooled, the pronoun paradigms used a much higher percentage of basic segments and a lower percentage of complex or elaborated segments than their percentage in the general inventory would predict. Table 6 summarizes the data found in Willerman's Tables 8.3 and 8.4. The difference between the distribution of consonant types in the general inventory and in pronouns is highly significant. She also found that a similar test on vowels yielded highly significant results as did the test on consonants and vowels combined.

A study of each language individually, however, yielded more mixed results. A comparison as in Table 6 for all segments for each language found that in fourteen out of thirty-two languages there was no significant difference

Table 6. *The number of consonants of each type in the general inventory of the thirty-two languages compared to the number of consonants of each type used in pronouns in the same languages (data from Willerman 1994)*

	Inventory	Percent	Pronouns	Percent
Basic	343	29.1	132	61.7
Elaborated	468	39.5	63	29.4
Complex	373	31.5	19	9.0

($p > .05$) between the distribution of segment types in the general inventory and in pronouns. For eighteen languages, this difference was significant at the .05 level or better. Of course, some of this correspondence could be attributed to the fact that complex segments tend to be of lower frequency and more restricted in distribution than the basic ones. Thus Willerman's results on pronouns show that what appears in broad perspective to be a general trend, is actually only found in small majority of languages, with a large minority showing no such trend.

Using the GRAMCATS sample, I counted consonant types for nineteen of the languages used in the study. Not counted here were Inuit, where the three phonemes excluded from affixes are all dialectal or used only in loanwords, Gugu-Yalanji, in which all but one phoneme occur in affixes, Engenni, with only three affixes, and !Kung. !Kung was not counted because with 94 consonant phonemes and only eight affixes the excluded consonants would overpower the rest of the sample. The results for these nineteen languages are shown in Table 7. The "elaborated" categories contain stops, fricatives, and affricates with a secondary articulation or a special laryngeal state or both. These categories correspond to Willerman's "elaborated" and "complex" categories together.

The findings of Table 7 may be summarized by noting that the following consonant types are favored in affixes, with 60% or more of these types occurring in affixes: voiceless stops (65%), voiced stops (63%), nasal stops (72%), /h/ (60%), glottal stop (73%).

The following types seem to be dispreferred in affixes: voiced fricatives (37%), affricates (27%), elaborated stops (31%), elaborated fricatives, and affricates (38%).

The following types show no distinctive distribution between affixes and lexical stems: voiceless fricatives (46%), liquids (44%), glides (53%).

The low occurrence of elaborated obstruents confirms the findings of Willerman and reflects the discussion above, in which it was noted that ejectives, labialized and palatalized consonants are sometimes systematically excluded from affixes.

Table 7. *Number of consonant phonemes included in affixes or excluded from them in nineteen languages from the GRAMCATS sample*

		Included		Excluded	
		N	%	N	%
Simple stops	voiceless	35	65	19	35
	voiced	27	63	16	37
Fricatives	voiceless	24	46	28	54
	voiced	11	37	19	63
Affricates		6	27	16	73
Elaborated stops		17	31	38	69
Elab. fricatives/affricates		14	38	23	62
Nasal stops		42	74	15	26
Liquids		18	44	23	56
Glides		21	53	19	47
/h/		9	60	6	40
Glottal stop		8	73	3	27

7. Markedness or complexity

In the theory of markedness as originally formulated by Jakobson the absence of a feature was considered unmarked and its presence marked. Thus in consonants, voicing and nasality are both marked, as well as palatalization, rounding, and so on. Jakobson sought a correspondence between unmarked values, frequency of occurrence in the language of the world, early child language acquisition as well as loss in aphasia. Unfortunately for our purposes, Jakobson only formulated markedness values for a small number of features. A more comprehensive formulation of markedness values is found in Chomsky & Halle (1968: Chapter 9). This formulation seems to be based on intuitions of “natural” or “expected” values, which presumably results from the distribution of consonants and vowels in the languages of the world. In other words, lacking better criteria, it appears that in this work, unmarked simply means “often found in phoneme inventories”. Despite the lack of firm grounding, it is the only explicit formulation that applies to a wide range of phonological features, so it is the one to which I am referring here when markedness is discussed.

The difference between markedness theory and complexity as defined by Willerman is that markedness relations pervade the inventory, so that some consonants that are on Willerman’s analysis “simple” could have some marked values. Apparently, however, a low degree of markedness, such as that of nasal consonants, does not exclude consonants from affixes. Only a high degree, such

as found in Willerman's "elaborated" and "complex" categories, describes the excluded consonants.

While the theories are slightly different, neither theory offers a complete explanation for the tendencies found in the sample examined here. For instance, while voiceless stops are less marked than nasal consonants, the latter appear more often in the affixes of the sample. Both are relatively low in markedness, compared to the glottal stop and fricative, which according to Chomsky & Halle (1968) – who consider them glides – are more marked than obstruents, and in the class of glides are the most marked. Thus according to markedness theory, the glottal stop and fricative should not occur frequently in affixes. On the other hand, in Willerman's theory the glottal stop and fricative are considered "simple". Also, while elaborated stops, fricatives, and affricates are predicted to be excluded by both theories, neither theory provides an explanation for the low percentage of voiced fricatives in affixes.⁴ Even the theory that affixes contain consonants that have undergone reduction explains only part of the data, as we will see in Section 10.2.

8. Places of articulation

Patterned exclusions by place of articulation are rare if both obstruents and nasal consonants are taken into account. The preference for nasal consonants in affixes means that many languages exclude a certain place of articulation for obstruents, but have no such exclusion for nasals. Thus Margi, Kui, Cheyenne, Slave, and Palantla Chinantec do not use labial obstruents in affixes, but they all allow /m/. Only Tojolabal excludes all labials. Labial obstruents are the most often excluded though other exclusions occur in a few languages. Dakota coronal obstruents do not occur in affixes, but /l/ and /n/ do. Palantla Chinantec /t/ and /d/ do not occur in affixes, but the corresponding affricates do, and so does /n/. Pangasinan excludes dental stops but allows /n/ and /s/. Abipon and Kanuri exclude alveo-palatal affricates. Cheyenne and Guaymí exclude velar consonants and Shuswap excludes uvulars. Thus almost every common place of articulation can be singled out for exclusion, but the most common exclusion involves labials, particularly obstruents. Except for this fact, there is no particular evidence for the favoring of coronal consonants in affixes as one might expect from the evidence from English and other Indo-European languages (Hopper & Traugott 2003).

4. This is not to say that an explanation is not available: voicing is difficult to maintain in fricatives as the stricture diminishes the airflow through the glottis.

Table 8. *Number of vowel phonemes (N), number of vowel slots in affixes (S), number of vowel phonemes used in affixes and the probability of that number occurring by chance. Probabilities of less than .05 are marked with an asterisk*

Language name	<i>N</i>	<i>S</i>	Number of vowels used in affixes	Probability of that number of vowels
Inuit	3	159	3	0.963
Margi	6	26	4	0.203
Cheyenne	3	114	3	0.949
Tucano	6	32	6	0.390
Gugu-Yalanji	3	21	3	0.751
Pangasinan	5	27	2	0.008*
Tohono O'odham	5	17	4	0.468
Abkhaz	2	99	2	0.980
Guaymif	8	10	3	0.104
Kui	10	28	4	0.005*
Abipon	5	52	4	0.283
Karok	8	56	6	0.176
Baluchi	8	19	5	0.261
Yagaria	9	53	7	0.249
!Kung	7	12	4	0.311
Slave	11	32	6	0.053
Engenni	10	4	2	0.189
Kanuri	7	27	6	0.416
Palantla Chinantec	7	23	4	0.113
Shuswap	6	22	4	0.247
Tojolabal	5	20	5	0.365
Dakota	8	19	5	0.261
Buriat	12	50	6	0.004*

9. Vowels

The examination of the number of vowel phonemes used in affixes yields weaker results than the same test for consonants. While only six of the twenty-four languages use all their vowels in affixes, of the eighteen that do not use all, only three show a significant deviation from chance in the number of vowel phonemes used. These three languages are Pangasinan, Kui, and Buriat. In addition, Slave comes close to reaching significance. See Table 8.

The distribution of vowels in languages which use only a small number of vowels is in some ways typical of the general trends: Pangasinan uses only /i/ and /a/ in affixes and excludes /ɛ/, /u/, and /u/; Kui excludes long vowels and /o/ but allows /e/, /i/, and /u/; Buriat excludes long and short /u/, /ɛ/, and /ɔ/ but

allows long and short /i/, /a/, and long /u/; Slave excludes /e/, /o/, the diphthong /ie/, and all nasal vowels except /ɛ/.

One universally applicable statement can be made about vowels with respect to the sample used: all of the languages have /a/ and /i/, that is, a low vowel and a high front vowel, and all but one of the languages (Engenni) use these two phonemes in their affixes. (The only qualifying comment needed here concerns the vowel phonemes of Abkhaz: according to Hewitt (1979), Abkhaz has only two vowel phonemes, an open vowel /a/ and a close vowel, which Hewitt writes as /ɔ/; principal allophones of this vowel are [i] and [u].) Thus there seems to be a slight dispreference for the mid vowels (/o/, /ɔ/, /e/, and /ɛ/): /e/ is excluded in Margi, Baluchi, and Buriat; /ɛ/ is excluded in Pangasinan, and Palantla Chinantec; /o/ is excluded in Margi, Kui, Dakota, Guaymí, and Shuswap; /ɔ/ is excluded in Buriat. (In none of these languages do /e/ and /ɛ/ contrast, nor do /o/ and /ɔ/ contrast.) An obvious point to raise about these exclusions is that they leave the more peripheral vowels in affixes, a result that could be explained as favored for perceptual reasons.

