What's With All These Long Words Anyway?*

KEVIN RUSSELL

University of Manitoba

1. INTRODUCTION

The languages of the Americas have long been famous for their tendency to have extremely complex word-structure. As put by one of the leading American linguists of the nineteenth century, William Dwight Whitney:

> This type is called the incorporative or polysynthetic. It tends to the excessive and abnormal agglomeration of distinct significant elements in its words; whereby, on the one hand, cumbrous compounds are formed as the names of objects, and a character of tedious and time-wasting polysyllabism is given to the language.　　　　(1867:348)

Even leaving aside the negative value judgments, the reputation for long words may be largely undeserved, since the question of what exactly constitutes a word in American languages has not received the attention it ought to have. When linguists do address the issue, the criteria they use for lumping formatives together into words in polysynthetic languages are often quite different from those used in justifying word boundaries in more familiar languages.[1]

　　Missionary grammarians were the first to record many languages of the Americas, and thus the first to struggle with the question of where to put word boundaries. While it is hard to be certain what motivated the decisions of these first grammarians (a slavish devotion to the categories of Latin is evident in more than a few), it is perhaps not astonishing that so many came up with results that appear highly exotic. In the hands of early linguists, this exoticness quickly became the *sine qua non* of American languages. In 1819, Duponceau declared that polysynthesis prevailed "with little variation among the aboriginal natives of America, from Greenland to Cape Horn." (cf. Haas 1969). A long line of nineteenth century linguists adopted

---

** Abbreviations used in this paper: 1=first person, 3=third person, AGT=agent, NEG=negation, OBJ=object, PERF=perfective, pl=plural, V=verb root, VII=intransitive inanimate verb, sg=singular.

[1] I will use the term "formative" in a pre-theoretical sense to refer to a fundamental building block of a word or sentence, i.e., any element of an utterance which has some consistent semantic or grammatical properties. Informally, a "formative" in this sense can be seen as a "morpheme", but I intend the term simply to be neutral with respect to questions such as: Does the formative have a consistent meaning? Is the formative an affix or an independent word?

Duponceau's conviction of the universality of polysynthesis in the Americas (often coupled with a belief that it resulted from all languages of the Americas being genetically related).[2]

In the twentieth century, many linguists carried on the conviction that American languages were mostly polysynthetic, though often with more subtlety and self-consciousness. In particular, advances in linguistic theory made it more difficult to assume that "word" was a self-evident notion. The growing tension is well represented by Sapir, who defended the pre-theoretical validity of the "word" concept, and by Bloomfield, who questioned it. According to Sapir (1921:32): "The word is merely a form, a definitely molded entity that takes in as much or as little of the conceptual material of the whole thought as the genius of the language allows it." One might well question whether it is the genius of the language doing the allowing, or the genius of the linguist. Trying to answer this objection, Sapir writes:

> But is the word, one may object, as much of an abstraction as the radical element? Is it not as arbitrarily lifted out of the living sentence as is the minimal conceptual element out of the word? Some students of language have, indeed, looked upon the word as such an abstraction, though with very doubtful warrant, it seems to me. It is true that in particular cases, especially in some of the highly synthetic languages of aboriginal America, it is not always easy to say whether a particular element of language is to be interpreted as an independent word or as part of a larger word. These transitional cases, puzzling as they may be on occasion, do not, however, materially weaken the case for the psychological validity of the word. Linguistic experience, both as expressed in standardized, written form and as tested in daily usage, indicates overwhelmingly that there is not, as a rule, the slightest difficulty in bringing the word to consciousness as a psychological reality. No more convincing test could be desired than this, that the naive Indian, quite unaccustomed to the concept of the written word, has nevertheless no serious difficulty in dictating a text to a linguistic student word by word; he tends, of course, to run his words together as in actual speech, but if he is called to a halt and is made to understand what is desired, he can readily isolate the words as such, repeating them as units. He regularly refuses, on the other hand, to isolate the radical or grammatical element, on the grounds that it "makes no sense".   (1921: 33-34)