There also seems to be a dispreference for rounded vowels in affixes when the phonemes of all of the languages were pooled: twenty-nine out of forty-six excluded vowel phonemes were round. Four languages excluded /u/ (Pangasinan, Baluchi, Dakota, and Guaymí), and six /o/ or /ɔ/ (see above). In markedness theory, roundness is unmarked for back vowels and marked for front vowels, so the general dispreference for round vowels is not fully explained by markedness theory. It could, however, be explained with reference to the loss of rounding as a result of reduction that is commonly found in affixes.

Pursuing the idea that reduced vowels may be used more in affixes, the six languages that have a mid centralized vowel (schwa) all use it in affixes, except for Engenni, which has only three affixes that met the coding criteria.

The languages with long vowels – Inuit, Cheyenne, Karok, and Buriat – all used them in affixes, with a few exclusions: Inuit and Karok do not use /u:/ and Buriat has several exclusions.

10. Explanations

Before examining explanations, it is important to bear in mind that the phenomenon to be explained has a limited applicability. Most of languages of the sample do not show fewer consonant or vowel phonemes used in affixes than would be expected by chance. Recall in addition that this test assumed that all phonemes have an equal chance of occurring, which we know is not true.

When it comes to which phonemes are used in affixes, only six languages exhibit patterned exclusions among consonants. Thus only a little more than a quarter of the languages sampled had patterned exclusions. However, when all consonants were pooled across languages, it does appear that consonants with

secondary articulations and laryngeal states or airstream mechanisms other than egressive pulmonic are less likely to appear in affixes. As might be expected, while some basic, complex, and elaborated consonants are used in affixes, there are no languages in the sample in which ONLY complex or elaborated consonants are used in affixes to the exclusion of basic or plain consonants. Similarly for vowels: while some few long vowels and nasalized vowels are used in affixes, there are no languages in the sample that use only long or nasalized vowels in affixes. There are two reasons for this asymmetry. First, complex and elaborated segments are probably diachronically derivative and have arisen only in specific contexts, which makes them less frequent and therefore less prone to occur in affixes. Second, affixes are diachronically produced by reductive phonological process which would have a tendency to reduce complex or elaborated segments to basic ones. However, the latter process is only a tendency and in some languages grammaticizing elements become affixes before undergoing a great deal of phonological reduction. In addition, the phonological reduction in affixes is ongoing at all times, so that particular affixes may be at different stages of reduction. Thus there is no absolute prohibition against complex and elaborated segments in affixes.

While it appears that many of the restrictions on segments in affixes can be explained via the grammaticization process (see Section 10.2), other avenues of explanation are also available. As we shall see in Section 10.3, languages with highly restricted inventories in affixes may be building new affixes out of old ones and thus re-using a small set of phonemes. But before turning to that phenomenon and its explanatory potential, we examine the specific case of the rarity of ejectives and implosives in affixes.

10.1. *Glottalized consonants*

Since the strongest tendency detected in the data for patterned exclusions was the absence of glottalized consonants (both ejectives and implosives) in affixes in the languages of the sample, I examined the other languages in GRAMCATS that have glottalized consonants to see what proportion of them excluded these elaborated consonants from affixes. The additional languages are Kanakuru (Newman 1974), Tigre (Leslau 1945, Raz 1983), Koho (Nguyen 1973), Krongo (Reh 1985), Maidu (Shipley 1964), Bari (Spagnolo 1933), and Ngambai (Vandame 1963). (See the list of phonemes in the supplemental sample in Appendix 1.)

- | | | |
|-----|----------|---|
| (4) | Kanakuru | implosive (/ɓ/, /ɗ/, and /ɓʷ/), labialized and prenasalized excluded |
| | Tigre | ejectives excluded (but only /m/, /n/, /k/, /j/, /w/, and /ʔ/ are used) |

Koho	(only 5 affixes use 4 consonants) aspirated and implosives excluded
Krongo	(only 9 consonants used in affixes: /t/, /c/, /k/, /m/, /n/, /nj/, /ŋ/, /w/, /j/) implosives and prenasalized excluded
Maidu	uses implosives and ejectives in affixes
Bari	(only 6 consonants used in affixes: /t/, /d/, /n/, /r/, /j/, and /k/) implosives not used
Ngambai	implosives (/b/ and /d/) excluded from affixes

Thus out of seventy-six languages, fifteen have ejectives or implosives and nine of these exclude them from affixes. When the formula presented in Section 4 is applied to these languages (see Table 9) three of the seven languages have significantly fewer consonant phonemes used in affixes than predicted. This suggests that the exclusion of glottalized consonants is an important factor in the tendency examined here.

In explaining the dispreference for glottalized consonants in affixes, it is important to bear in mind that Greenberg (1966) found that ejectives and implosives occur very infrequently in the languages where they occur. While this factor may explain the data, there are other factors to consider.

Table 9. *The formula applied to consonants (C) and vowels (V) in the supplemental sample*

Language	Number of phonemes in the language	Number of slots in affixes	Actual number of different phonemes used in affixes	Probability of observing this number of phonemes	Statistical significance
Maidu	C	19	62	14	0.215
	V	7	46	7	0.400
Bari	C	20	14	7	0.162
	V	9	11	3	0.050
Kanakuru	C	27	33	10	0.006 significant
	V	11	31	6	0.059
Tigre	C	27	38	6	0.000 significant
	V	6	27	6	0.327
Koho	C	30	5	4	0.394
	V	9	4	3	0.509
Krongo	C	26	30	9	0.004 significant
	V	13	32	7	0.060
Ngambay	C	18	10	6	0.277
	V	10	7	2	0.024 significant

A strong possibility for the explanation of the rarity of consonants with glottalic air in affixes is the well-supported theory that as lexical items grammaticize into affixes, the phonological material in them undergoes reduction (Lehmann 1982, Heine & Reh 1984, Bybee et al. 1994). A common type of reduction diminishes the magnitude of articulatory gestures and can also lead to the loss of gestures (Browman & Goldstein 1992, Mowrey & Pagliuca 1995). As Willerman has argued in this context, the production of a consonant using glottalic air requires much more muscular activity than the simple pulmonic egressive airstream of a plain consonant.

Though the evidence is fragmentary, implosives and ejectives appear to reduce in different ways. In Margi, a reduction of the auxiliary /gheda/ 'come from' to the anterior marker /ghera/ shows the implosive /d/ reducing to [r] (Hoffmann 1963: 220–221). For Ngambay, Vandame (1963) reports that plain [b] and [r] are always an optional variant of the labial and dental implosives. Also in Ngambay, labial and dental implosives are restricted to the beginnings of morphemes which suggests possible reduction in other contexts.

As for ejectives, one possible outcome of reduction would be a plain stop, possibly voiced, at the same place of articulation. However, another possibility is the loss of the supraglottal articulation, as reported for Bari by Spagnolo (1933). This author reports that /'b/ and /'d/ may be pronounced implosively or explosively or with the supraglottal articulation totally left out, so that only a glottal catch remains. He gives these examples: /na'but/ > [na'ut] and /'dupyet/ > ['upyet]. If this is a common path of reduction for ejectives, then this may contribute to the frequent presence of glottal stop in affixes along with the fact that glottal stops can also represent the late stages of reduction for plain stops.

Another crosslinguistic trend to consider is that some languages have restrictions against more than one ejective per word. Hopper (1973) mentions Hausa, Yucatec Mayan, and Quechua. In addition, in the Hopper (1973) and Gamkrelidze & Ivanov (1973) reconstruction of Proto-Indo-European stops, one of the restrictions on sequences of root consonants would be stated as a prohibition against two glottalized consonants in the root. If there is a tendency to lose glottalization on a consonant if another glottalic occurs in the word, then this might lead to the loss of glottalization in affixes. The source of this constraint might be a natural tendency to produce the ejective in word-initial position with more glottalization than ejectives further along in the word, as is reported for Maidu (Shipley 1964: 7). This tendency might lead to the loss of glottalization in non-initial ejectives.⁵

5. A perceptually based explanation for a dissimilation if two glottalized consonants are found in one word is presented in Ohala (1993).

Thus there is probably not just one explanation for the rarity of ejectives and implosives in affixes.

10.2. *Reduction in grammaticization*

The trend towards phonological reduction in the grammaticization process is extremely well-documented (Bybee et al. 1994, Heine & Reh 1984, Lehmann 1982). This trend could explain the loss of secondary articulations such as palatalization and labialization as well as the reduction of ejectives and implosives, as just noted. In addition, the high percentage of glottal stop and /h/ in affixes could be the outcome of reduction of other stops and fricatives in the grammaticization process. However, inherent problems arise in finding a complete fit with the data that would be predicted by grammaticization. For instance, the high percentage of voiceless stops is not predicted by grammaticization theory, unless they were the reduced version of some more elaborated stop, or unless we were to claim that voiceless stops are resistant to reduction. Similarly, the high percentage of nasal consonants in spite of grammaticization would suggest that nasal consonants are highly stable, which certainly appears to be the case. However, the fact that some consonants are stable or resistant to reduction under grammaticization while others are the outcome of reduction processes makes it difficult to find a coherent prediction about what consonants occur in affixes from grammaticization theory alone.

In addition, reductive processes differ across languages: in some a voiceless stop might become voiced, in others it might lose its supraglottal articulation, becoming /ʔ/. Another scenario is that a voiceless stop might become a fricative and then /h/. Thus grammaticization predicts a wide range of outcomes. Of course, since the data also show a wide variety of outcomes, the application of grammaticization theory is not disconfirmed; it is, however, difficult to test.

In this context it is interesting to consider Abkhaz and !Kung, both of which use highly elaborated consonants in affixes. To review, out of nine consonants used in affixes in !Kung four are clicks. In Abkhaz, seven out of fourteen ejectives occur in affixes and seven labialized consonants are used. The question that arises about these two languages is not why they violate a synchronic constraint, but rather why grammaticization in these languages does not eliminate elaborated segments. What they have in common are enormous consonant inventories and a low degree of phonological reduction in affixes. The two languages differ, however, in that Abkhaz has a very large number of affixes meeting the coding criteria, while !Kung has only eight. Indeed, it appears by all measures that !Kung in general has a low degree of grammaticization. Abkhaz, on the other hand, has a high degree of grammaticization, but a low degree of phonological reduction in affixes. As evidenced by the fact that languages come

in a variety of morphological types, such as those described by Sapir (1921), grammaticization may proceed to different extents in different languages (Bybee 1997). We must conclude, then, that grammaticization does not provide a single unified prediction about what phonemes will occur in affixes. However, since the data are also not unified, the effects of grammaticization cannot be ruled out as possible causes.

While we have found languages that use fewer complex and elaborated consonants and vowels in affixes, we have found no languages that use only complex or elaborated and exclude basic consonants and vowels. This one absolute universal is readily explained by the reductive directionality of grammaticization. General processes of reduction of secondary features or complex articulations are found in the formation of affixes, but no processes that routinely make the segments in affixes more complex.

10.3. *Re-use of grammaticized material*

Some languages use an unusually restricted set of phonemes in affixes, given the number of affixes they have. The restrictions are not explainable in the terms outlined above, because they are not patterned and do not necessarily involve elaborated segments. As examples, let us take Baluchi, Kanuri, and Pangasinan which have fewer consonant phonemes in affixes than would be expected by chance (see Table 3 above).

Taking a close look at Baluchi (an Indo-Iranian language), we see that with sixteen consonant slots in inflectional affixes, only five out of 25 consonant phonemes are used. The five are: /b/, /t/, /g/, /n/, and /j/. This seems to be a random set as it includes all places of articulation (except retroflex) and voiced, voiceless, and nasal stops. Note also that the same affixes turn up in a variety of constructions: the infinitive suffix /-əg/ is an element of ten periphrastic constructions, but was only counted once in the application of the formula. Also, the past suffix /-yt/ occurs in nine constructions, but was counted only once. On the other hand, some of the same etymological elements show up in more than one affix and thus were counted each time. The past completive is formed with past /-yt/ and /-ət/, which is related to the past form of the copula (and possibly also to /-yt/). The past perfect completive is /yt + ətət/ which appears to involve a reduplication of the past form of the copula. Even counting each element (/yt/, /ət/ and /ətət/) only once, four consonant slots occupied by /t/ probably came from the same etymological source. The other reason for the small number of consonants in affixes in Baluchi is the reduction of consonants. 1st person singular and plural, which would have had /m/ in them earlier, as evidenced by other Indo-European languages, now have /n/ as does 3rd person plural, which had /n/ etymologically. The negative prefix also has /n/, as do two other suffixes. Thus the re-use of old grammatical affixes in the creation of new ones, as

well as the phonological reduction of affixes accounts for the small number of consonants used in affixes in Baluchi.

Pangasinan also appears to have re-used some affixes in the creation of new ones, which leads to a small number of consonants in affixes. The 19 affixes we coded (all prefixes except for one) have 29 consonant slots, but only use six distinct consonant phonemes in these affixes: /p/, /m/, /s/, /n/, /k/, and /g/. Pangasinan also uses only two (/i/ and /a/) of its five vowels in affixes. The affixes appear to be made up of a much smaller number of recurring elements. For instance, the element /pa/ appears as a causative, also in /ama-/ (aN + pa) meaning successful completion, in /impama-/ (impaN + pa) for reason or circumstance focus, and in /pama-/ (paN + pa) for causative with reason focus; /impaN/ occurs in the form just cited and as an instrumental, passive, and infinitive. Another example is /maka-/ for possibility and a partial reduplication /makaka-/ for a tendency (Benton 1971: 123–141). In fact the recurrence of the syllables /an/, /pan/, /man/, /gi/, /in/, and /i/ strongly suggests that many of the focus prefixes are made of combinations of old prefixes (Benton 1971: 196–197).

Kanuri has a larger consonant inventory than the languages just discussed, with 25 consonant phonemes, but it only uses nine of these in its 26 consonant slots in affixes. Here part of the explanation is the re-use of certain consonants in person and number markers. For instance, /n/ occurs in three 2nd person markers, /z/ in two 3rd person markers, and /s/ in two 2nd person singular markers. However, other recurrent consonants, such as the use of /n/ in four other affixes, are not accounted for in this way.

Other languages with complex person/number agreement systems often re-use elements in the formation of agreement affixes. For instance, in Tojolabal the element /k/ or /ik/ occurs in six affixes of the person/number system, primarily in 1st person but in one case in 2nd person plural. In Abkhaz the 26 person/number prefixes (for ergative, absolute, and indirect object) re-use the same nine consonants. Thus complex person/number agreement decreases the number of different phoneme types used in affixes.

11. Conclusion

Let us return now to the four hypotheses formulated in Section 3. The systematic examination of the number and type of phonemes used in affixes in a sample of the world's languages shows that Hypothesis 1, that languages use a more restricted number of phonemes in affixes than would be expected by chance, is only supported in small minority of the languages of the sample. Hypothesis 2 stipulates that highly marked segments are excluded from affixes while Hypothesis 3 predicts that highly complex segments are so excluded. Neither hypothesis provides a good fit with the data, but it appears to me that

the latter hypothesis has stronger support because certain marked consonants, such as /h/ and /ʔ/, occur in affixes more than they are excluded. Finally, Hypothesis 4 which merely predicts pattern exclusions is also only weakly supported. The primary class of segments that tend to be excluded are ejective and implosive consonants, but even this trend is by no means universal.

The data on vowels reveal stronger universals: all the languages have a low vowel and a high front vowel and all but one language uses these in affixes. Common exclusions include rounded vowels and mid vowels. Of the six languages that have a schwa phoneme, only one does not use it in affixes (Engenni, which has only three affixes coded).

Our examination of the languages where these trends can be documented reveals that there are different explanations for different languages. A general tendency for phonological reduction in affixes explains many of the cases, including the rarity of ejectives and implosives in affixes, but for languages with highly reduced inventories, the probable explanation is the re-use of affixes in the creation of new affixes, which means that the same phonemes are re-used in new affixes. Of the five languages with the most restricted inventories of consonants in affixes, four of them (Baluchi, Pangasinan, Kanuri, and Tojolobal) show evidence of the re-use of the same material in multiple affixes. Shuswap's limited inventory seems to be the result of non-use of ejectives and uvulars. It is possible that if these consonant types were used earlier in affixes, they have been reduced to plain stops and velars respectively.

The acquisition of phonemes by borrowing and their consequent absence in affixes has limited applicability and it was not possible to substantiate the idea that complex segments develop in stems but not in affixes. It was also not possible in these data to investigate Hopper & Traugott's claim that phonologically less marked allomorphs tend to replace more marked ones in analogical change.

The wide variety of situations displayed by the languages of the sample strongly suggests that there is not a single unified explanation for this weak trend and thus argues against this trend as a universal constraint. The data have shown that the lower number of phonemes used in affixes is in large part due to the fact that languages have more lexical stems than affixes, that affixes are short and thus utilize fewer phonemes. It is also possible that the fact that marked phonemes are less frequently used than unmarked phonemes in the languages that have them also contributes to the smaller number of marked phonemes in affixes. The existence of patterned exclusions of consonants was only a minor trend, evidenced in only six out of 22 languages. Not only are such exclusions not universal, they also reflect diachronic processes such as the phonological reduction in affixes during the grammaticization process rather than synchronic constraints. The evidence against a universal synchronic constraint lies in the fact that so many of the languages in the sample had the num-

ber of phonemes in affixes that could be expected by chance and two languages – Abkhaz and !Kung – use highly elaborated consonants in affixes.

Received: 15 March 2004
 Revised: 31 March 2005

University of New Mexico

Correspondence address: MSC03-2130, Linguistics, University of New Mexico, Albuquerque, NM 87131, U.S.A.; e-mail: jbybee@unm.edu

Acknowledgements: I am grateful to Keri Holley for assisting with the data analysis, James Gagnon and Clay Beckner for assistance with the statistical formula, Vsevolod Kapasinski for preparation of the manuscript, and Ian Maddieson and Caroline Smith for phonetic consultation.

Appendix 1: Phoneme inventories for the sample languages

Languages ordered as in Table 2 followed by the supplementary sample as in (4).

Language	Phonemes	
	used in affixes	not used in affixes
Inuit	C t k ʏ q s j β m n l ŋ v p V i a u	h n ɹ
Margi	C ʔ d g k m n ŋ r s j 'j d dʒ V a ə i u	b ʃ c ɓ dz f gʲ ʏ j x ç kʲ l p f t l t s v v b w 'w z ʒ e o
Cheyenne	C h ʔ m n s f t v x V a e o a: o: i	k p
Tucano	C p t tʰ k kʰ b d g m r s w j V i e i a u o	pʰ bʰ dʰ gʰ mʰ n nʰ ŋ jʰ rʰ wʰ jʰ ʔ x sʰ
Gugu-Yalanji	C b w ʃ n k l ŋ d m r r r V i a u	j
Pangasinan	C p m n s g k V i a	j w b d t r l ŋ h ɛ u u
Tohono O'odham	C t k b d ʔ ʃ h m n ŋ j d V i u o a	p t ʃ ʒ g s l w i
Abkhaz	C b pʰ p' d tʰ t' tʃ z s f fʷ ʏ x h m n l g g j r w kʷ q' k tɕʷ ʔʷ qʲ tʷ ʃ ʒ dz ts' x tsʷ ts ɕ' tɕʷ gʷ p t k̃ V ə a	dʷ tʷ dzʷ dʒʷ tʃʰ tʃ' dʒ k' kʷ, qʷ, v f f' ʒʷ ʏ x ʏʷ xʷ k' tʃ' z hʷ
Guaymí	C d b j n r t V a e i	tʃ g k l m ŋ s w o ö ó u ü