Bloomfield (1914) was less sanguine about the ease of identifying words:

> This procedure is makeshift, for it has long been recognized that the first and original datum of language is the sentence, — that the individual word is the product of a theoretical reflection which ought not to be taken for granted, and further, that the grouping of derived and inflected words into paradigms, and the abstraction of roots, stems, affixes, or other formative processes, is again the result of an even more refined analysis. It needs but little scientific reflection to make us realize that the grammarian ought by no means to extract such products with magic suddenness, live and wriggling, out of the naïve speaker's hat.

---

[2] For an overview of nineteenth century linguistic notions concerning the polysynthesis and relatedness of American languages, see Haas (1969).

It should be noted, however, that after all of the theoretical reflection and refined analysis was over with, the words postulated by Bloomfield tended to be every bit as long as those that were given to Sapir (after he successfully made his consultants understand what he desired).

In this paper, I question the criteria that have been used for "wordhood" in the polysynthetic languages of the Americas and elsewhere. I argue that accepting the word boundaries of traditional orthography at face value has resulted in distorted analyses.

2. KINDS OF WORDS

Part of the problem in using the term "word" comes from the fact that it has traditionally been used for several fundamentally different kinds of entities. As a rough indication of what kinds of "words" may exist, we can begin with the following list, based on that of DiSciullo and Williams (1987):

(1) a. orthographic word
     b. free form (minimal unit that can stand as a complete utterance)
     c. syntactic atom (X^0)
     d. phonological/prosodic word (PWd)
     e. morphological word
     f. listed item (a unit with enough idiosyncrasies that it must be memorized)

Except insofar as they shed light on one of the other kinds of words, orthographic words are of no intrinsic interest to theoretical linguistics, so we will concentrate our attention on the other definitions. The ability to stand alone as a complete utterance is more interesting, but there are enough other factors which can affect a linguistic community's idea of what constitutes an acceptable utterance that linguists must be more careful than they have traditionally been in accepting this diagnostic at face value. X^0s and PWds are such firmly established and useful elements of syntactic and phonological theory, respectively, that their inclusion needs no defence. Intuitively, the idea of "morphological word" seems to be meaningful, but attempts to formalize it fall all too easily into circularity: morphological words are the things created by the morphological component, the morphological component is the thing that creates morphological words.

"Word" (without qualification) has been a useful concept in linguistics to the extent that the definitions in (1b-f) pick out the same objects: in the unmarked case, a X^0 in the syntax is realized by a prosodic word constituent in the phonology, can stand alone as a complete (if elliptical) utterance, can be seen as the output of a morphological component, and will through the history of the language often gather idiosyncrasies which will require it to be memorized as a listed item. The correlations are too strong to be dismissed as coincidence, but if we try to explain them as inevitable consequences of "the architecture of the theory", we immediately run up against the myriad of cases where the correlation is imperfect. Entire vocabularies have grown up around the mismatches between the definitions: idiom, clitic, level 2 affix, compound,

incorporation, and so on. Attempts to formalize the "word" (without qualification) can take one of two approaches: i) assert that the word *is* the overlap between the definitions (this requires so many escape hatches for the exceptions that the resulting theory risks being vacuous), or ii) assert that the word *is* really one of the definitions (say X^0 or PWd) and to explain the correlations with other definitions in some other way.

A better approach is to allow that each of the definitions defines a real entity at a different level of grammar. An X^0 is a well-defined entity of syntax. A PWd is a well-defined entity of phonology. The correlation between them can be the result of some kind of default interface principles which usually but not always cause them to coincide (e.g., the boundary mapping conditions of Selkirk (1986) or Nespor and Vogel (1986), Generalized Alignment constraints in Optimality Theory, interface conditions in Autolexical Syntax, and so on).