Language	Phonemes		
	used in affixes	not used in affixes	
Kui	C	k t d s n m r j	g c t d p b v h ŋ r l
	V	e i u a	o a: e: i: o: u:
Abipon	C	p t q m n ɲ l ʀ h ɣ r	k t f j w
	V	i e a o	i
Karak	C	p t tʃ k ʔ f ɸ f h β r m n	s x j
	V	i i: a a: u o:	e: u:
Baluchi	C	g n t b j	tʃ d d f ɣ h k l m p r r̄ s f t v
	V	a i o ə ɪ	x z ʒ r e u ʊ
Yagaria	C	p b t d g ʔ v s m n l	k f h j
	V	i u o a ε ou ei	aε ao
!Kung	C	t k ^h ʃ ^{h?} m ŋC ^{h?} w ʃ ^s ŋ ^ʃ C ^{h?}	p l p ^h t ^h b d d̄ g g t' b' d' g' ts ts tʃ tʃ ts ^h tʃ ^h dz d̄ʒ ts' tʃ' ʔC ^h ŋ ^{ʃ?} ŋ ^{ʃ?} ŋC ^ʔ ŋC ^ʔ g ^ʃ gC ŋC ḡC g ^ʃ gC ŋC ḡ ^{ʃ?} gC ^ʔ ʃ ^s ʃ ^{sh} ŋ ^{ʃsh} ŋ ^{ʃs?} ŋ ^{ʃs?} g ^{ʃs} ŋ ^{ʃs} ḡ ^{ʃs} g ^{ʃs} ŋ ^{ʃs} ḡ ^{ʃs?} C ^λ C̄ ^λ C ^{λh} yC ^{λh} ŋC ^{λ?} ŋC̄ ^{λ?} gC ^λ ŋC ^λ ḡC ^{λ?} gC ^λ ŋC ^λ dz dz' dʒ' x s ʃ z ʒ ŋ ŋ' m: n m̄ m̄ r j C C̄ ʃ ^h ŋ ^{ʃh} C ^h ŋ ^ʃ ŋC ^{h?} ʃ h k æ ^ʃ ʊ ^ʃ u
	V	e i o a	æ ^ʃ ʊ ^ʃ u
Slave	C	k ʔ ts' s l x h ɣ m n j ɲ g d	p k ^w t ^h k ^h t' k' ts tʃ tʃ' tʃ' f ʃ z l ʒ ɣ ^w 'ɣ ^w r t
	V	i e a o u ε	ie ī ã ũ ö
Engenni	C	s n r	kp gb bh dh p t d k g f v z m l w j
	V	e u	i ɛ ə a ʊ o ɔ
Kanuri	C	b g k m n s t j z	ʔ tʃ d f ɸ h dʒ l l̄ ɲ ŋ p r f w x
	V	a e i o u ə	a
Palantla Chinantec	C	k g t ^s z h m n ŋ r l j	p t b d f s ʔ w
	V	ɪ ʉ ʊ a	ε ø o
Shuswap	C	p m t n l c s j k k ^w x x ^w ʔ w	p' m' t' λ n' l' c' j' k' q q' ɣ q ^w q ^w ' ɣ ^w h ɣ ɣ' ʙ ʙ ^w ʙ ^w w'
	V	e i u ə	a o
Tojolabal	C	t k ʔ m n f j h w	p b' t' k' ts ts' tʃ tʃ' r s l b d g

Form	Label	Pages
tit	cause/let	84–85, 87, 265, 268, 270, 282, 325
qqu	ask/want/tell to	268–270
sav̆i	get to/cause to	268, 279, 283
qqammiv̆	recent past	273
v̆i:v̆kataʉ	distant past	273
ssa	future	64–67, 274–275
niav̆	future	275–275, 325
juma:v̆	indefinite future	275, 325
ʉina	future	225, 294–295
v̆i:v̆	action completed prior to some reference point	278
v̆iav̆	imperfective	279
nʉit	negation	136, 279, 288
juit	(can) never	281
jua:v̆	continuous	281 ff.
jusa:v̆	keep on V-ing	282
ʉiav̆tu:v̆	progressive	282
ʉiav̆tua:v̆	protracted progressive	282
liv̆	ingressive aspect	282–283
jupna:v̆	terminative aspect	283
ssa:v̆	terminated state	283
qatta:v̆	indefinite iterative	283–284
latta:v̆	from time to time	284
jurtu:v̆	one after another	284
qqiʉ	again	284
qqa	state	284–285
latau:v̆	for once	285
qqajav̆	imminent action	285
juma:ta:v̆	take a long time V-ing	285
latsiav̆	for a while	285
v̆vat	in a crowd/herd	285
qqa:v̆	barely	285
qqa:v̆	first/for first time	285
ja:llu	habitually early	279–280
vav̆	indicative	288–289
ʉu	conditional	290
li	optative	24–26, 291–292
nʉa	1SG.ABS	288–292
ʉut	1PL.ABS	288–292
tit	2SG.ABS	288–292
si	2PL.ABS	288–292
q	3SG.ABS	288–292
pput	3PL.ABS	288–292

Form	Label	Pages
vi	interrogative	4 ff., 289
niaʁ	try to	292, 325
lirsa:ʁ	intend to	292
sussa:	is to	292
sinna:	can	293
ja	tend to	292
ʒuminaʁ	be good/easy to	293
llaqqiɣ	be good at	293
navi:ʁ	can no longer	293
yi	imperative	24 ff., 291
su	participial	28–29, 34 ff., 60, 289
ɣunaʁ	it seems	293
ʒunnaʁsi	probably/presumably	293–294
qqu:qa	undoubtedly/must have	293–294
qina	monitory	294–295
yaluttuaqi	monitory	295
yi	consecutive	295, 324
ɲat	causative	287, 295
kasiɣ	subjective coloration	295
nɲuaʁ	affection	295–296
ʁianɲuaʁ	surprise	295–296
ɣa	causative	32, 56–57, 64–65, 290
qi	very	296
llaʁuma:ʁ	intensified future	296
ɣaluʁ	(sure) but	296
llu	contemporative	297, 120–126
uti	reciprocal	165
ɟuma	want to	325
kataɣ	be fed up with	325
ʁusuɣ	want to (a lot)	325
ɣiɟaʁtuʁ	go and/in order to	325
navi:ʁsa:ʁ	try not to	325
ssama:ʁ	plan/intend	325
tsi	2SG/1PL.ERG	289
ati	3PL.ERG	289
ssi	2PL.ERG	289
uk	SG.OBJ	290
ɣik	PL.OBJ	290
ni	4SG.SUBJ	290
nik	4PL.SUBJ	290
kk	NON3SG.SUBJ	290
aɣ	3SG.SUBJ	290
siɣ	1PL.SUBJ	290

Margi (Hoffmann 1963)

Form	Label	Pages
jú	1SG.SUBJ	74
gu	2SG.SUBJ	74
dʒá	3SG.SUBJ	74
ma	1DU.INCL.SUBJ	74
ʔja	1PL.EXCL.SUBJ	74
mər	1PL.INCL.SUBJ	74
jí	2PL.SUBJ	74
nda	3PL.SUBJ	74
ɗa	1SG.OBJ	75–77
ɲu	2SG.OBJ	75–77
ɲi	3SG.OBJ	75–77
mà	1DU.INCL.OBJ	75–77
ʔjà	1PL.EXCL.DO	75–77
mər	1PL.INCL.DO	75–77
ɲì	2PL.DO	75–77
ndà	3PL.DO	75–77
amu	imperative plural	170–171
a	conjunctive	188–190
a	present	190–196
aʔi	negative	169–170
əri	past	196–200
a	subjunctive	200–205
ska	exclusive	206–209

Cheyenne (Leman 1980)

Form	Label	Pages
na	pronominal	20
ni	pronominal	20
ex	pronominal	20
o	pronominal	21
a	pronominal	23
i	pronominal	55
a	pronominal	21
atse	pronominal	55 ff.
ahtse	pronominal	55 ff.
ho	pronominal	55 ff.
heme	plural	27, 55
no	plural	27, 55 ff.
vo	plural	27, 55
oʔo	plural	55
ae	pronominal	55

Form	Label	Pages
az/e	pronominal	27, 55 ff.
aetseno	pronominal	55
to	pronominal	55
ne	pronominal	27
ma	pronominal	27
amo	pronominal	55ff.
otse	pronominal	87
nestse	pronominal	43, 87
sa:	negative	31, 45, 69, 94
hane	negative	
he	interrogative	34, 73
mo	dubitative	79
mase	attributive	81
o?o	delayed imperative	85, 99, 41
?eha	hortative	42, 86, 100, 142
ho:no	mediate	83, 98
net	relational	44
tseh	conjunct verb	110
mah	conjunct	110
ve?	conjunct verb	110
oh	conjunct verb	110
to?se	conjunct verb	110
ho?	conjunct verb	110
ah	obligative	110
momoxe	optative	110
moho?no	intensive negative	110
eo	conjunct interrogative	110
ho?nefe	concessive	111
tsexhefe	complement	111
eme?hefe	purpose	111
tsexhoma?xe	causal	111
tsehve?	grounds causal	111
to	conjunct person	110–122
se	conjunct person	110–122
tsese	conjunct person	110–122
tse	conjunct person	110–122
se	conjunct person	110–122
vose	conjunct person	110–122
?seh	causative	139
ve?e	prohibitive	141
nex	directional	147
ta	directional	147
ne	directional	147
tse	directional	147

Form	Label	Pages
tano	desiderative	148
hta	future	191
s	past tense	191
neh	far past	191
stse	future	191
hta	past	191
ane	passive	55
estse	immediate imperative	41, 84, 99
hene	delayed imperative	41, 85, 99
taʔe	immediate hortative	42, 86, 100, 142
tse	immediate hortative	42, 86, 100, 142
ta	1st person hortative	42, 86, 100
movo	relational	90

Tucano (West 1980 [W], Sorensen 1969 [S])

Form	Label	Pages
w	remote past	W28
p	supposed past	W27
s	emphatic	W29
i	1st person	S162
mi	2.MASC	S145
mo	3SG.FEM	S145
ma	3PL.INAN	W24–25
àtó	encouragement	S154–155
tí	interrogative	S146–147
já	imperative	S154–156
tíkàjá	negative order	S155
sómé	negative hortative	S155, S184
á	certainty	S156–161
sā	probable	S156–161
kà	perfective	S165–167
bō	debitative	S168–169
ti	negative	S169–170
mi	frustrative	S167–170
t ^h óà	completive	S171–173
dú	completive	S171–176
péò	completive	S173–176, W90
gi	continuative	S267–269, W67–69
ti	anticipated action	W74–75
ki	proof of action	W75–76
ák ^h i	prediction of an action	W76–77
ro	obligatory action	W78

Form	Label	Pages
ti	def future	W36
a	complete action	W78–80, S271–272

Gugu-Yalanji (Hershberger 1964)

Form	Label	Pages
n	completed action	35–36, 49, 51 ff.
l	non-completed	36, 39, 42
ka	imperative	36, 37 ff.
pi	warning	36, 71
pa	subordinator	36 ff.
ja	inchoative	37
ma	become	38, 42 ff.
ri	pluralizer	41
mana	transitivizer	41 ff.
bunga	make	44 ff.
wa	reciprocal	45 ff.
ji	passive	45 ff.
ngarr	first	51
arr	completely	51 ff.
ku	still	53, 69 ff.
da	now	53, 69 ff.
ɲjiku	before	53
ɲaku	should	37
nka	purposive	36

Pangasinan (Benton 1971)

Form	Label	Pages
on	incomplete	133 ff., 196–197
in	complete	123, 125, 128
maka	potential complete action	126–141
makaka	potential incomplete action	133
a	potential complete passive	127
na	potential incomplete passive	134
paka	involuntary	127–141
aki	mutual	128, 134
pa	causative	137
an	referent	127–141, 168–169, 190 ff.
i	passive	127–141, 190–191, 196 ff.
nan	intentional	133, 136

Form	Label	Pages
pan	instrumental	138–139
magsi	distributive	132, 135
man	frequentative	133
gi	transitivizing	127–141
nai	actor responsible	136
nan	mutual	134, 136
pinan	agentive	140–141

Tohono O'odham (Mathiot 1973–78)

Form	Label	Pages
ɲ	1SG.OBJ	55–57
t	1PL.OBJ	55–57
ʔe	reciprocal	55
m	2SG.OBJ	56
ʔem	2PL.OBJ	56
ha	3.OBJ	56
ta	3.INDEF.HUM	56
ʃu	3.INDEF.NHUM	56
ok	completive	58
him	interruptive	57–58
da	durative	59
kaʔi	immediate	59
j	3.DEF.SG.OBJ personal	
ka	stative	64
ɲ	imperative	61
ʃk	correlative	62
kɔ	concurative	63
ʔamjɔ	locational	67
da:m	locational	65–66
ʔam	locational	65–66
ʔan	locational	65–66
ʔab	locational	65–66

Abkhaz (Hewitt 1979)

Form	Label	Pages
p'	present	167, 172, 181
r	causative	170–171
sə	1SG.SUBJ _{tr} /OBJ _{itr}	101–105
wə	2SG.MASC.SUBJ _{tr} /OBJ _{itr}	101–105
bə	2SG.FEM.SUBJ _{tr} /OBJ _{itr}	101–105
də	3SG.HUM.SUBJ _{tr} /OBJ _{itr}	101–105

Form	Label	Pages
jə	3SG.NHUM.SUBJ _{tr} /OBJ _{itr}	101–105
ħa	1PL.SUBJ _{tr} /OBJ _{itr}	101–105
ʃ ^w ə	2PL.SUBJ _{tr} /OBJ _{itr}	101–105
jə	3PL.SUBJ _{tr} /OBJ _{itr}	101–105
sə	1SG.IO	102–105
wə	2SG.MASC.IO	102–105
bə	2SG.FEM.IO	102–105
jə	3SG.HUM.MASC.IO	102–105
lə	3SG.HUM.FEM.IO	102–105
a	3SG.NHUM.IO	102–105
ħa	1PL.IO	102–105
ʃ ^w ə	2PL.IO	102–105
rə	3PL.IO	102–105
sə	1SG.SUBJ _{tr}	102–105
wə	2SG.MASC.SUBJ _{tr}	102–105
bə	2SG.FEM.SUBJ _{tr}	102–105
jə	3SG.HUM.MASC.SUBJ _{tr}	102–105
lə	3SG.HUM.FEM.SUBJ _{tr}	102–105
na	3SG.NHUM.SUBJ _{tr}	102–105
ħa	1PL.SUBJ _{tr}	102–105
ʃ ^w ə	2PL.SUBJ _{tr}	102–105
rə	3PL.SUBJ _{tr}	102–105
zə	sentential adverb ‘for’	113
z	imperative	113, 189
la	instrumental	114
ts	comitative	114–115
jtʻ	finite marker	180–181, 174–175
x	reference ‘about’	119
t ^h ə	relative ‘out of’	129
ʎ ^w nə	motion past ‘through’	129 ff.
ɣra	inside the body	128
p ^h əra	in front of	130
ʃt ^h ə	essive-allative	132
mə	negative	70–73, 172, 266
tʃə	reflexive	77
ajba	reciprocal	86ff., 160, 172
wa	dynamic	172
n	past	173
ɣa	perfect	175, 180–181
za:	future of stative	176
pʻ	future I	177
tʻ	future II	176–177
eʻə	progressive	128, 181–182

Form	Label	Pages
la	iterative	182
k ^w a	several times	183, 212
r	protasis	27, 184, 192, 195
ndaz	optative	190–191
awaz	optative	191
nə	purposive	42, 191–192, 199–201
zə	potential	194–195
za:rən	inference	196
amxa	unwillingness	197–198
a:jt'	subjunctive	198–199
ts ^w ə	disadvantage	210
a:	action toward speaker	212–216
na	action toward hearer	212–216
ɿ ^w	action upwards	212–216
la	action downwards	212–216
tɕ ^w 'qj'a	really, indeed	219
tɕ ^w a	to excess	219
ja	intensifier	219–220
g ^w əfa	subject wretched	220
xə	again	220
ajta	again	220
egə	anything	221
l	preverb 'from in'	117, 120
akrə	dummy prefix	168, 220
anə	temporal adverbial clause	39
da	human NP question	10
ma	neutral yes-no question	8
w	yes-no question	8, 22
əj	yes-no question when yes is expected	9
zej	non-human NP question	2, 11–12
ba	temporal wh-question	13
zə	wh-question of reason	14–15
rt ^w '	purpose	17, 42, 201
ʂə	manner	7, 28, 40–42
aχə	complementizer	28–30
gə	concessive	41–42, 45
nə	past absolutive	218, 62
ta	in a delimited area	128
tɕ	conditional	187–189
k ^w ə	locative	210–211
tə	from-in	213
ka	down	214
tsəpχadza	every time that	231

Form	Label	Pages
jɔtej	since	231
nats'ə	while	232
ʃ ^w a	as if, as though	232–233, 30
a:ndza	until, before	233
aɦa	any more than, let alone	233–234
ts'q'a	indeed, really	40
n	past	173

Guaymí (Kopesc 1975 [K], Young 1987 [Y], Alphonse 1956 [A])

Form	Label	Pages
di	potential proximate	K22–23, Y7
ba	actual remote	K23, Y4–7
i	remote potential	K23, Y7
en		K23, 72, Y5
a	passive	K33
re	potential	K24, 79, 80
ni	perfect	K24, Y8, 10
ja	reflexive	A29
ra	completed action	K45
ta	repeated action	K45

Kui (Winfield 1928 [W], Subrahmanyam 1971 [S])

Form	Label	Pages
in	future (or indefinite)	W60–85, S278–279
enu	1SG	W63–66, S400
i	2SG	W63–64
eru	2PL	W63–66
it	past	W60–85, S157–160
ʔa	negative	W61–85
ara	transition	W101–111
amu	1PL.EXCL	W63–66, S400
asu	1PL.INCL	W63–66, S400
enju	3SG.MASC	S402
e	3SG.FEM/NEUT	W63–66, S402
u	3PL.LOW	S402
umu	2SG.IMP	W65, S481
k	hortative	W94–96
atu	2PL.IMP	W65–85, S481–82
ika	motion	W111–122
k	PL action form	W142–145
de	emphatic	W153

Form	Label	Pages
ai	present	W85–87
n	present	W85–87
s	imperfect	W85–87
a	perfect	W86–87

Abipon (Najlis 1966)

Form	Label	Pages
ji	1.SUBJ	30–33
n	2	30–34
i	2	30–34
n	3	30–33
aq	PLSUBJ	31–32
ran	causative	35–36
en	volitive	36
am	future	37
alta	reflexive	37
eta	reciprocal	37
roa	comitative	37
ye	def obj	38
a:ye	habitual	38
eoꝑe	again	38
i:	1SG.OBJ	39–40
ieꝑari	2SG.OBJ	39
aꝑeꝑ	3SG.OBJ	39
eꝑ	1PL.OBJ	40'
ieꝑar	2PL.OBJ	39
eie	3PL.OBJ	39
etapeꝑ	durative	41–42
m	interrogative	70–71
tꝑiꝑ	negative	71
a:	outside	40
alꝑe	surface	40
aꝑi	down	40
arhe	here, present	40
heꝑem	up	40
oa	inside	40
o:ꝑe	outward	40
aiꝑit	around	40

Karok (Bright 1957)