If we take the latter approach, it is meaningless to say that language X has short words, or that American languages typically have long words. We must specify what kind of word we are talking about. We could say that American languages have long *phonological* words—that would be meaningful, though probably false.[3] We could say that American languages have "large" syntactic words, perhaps in the sense of having a number of adjoined functional and lexical heads assembled via head movement—that too would be meaningful, though empirically unproven.[4] We could even say that American languages have large morphological words, perhaps in the sense of containing many morphemes or having undergone many morphological processes—that too would be meaningful (so long as one assumes the existence of a morphological component). It is, however, a great deal more problematic empirically, due to the inherent circularity in the definition of the morphological word/component. If I were to refuse outright to believe that formatives A and B belonged to the same morphological word, there would be very little clear morphological (not syntactic, not phonological) evidence that you could bring to bear to change my mind.

Unfortunately, it is relatively rare to find careful distinctions between kinds of words in the literature on American languages. It is not uncommon to find "words" discussed as self-evident facts, with no clue in sight as to what the author's criteria might have been. Indeed, as mentioned, many still accept traditional orthographic divisions at face value, assuming uncritically that they represent real linguistic units.

Before embarking on a general discussion of the validity of the criteria that have been used in identifying words in a polysynthetic language, I would like to give a sketch of one recent discussion of the "word" in a polysynthetic language which *does* carefully distinguish many of the different kinds of words in (1) and argue explicitly for a particular segmentation of the sentence.

---

[3] For arguments that phonological words in various North American languages are shorter than the traditional orthographic word, see Dyck (1994), McDonough (1990), Rice (1994), Russell (to appear).

[4] Sandalo (this volume) is an example of the kind of evidence that would have to be marshalled in order to demonstrate that two formatives are dominated by the same X^0.

3. FOLEY (1991) ON YIMAS

Yimas is a polysynthetic and possibly non-configurational language of Papua-New Guinea described by Foley (1991).

Foley distinguishes between a phonological word and a morphosyntactic word.[5] The phonological word serves as the domain for the application of several phonological processes: vowel epenthesis, R-spreading, palatalization, final cluster simplification and stress assignment. His initial definition of the morphosyntactic word, however, borders on the circular:

> ...a word is a string of morphemes that behaves as a single unit with respect to morphological processes in the language. The strictly morphological processes of a language, of course, are a language-specific fact: some languages such as Vietnamese may have no morphological processes at all; others, like Yimas, have an exuberance of these. (Foley 1991:81)

Morphosyntactic and phonological words need not coincide. Indeed, one morphosyntactic word may contain several phonological words, and often does. In (2a), Foley argues that the three verb roots, the conjunction formative *mpi* 'and', the agreement prefixes and the suffix all form a single morphosyntactic word, though the three primary stresses show that there are three different phonological words (2b).

(2)   a.   ya-mpu-park-mpi-kapik-mpi-wark-t
           Vpl.OBJ-3pl.AGT-split-and-break-and-tie-PERF
           'They split the branches, broke them and tied them.'

      b.   [ya - mpu - [párk - mpi]] - [kápik - mpi] - [wárk - t]

In general, Foley's arguments for a morphological word come from the rigidity of ordering and adjacency between formatives: "This form is morphologically a single composite unit. The morphemes must occur in this order and no other... Further, no other morphemes may be inserted internal to this sequence." The language-specific process of negation also seems to delimit a word-sized unit:

> Negation is an especially good diagnostic for words, for it affects both the first morpheme of a word and the last. Negation is marked by the prefix *ta-* which occupies the first prefixal slot of the verb, usurping the normal occupier of this slot, a pronominal prefix, and shifting this to the final suffixal slot of the word.   (Foley 1991:82)

To illustrate this, in (3), we see the pronominal *p-* in the first position of the verbal complex. In the negated sentence in (4), the negative marker *ta-* occupies the first position and the pronominal (in its *-m* allomorph) is forced to the end of the verbal complex.