Form	Label	Pages
ni	1SG.SUBJ/3.OBJ	60
ná	2/3SG.SUBJ/SG.OBJ	60
kaná	2/3PL.SUBJ/1SG.OBJ	60
ap	3PL.SUBJ/1SG.OBJ	60
nu	personal morpheme	60
kín	personal morpheme	60
?i	personal morpheme	60
?i .. ap	personal morpheme	60
ku	personal morpheme	60
ki:k	personal morpheme	60
?u	personal morpheme	60
kun	personal morpheme	60
i	imperative	65–66
aβif	future	124
at	past	67
ip	iterative	88–89
kupa	modal	89
ahi	modal	89
βa	plural action	92–93
ar	class 4 suffix	106
ara	instrumental	106
fip	completely	106
ihi	benefactive	107
kiri	instrumental	108
maθ	causative	109
mara	to finish doing	108–109
o:	habitually	109
sar	along/together with	110
tánmah	for nothing/for no reason	110
unif	class 4 suffix	110
βa:na	reflexive	111
aβrik	class 4 suffix	111
ahi	essive	111
βuna	PL	112
tih	durative	113–114
atf	diminutive	114
ahε:n	anterior tense	125
aník	ancient time	125–126
pu	negative	137

Baluchi (Barker & Mengal 1969)

Form	Label	Pages
nə	negative	130–131, 182
bì	subjunctive	179–190
əg	want	196
a	stop doing	197
i	obligation	239
ɪt	past tense	282ff.
əg	present perfect	333ff.
ət	past perfect	336ff.
ətət	past perfect completive	340–341
an	iterative	237
bí	past subjunctive	459–460
en	past subjunctive	459–460
in	1SG	130
əj	2SG	130
it	3SG	130
ən	1PL	130
ɪt	2PL	130
ənt	3PL	130

Yagaria (Renck 1975)

Form	Label	Pages
ou	1SG/DU.SUBJ	87 ff.
oun	1PL.SUBJ	87 ff.
ɛin	2SG.SUBJ	87 ff.
ɛiʔ	2DU.SUBJ	87 ff.
ɛi	3SG/2PL/3PL.SUBJ	87 ff.
da	1SG.OBJ	20 ff.
laʔa	1DU.OBJ	20 ff.
la	1PL.OBJ	20 ff.
ga	2SG.OBJ	20 ff.
lata	2DU.OBJ	20 ff.
lapa	2PL.OBJ	20 ff.
ta	3PL.OBJ	20 ff.
pa	3PL.OBJ	20 ff.
no	present progressive	90 ff.
d	past	92 ff., 114 ff.
boʔo	completed action	94
s	intentional future	9 ff.
g	future	95 ff.
o	imperative	97 ff.
gapa	intensifier	98 ff.

Form	Label	Pages
avi	interrogative	101–103
agi	emphatic mood	105
lo	completed	110
du	perfective	110 ff.
li	extended	111
mo	aspectual continuative	133 ff.
gogo	prolonged action	120
go	habitual	119 ff., 135
ama?	motivational	127
bose?	motivational	127
to?	real conditional	128, 135
?εne	irreal conditional	129 ff.
bobo	potential conditional	128
tone	irreal conditional	129 ff.
ge?	in order to	131 ff., 95
a?	negative	84 ff.
go	habitual continuative	134
gεse?	referent action	136
ε	indicative mood	86
ga	SG/PL non-identity of subjects	80
aga	DU non-identity of subjects	83

Kung (Snyman 1970)

Form	Label	Pages
wa	terminal -a	128–129
ŋ [†] ei	cause	129–1131
C ^{h?} ae	together	150
ɸ ^s wa	with	150–151
wak ^h we	each other	150–151
ɸ ^{h?} a	for, on behalf of	150–151
tama	unsuccessfully	150–151
ŋC ^{h?} o	aimlessly	150–151

Slave (Rice 1986)

Form	Label	Pages
h	h-classifier (voice)	442–443
d	d-classifier	443–445
h	1SG.SUBJ	471–472
nε	2SG.SUBJ	473–475
íd	1DU/PL.SUBJ	476–481
ah	2PL.SUBJ	481–483

Form	Label	Pages
ɣu	optative mode	556
ɲ	perfective mode	528
dɛ	inceptive	587–592
j	future	511
j	progressive	802
nɛ	terminative	800
í	seriative	594–595, 800
í	semelfactive	595–596, 801–802
í	transitional	596–597, 801–802, 817–818
u	conative (tentative)	599–601
dɛ	benefactive	601–603
dɛ	noise of mouth	604–605
dɛ	wood	606–607
gɛ	human PL deictic subject	623–624
ts'ɛ	unspecified subject	624–625
sɛ	1SG DO	627
nɛ	2SG DO	627
mɛ	3.ANIM.DO	627
jɛ	4.DO	628–629
ʔɛ	unspecified object	629–631
naxɛ	1/2PL OBJ	631–612
ku	3PL.ANIM.OBJ with 1/2.SUBJ	632–633
go	3PL.ANIM.OBJ with 3.SUBJ	633
ki	3PL.ANIM.SUBJ with 3SG.OBJ	633–634
ʔɛdɛ	reflexive	634
ʔɛtɛ	reciprocal	634
go	areal gender prefix	634–636
nà	customary	673–676
jà	distributive	677–701

Engenni (Thomas 1978)

Form	Label	Pages
ene	of its own accord	142 ff.
se	causative	142 ff.
ru	more than once	143

Kanuri (Lukas 1937)

Form	Label	Pages
in	continuous	35–36, 48, 50 ff.
e	conjunctive	39–42

Form	Label	Pages
go	past	40–41
e	optative	42
o	relative past	42–43, 150–153
nà	perfect	43–44
i	predicative	45–47
zo	future	46–47
sko	1SG	35 ff.
nəm	2SG	35 ff.
zə	3SG	35 ff.
njè	1PL	35 ff.
nəu	2PL	35 ff.
zâ	3PL	35 ff.
ənjí	negative past and future	68–69
njá	dependent past	70
ija	dependent future	71–74
jógo	imperative	74–75
jitə	causative	101 ff.
s	1OBJV	112 ff.
n	2OBJV	112 ff.
a	objective plural	112 ff.
bá	interrogative	36, 152
tə	passive-reflexive	93–101

Palantla Chinantec (Merrifield 1964)

Form	Label	Pages
t ^s a ¹	negative	26
mi ³	imperfect	25
ma ²	perfect	25
ma ¹	terminative	25–26
la ³	non-causative	26
na ²	action completed earlier same day	25
ka ¹	action just/previous day completed	25
li ¹	intensive active	22
lu ¹	completive active	22
ri: ²	stative	22
ha ²	directive: toward speaker	23 ff.
ŋji ²	directive: away from speaker 1SG	23 ff.
za ²	directive: away from speaker 1PL	23 ff.
gu ²	directive: away from speaker 2PL	23 ff.
za ¹	directive: away from speaker 3	23 ff.
ŋji ¹	directive: away from speaker (completive)	23 ff.
ha ¹	directive: toward speaker	23 ff.

Form	Label	Pages
hi ¹	directive: toward speaker	23 ff.
u ¹	directive: toward speaker	23 ff.
za ³	directive: away from speaker	23 ff.
gu ³	directive: away from speaker	23 ff.
u ³	directive: away from speaker, complete, and return	23 ff.
nji ³	directive: away from speaker, complete and return	23 ff.

Shuswap (Kuipers 1974)

Form	Label	Pages
Vm	intransitive	44
en	1SG.SUBJ	44
t	1PL.INCL	44–48
ex ^w	2SG	44–48
ep	2PL	44–48
əkw ^e	3RD	44–48
xe	IMP.SG	45
x ^w je	IMP.PL	45
səkp	conditional	45–48, 74
nwént	manage, be able to	45–46
cém	1SG.OBJ	46
él	1PL.INCL.OBJ	46ff.
cí	2SG.OBJ	46–48
úlm	2PL.OBJ	46–48
cút	reflexive	46
wéx ^w	reciprocal	46
em	passive	47
st	causative	57
c	customary	51, 53, 71
ep	arrive into a state	61
əmi?	do something all the time	62
t	state	62
nke	evidential	74
nt	trans	46–48
xít	human secondary obj	46–51
m	aorist	74, 80
k	predicative	44, 74

Tojolabal (Furbee-Losee 1976)

Form	Label	Pages
ʃ	incompletive	129, 134
k	1.ERG	123
haw	2.ERG	123
j	3.ERG	123
uneh	perfective	129, 133–134
ʃ	middle	129, 136
h	passive	129, 136
ij	main clause	136, 140
tik	1PL.INCL.ERG	123 ff., 129
tikon	1PL.EXCL.ERG	123, 129
eʃ	2PL	123, 129
eʔ	3PL	123, 129
on	1SG.NOM	123–129
a	2SG.NOM	123–129
otik	1PL.INCL.NOM	123–129
otikon	1PL.EXCL.NOM	123–129
ik	2PL	123–129
an	subjunctive	141
mi	negative	262–263
ma	question	260–263

Dakota (Buechel 1939 [B], Boas & Deloria 1941 [BD])

Form	Label	Pages
wa	indefinite object	B29, BD52 ff.
itʃi	reflexive	B29, B51, B62
wa	1	B35 ff.
ja	2	B35 ff.
ũk	dual	B35 ff.
pi	plural	B35 ff.
ma	1.OBJV	B37 ff.
ni	2.OBJV	B37 ff.
witʃa	3PL.OBJV	B37 ff.
ki	possessive	B42ff., BD86–8, BD93
kitʃi	dative	B42, BD86 ff.
hã	progressive	B281
k ^h ija	causative	BD100
ka	rather	BD55–56
kel	somewhat	BD56–57

Buriat (Poppe 1960)

Form	Label	Pages
na	present	56 ff.
ba	past	56 ff.
u:ʒa	future	56 ff.
hä:	optative	56 ff.
hu:	1.PRES.IMP	59 ff.
ja:	1PL.PRES.IMP	59 ff.
i:	2.PRES.IMP	59 ff.
gti:	2PL.PRES.IMP	59 ff.
g	3.PRES.IMP	59 ff.
arä:	future imperative	59 ff.
b	1SG	56 ff.
ʃ	2SG	56 ff.
bdi	3PL	56 ff.
t	2PL	56 ff.
d	3PL	56 ff.
gȳ:	indicative	57 ff., 83 ff.
gy	interrogative	57 ff., 83
xa	future VN	61 ff.
a:	imperfect VN	62 ff.
artä:	passive VN	64
nxä:	distant past VN	64
han	perfect VN	65
gʃa	1.PRES VN	66 ff.
a:ʃa	2.PRES VN	67 ff.
dag	frequentative VN	68 ff.
bal	conditional gerund	71
tar	limitative gerund	71
msa:r	precedent gerund	71 ff.
xla:r	successive gerund	72
ha:r	continuous gerund	72
ga:	causative	99 ff.
gda	passive	100
lda	reciprocal	100
lsa	cooperative	101
ʃxa	perfective	101
saga:	iterative	101
ʒa	compound of imperfective gerund	102
a:d	compound of perfective gerund	103
mni	1SG.POSS	42 ff.
ʃni	2SG.POSS	42 ff.
i:nʃ	3.POSS	42 ff.
mnä:	1PL.POSS	43

Form	Label	Pages
tnä:	2PL.POSS	43
ja:	reflexive possessive	46

Kanakuru (Newman 1974)

Form	Label	Pages
ná	1SG.SUBJ pronominal	138
ká	2SG.SUBJ pronominal	138
cìj	2SG.FEM.SUBJ	138
cí	3SG.MASC.SUBJ	138
cír	3SG.FEM.SUBJ	138
món	1PL free subject	138
kó	2PL.SUBJ pronominal	138
wún	3.SUBJ pronominal	138
né	1SG.DO	138
xé	2.MASC.DO	138
ǰí	2.FEM.DO	19, 138
jí	3SG.DO	19, 138
ré	3SG.FEM.DO	19, 138
məní	1PL.DO	19, 138
mái	2PL.DO	19, 138
wúní	3PL.DO	19, 138
nó	1SG.IO	21, 138
wò	2SG.IO	21, 138
ǰì	2SG.FEM.IO	21, 136
nì	3SG.MASC.IO	21, 136
rò	3SG.FEM.IO	21, 138
mù	1PL.IO	21, 138
màì	2PL.IO	21, 138
wù	3PL.IO	21, 138
ko	reflexive	21
mà	hortative	59
tə	ventive	73–75

Tigre (Leslau 1945)

Form	Label	Pages
tə	reflexive-passive	11
ʔa	causative	12
ʔatta:	causative factitive	13–14
ä	3SG.MASC.SUBJ	5–8
ät	3SG.FEM.SUBJ	5–8
ka	2SG.MASC.SUBJ	5–8

Form	Label	Pages
ki	2SG.FEM.SUBJ	5–8
ko	1SG.COMMON	5–8
äw	3PL.MASC.SUBJ	5–8
äja	3PL.FEM.SUBJ	5–8
kum	2PL.MASC	5–8
køn	2PL.FEM	5–8
na	1PL.COMMON	5–8
ʔi	negation	7
ju	3SG.MASC	187–189
ja	2SG.FEM	187–189
kka	2SG.MASC	187–189
kki	2SG.FEM	187–189
nni	1SG.COMMON	187–189
jom	3PL.MASC	187–189
jän	3PL.FEM	187–189
kkum	2PL.MASC	187–189
kken	2PL.FEM	187–189
nna	1PL.COMMON	187–189
ka	complex perfect	73–74

Koho (Nguyen 1973)

Form	Label	Pages
ton	causative	21–31
go	supernatural	32–34
tam	reciprocal	51
ɐ	negative	57

Krongo (Reh 1985)

Form	Label	Pages
m	FEM agreement	184 ff.
n	NEUT agreement	184 ff.
k	PL agreem	184 ff.
ák	preterite	188ff.
à:k	IMP.PL	195–197
t	hortative	197 ff.
ì	frequentative	201–209
ŋ	transitivizer	211 ff.
àkà	benefactive	221–223
àŋ	benefactive	224–225
i	directive	225–226
ká	iterative	227–228

Form	Label	Pages
àcá	venitive	227
àtíjí	passive	228–230
ácá	passive	228–230
í	mediopassive	230–231
áncá	mediopassive	230–231
ò:nó	reflexive	233
tíj	bound PERSPRO	167–168, 187
tí	bound PERSPRO	167–168, 187
tú	bound PERSPRO	167–168, 187
cá	bound PERSPRO	167–168, 187
tókwa	bound PERSPRO	167–168, 187
táàj	bound PERSPRO	167–168, 187
tíní	bound PERSPRO	167–168, 187
ànà	intransitivizer	214–221

Maidu (Shipley 1964)

Form	Label	Pages
ti	causative	40–41
ʔus	reflexive	41
jo	repetitive action	41
toto	reciprocal action	41
đoj	upward	42
je	hither	42
kit	downward	42
k'oj	away from here	42
men	negative	44
ḥós	completive	44
c'íj	unable	44
đój	inchoative	44
nú	durative	44
tí	for the sake of	44
ḥéw	a little more	44
c'ój	quotative	45
wéw	evidential	45
k'í	seems to me	45
sí	1	45–46
ano	2	45–46
n	3	45–46
ʔi	SG	45–46
h'a	DU	45–46
h'e	PL	45–46
ʔám	indicative	46

Form	Label	Pages
k	present-past	46–47
mak	future	47
ʔús	habitual	48
k'en	periphrastic subjunctive	48–49, 54
ʃ	optative	49
iʔi	monitive	49
æ	intensive	49–50
a	hortatory	50
k'ade	interrogative	50–51
pi	imperative I	51
paďá	imperative II	51–52
paʔáje	in ancient times	52
mini	if, when, etc.	52
c'et	none	52
wono	long ago	52
ʔ	past	48

Bari (Spagnolo 1933)

Form	Label	Pages
ta	plural	111
a	past	105
ja	passive	107–110
i	imperative	111–115, 136 ff., 231, 255
dju	emphatic	133–137, 161 ff.
un	un	143ff.
kin	dative	145 ff., 155, 165 ff
raʔ	centrifugal-linear	146 ff., 155, 165 ff
rikin	instrumental	149 ff., 157 ff., 176
tɔ	causative/reciprocal	157 ff., 176
n	inchoative	176–177

Ngambay (Vandame 1963)

Form	Label	Pages
k	durative	95
m	1SG	85 ff., 98 ff.
dʒ	1PL.SUBJ	85 ff., 98 ff.
dʒe	plural	98 ff.
d	3PL.SUBJ	85, 98
ne	reflexive	85, 90–91, 98
m	1SG.OBJ	86 ff.

Form	Label	Pages
i	2SG.OBJ	86 ff.
e	3SG.OBJ	86–90
si	1/2PL.OBJ	86
de	3PL.OBJ	86–90
ne	reflexive object	86–91

Abbreviations: 1/2/3/4 1st/2nd/3rd/4th person; ABS absolute; AGR agreement; ANIM animate; DEF definite; DO direct object; DU dual; ERG ergative; EXCL exclusive; FEM feminine; HUM human; IMP imperative; INAN inanimate; INCL inclusive; INDEF indefinite; IO indirect object; ITR intransitive; MASC masculine; NEUT neuter; NHUM non-human; NOM nominative; OBJ object; OBJV objective; PERSPRO personal pronoun; PL plural; POSS possessive; PRES present; SG singular; SUBJ subject; TR transitive; VN verbal noun.

References

- Alderete, John (1999). Morphologically governed accent in Optimality Theory. Doctoral dissertation, University of Massachusetts, Amherst.
- Alphonse, Ephraim S. (1956). *Guaymí Grammar and Dictionary with some Ethnological Notes*. (Bureau of American Ethnology Bulletin, 1962.) Washington: Government Printing Office.
- Barker, Muhammad Abd-al-Rahman & Aqil Khan Mengal (1969). *A Course in Baluchi*. 2 volumes. Montreal: Institute of Islamic Studies, McGill University.
- Benton, Richard A. (1971). *Pangasinan Reference Grammar*. Honolulu: University of Hawaii Press.
- Boas, Franz & Ella Deloria (1941). Dakota grammar. *Memoirs of the National Academy of Sciences*, 23-2: 1–183.
- Bright, William (1957). *The Karok Language*. (University of California Publications in Linguistics, 13.) Berkeley: University of California Press.
- Browman, Catherine P. & Louis M. Goldstein (1992). Articulatory phonology: An overview. *Phonetica* 49: 155–180.
- Buechel, Eugene (1939). *A Grammar of Lakota: The Language of the Teton Sioux Indians*. Saint Francis, S.D.: Rosebud Educational Society.
- (1983). *A Dictionary/Oie Wowapi Wan of Teton Sioux: Lakota–English: English/Lakota/Lakota–Ieska: Ieska–Lakota*. Pine Ridge, S.D.: Red Cloud Indian School, Holy Rosary Mission.
- Bybee, Joan (1997). Semantic aspects of morphological typology. In Joan Bybee, John Haiman, & Sandra Thompson (eds.), *Essays on Language Function and Language Type*, 25–37. Amsterdam: Benjamins.
- (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.
- (2003). Mechanisms of change in grammaticization: The role of frequency. In Richard D. Janda & Brian D. Joseph (eds.), *Handbook of Historical Linguistics*, 602–623. Oxford: Blackwell.
- Bybee, Joan, Revere Perkins, & William Pagliuca (1994). *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Chomsky, Noam & Morris Halle (1968). *The Sound Pattern of English*. New York: Harper & Row.
- Dressler, Wolfgang Ullrich (1985). *Morphonology*. Ann Arbor: Karoma.
- Dryer, Matthew (1989). Large linguistic areas and language sampling. *Studies in Language* 13: 257–292.
- Dunkel, George (2001). The sound-systems of Proto-Indo-European. *Journal of Indo-European Studies* 40: 1–14.

- Fortescue, Michael (1980). Affix ordering in West Greenlandic derivational processes. *International Journal of American Linguistics* 25: 259–278.
- (1984). *West Greenlandic*. London: Croom Helm.
- Furbee-Losee, Louanna (1976). *The Correct Language – Tojolabal: A Grammar with Ethnographic Notes*. New York: Garland.
- Gamkrelidze, Thomas V. & Vyacheslav V. Ivanov (1973). Sprachtypologie und die Rekonstruktion der gemeinidg. Verschlüsse. *Phonetica* 27: 150–156.
- (1995). *Indo-European and the Indo-Europeans*. Berlin: Mouton de Gruyter.
- Greenberg, Joseph (1966). *Language Universals, with Special Reference to Feature Hierarchies*. The Hague: Mouton.
- Hale, Kenneth L. (1959). A Papago grammar. Doctoral dissertation, Indiana University.
- Heine, Bernd & Mechthild Reh (1984). *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske.
- Hershberger, Ruth (1964). Notes on Gugu-Yalanji verbs. In Richard Pittman & Harland Kerr (eds.), *Papers on the Languages of the Australian Aborigines* (Occasional Papers in Aboriginal Studies, 3), 35–54. Canberra: Australian Institute of Aboriginal Studies.
- Hewitt, Brian G. (1979). *Abkhaz*. (Lingua Descriptive Studies, 2.) Amsterdam: North-Holland.
- Hoffmann, Carl (1963). *A Grammar of the Margi Language*. London: Oxford University Press.
- Hopper, Paul J. (1973). Glottalized and murmured occlusives in IE. *Glossa* 7: 141–166.
- Hopper, Paul J. & Elizabeth Closs Traugott (2003). *Grammaticalization*. 2nd edition. Cambridge: Cambridge University Press.
- Householder, Fred W., Kostas Kazazis, & Andreas Koutsoudas (1964). *Reference Grammar of Literary Dhimotiki*. Bloomington: Indiana University.
- Hutchison, John (1981). *The Kanuri Language: A Reference Grammar*. Madison: African Studies Program, University of Wisconsin.
- Jakobson, Roman (1966 [1990]). Quest for the essence of language. *Diogenes* 51: 21–37. [Reprinted in Linda R. Waugh & Monique Monville-Burston (eds.), *On Language: Roman Jakobson*, 407–421. Cambridge, Mass.: Harvard University Press.]
- Kopeseć, Michael F. (1975). Los elementos verbales y sustantivos y la oración en guaymí. In Stephen H. Levinsohn (ed.), *Lenguas de Panamá*, Volume 2: *Observaciones preliminares sobre los sistemas gramaticales de las lenguas chibchas*. Panamá: Instituto Nacional de Cultura, Dirección del Patrimonio Histórico & Instituto Lingüístico de Verano.
- Kuipers, Aert H. (1974). *The Shuswap Language: Grammar, Texts, Dictionary*. The Hague: Mouton.
- Lehmann, Christian (1982 [1995]). *Thoughts on Grammaticalization: A Programmatic Sketch*. Volume 1. (Arbeiten des Kölner Universalien-Projekts, 48.) Köln: Universität zu Köln. Institut für Sprachwissenschaft. [München: Lincom Europa.]
- (1985). Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* 20: 303–318.
- Leman, Wayne (1980). *A Reference Grammar of the Cheyenne Language*. (Occasional Publications in Anthropology, Linguistic Series, 5.) Greeley: University of Northern Colorado Museum of Anthropology.
- Leslau, Wolf (1945). *Short Grammar of Tigré*. (Publications of the American Oriental Society, offprint series, 18.) New Haven: American Oriental Society.
- Lukas, Johannes (1937). *A Study of the Kanuri Language: Grammar and Vocabulary*. London: Oxford University Press.
- Manley, Timothy (1972). *Outline of Sre Structure*. Honolulu: University of Hawai'i Press.
- Mason, J. Alden (1950). *The Language of the Papago of Arizona*. Philadelphia: University Museum, University of Pennsylvania.
- Mathiot, Madeleine (1973–78). *A Dictionary of Papago Usage*. 2 volumes. (Indiana University Publications Language Science Monographs.) Bloomington: Indiana University.

- McCarthy, John & Alan Prince (1995). Faithfulness and reduplicative identity. In Jill Beckman & Suzanne Urbanczyk (eds.), *Papers in Optimality Theory* (University of Massachusetts Occasional Papers in Linguistics, 18), 1–126. Amherst: University of Massachusetts Press.
- Merrifield, William R. (1963). Palantla Chinantec syllable types. *Anthropological Linguistics* 5-5: 1–16.
- (1968). *Palantla Chinantec Grammar*. México: Instituto Nacional de Antropología e Historia de México.
- Mowrey, Richard & William Pagliuca (1995). The reductive character of articulatory evolution. *Rivista di Linguistica* 7: 37–124.
- Mühlhäusler, Peter (1985). Inflectional morphology in Tok Pisin; Syntax of Tok Pisin. In Stephen A. Wurm & Peter Mühlhäusler (eds.), *Handbook of Tok Pisin (New Guinea Pidgin)* (Pacific Linguistics, C-70), 335–421. Canberra: Australian National University.
- Najlis, Elena Lidia (1966). *Lengua Abipona*, Volume 1. (Archivo de Lenguas Precolombianas, 1.) Buenos Aires: Centro de Estudios Lingüísticos, Universidad de Buenos Aires.
- Newman, Paul (1974). *The Kanakuru Language*. Leeds: Institute of Modern English Language Studies.
- Nguyen, Van Hoan (1973). Ko'ho Sre affixation. Master's thesis, University of Saigon.
- Ohala, John (1993). The phonetics of sound change. In Charles Jones (ed.), *Historical Linguistics: Problems and Perspectives*, 237–278. London: Longman.
- Patz, Elizabeth (1982). A grammar of the Kuku Yalanji language of north Queensland. Doctoral dissertation, Australian National University.
- Perkins, Revere D. (1989). Statistical techniques for determining language sample size. *Studies in Language* 13: 293–315.
- Plank, Frans (1998). The co-variation of phonology with morphology and syntax: A hopeful history. *Linguistic Typology* 2: 195–230.
- Poppe, Nicholas N. (1960). *Buriat Grammar*. (Indiana University Publications: Uralic and Altaic Series, 2.) Bloomington: Indiana University; The Hague: Mouton.
- Raz, Shlomo (1983). *Tigre Grammar and Texts*. (Afroasiatic Dialects, 4.) Malibu: Undena.
- Reh, Mechthild (1985). *Die Krongo-Sprache*. Berlin: Reimer.
- Renck, Gunther L. (1975). *A Grammar of Yagaria*. (Pacific Linguistics, B-40.) Canberra: Australian National University.
- (1977). *Yagaria Dictionary*. (Pacific Linguistics, C-37.) Canberra: Australian National University.
- Rice, Keren (1989). *A Grammar of Slave*. Berlin: Mouton de Gruyter.
- Sapir, Edward (1921). *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace & World.
- Shipley, William F. (1964). *Maidu Grammar*. (University of California Publications in Linguistics, 41.) Berkeley: University of California Press.
- Skalička, Vladimír (1979). *Typologische Studien*. Braunschweig: Vieweg.
- Snyman, J. W. (1970). *An Introduction to the !Xū (!Kung) Language*. (University of Cape Town School of African Studies Communication, 34.) Cape Town: Balkema.
- Sorensen, Arthur P., Jr. (1969). The morphology of Tucano. Doctoral dissertation, Columbia University.
- Spagnolo, Lorenzo M. (1933). *Bari Grammar*. Verona: Missioni Africane.
- Subrahmanyam, P. S. (1971). *Dravidian Verb Morphology*. Annamalainagar: Annamalai University.
- Thomas, Elaine (1978). *A Grammatical Description of the Engenni Language*. Arlington, Tex.: Summer Institute of Linguistics.
- University of Colorado Lakhota Project (1974). Lakhota sketch. Unpublished manuscript. Boulder, Col.: Department of Linguistics, University of Colorado.
- (1976a). *Beginning Lakhota*. 2 volumes. Boulder, Col.: Department of Linguistics, University of Colorado.

- (1976b). *Elementary Bilingual Dictionary: English–Lakhóta, Lakhóta–English*. Boulder, Col.: Department of Linguistics, University of Colorado.
- (1976c). *Lakhóta Wayawapi/Lakhóta Readings*. Boulder, Col.: Department of Linguistics, University of Colorado.
- Ussishkin, Adam (2000). Root-and-pattern morphology without roots or patterns. *North East Linguistic Society* 30: 655–670.
- Vandame, Charles (1963). *Le Ngambay-Moundou: Phonologie, grammaire et texts*. Dakar: Institut Français d’Afrique Noir.
- Voegelin, Charles F. & Florence M. Voegelin (1978). *Classification and Index of the World’s Languages*. New York: Elsevier.
- West, Birdie (1980). *Grammatica popular del Tucano*. Bogotá: Ministerio de Gobierno, Instituto Lingüístico de Verano.
- Willerman, Raquel (1994). The phonetics of pronouns: Articulatory bases of markedness. Doctoral dissertation, University of Texas at Austin.
- Winfield, W. W. (1928). *A Grammar of the Kui Language*. Calcutta: Asiatic Society of Bengal.
- Woolford, Ellen B. (1979). *Aspects of Tok Pisin Grammar*. (Pacific Linguistics, B-66.) Canberra: Australian National University.
- Young, Philip D. (1986). Ngawbere numeral classifiers. In Scott DeLancey & Russell Tomlin (eds.), *Proceedings of the Second Annual Meeting of the Pacific Linguistics Conference*, 489–502. Eugene: Department of Linguistics, University of Oregon.
- (1987). From ergative to active in Guaymí. Paper presented at the 86th annual meeting of the American Anthropological Association, Chicago.
- Zepeda, Ofelia (1983). *A Papago Grammar*. Tucson: University of Arizona Press.

Copyright of Linguistic Typology is the property of Walter de Gruyter GmbH & Co. KG.. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.