---

[5] Foley is aware of Di Sciullo and Williams' distinction between morphological and syntactic words, but decides not to adopt it in his description.

(3)   tampaym   p-ka-apca-t
         hook.VII.sg VII.sg.OBJ-1sg.AGT-hang.up-PERF
         'I hung up the hook.'

(4)   tampaym   ta-ka-apca-r-m
         hook.VII.sg NEG-1sg.AGT-hang:up-PERF-VII.sg.OBJ
         'I didn't hang up the hook.'

Some of Foley's arguments for wordhood come close to the "free form" definition. Even if a formative is an independent PWd, if it can't stand alone in an utterance (because it obligatorily selects for a predicate complement, for example), this is enough to consign it to morphology rather than syntax:

> Of special interest is the continuative aspect marker *yakal-*. While a phonological word, it can never function as a verb root or on its own. It is restricted to occurring before a following verb or verb compound as the means of expressing its aspectual character. The great majority of morpheme types which constitute a grammatical word are like this, and for this reason I follow the convention of writing these as a single orthographic word. (1991: 87)

One may decide to criticize Foley's reasons for dividing words where he does, but only because we know what his reasons are. Foley is explicit in what his criteria for grammatical wordhood are. He is explicit that his grammatical word is not coextensive with the phonological word or with other possible definitions of word. Foley's work on Yimas is all too rare in this respect. Too often, we are simply presented with word-sized units, produced "with magical suddenness, live and wriggling, out of the naive speaker's hat" — or perhaps out of air, we can't tell.

4. CRITERIA FOR MORPHOLOGICAL WORDHOOD

We turn now to a general discussion of the kinds of criteria that have been used to argue that certain strings of formatives were assembled by the morphological component. In general, many arguments for wordhood come from an assumption that can be summarized as "If it looks like syntax, it's really morphology." I take syntax to be the study of the order, adjacency and constituency of elements in a sentence. Yet interesting constraints on order and adjacency are now often taken as the unique purview of morphology.

The arrangement of elements in syntax often shows a great deal of freedom. For example, *John ran away* can be rephrased as *Away ran John*. Beginning from this observation, it is all too easy a step to conclude that syntax is *always* like this, and that any time the order of elements fails to show this degree of anarchy, they must have been arranged by something other than syntax, namely morphology.[6] Similarly, syntax often shows considerable freedom in adjacency. In English a determiner and noun can be separated by an adjective. This observation often

---

[6] Ironically, we also often find the reverse. For example, we saw above that the Yimas object pronoun and the verb stem can occur in different orders under certain circumstances. Foley uses this rearrangeability as part of an argument that the two are ordered by *morphology*.

gives rise to the assumption that any two formatives that cannot be similarly separated necessarily belong to the same word and that their adjacency is enforced by principles of morphology.

A more subtle variant of the "if it looks like syntax, it's really morphology" assumption involves different orders occurring at different scales or with different sub-types of elements. If a full noun phrase direct object usually precedes the verb but a pronoun direct object follows it, it is tempting to divide the labour between modules: syntax puts the direct object before the verb while morphology uses a direct object "suffix".[7] (If we take this kind of approach to the extreme, we could end up concluding that the difference in English between *Give away the money* and *Give it away* is really a difference between syntax and morphology.)

Coupled with different behaviour at different scales, arguments from restrictions on order and adjacency can be especially seductive. In Yimas, the ordering and adjacency of elements really does have a wide degree of freedom at the scale of the entire sentence but rigidity at a smaller scale (the neighbourhood of the verb stem). One obvious explanation for this is that different modules are responsible for the different scales. But we should not be blinded to the fact that the same data can in principle be accounted for within a single module.[8]

A summary of some of the best-known criteria for morphology as opposed to syntax is given in (5). (To this could be added some of the differences discussed in Zwicky (1992), which deserve serious consideration but which do not offer much in the way of immediate empirical guidance to analysts.)

(5)   a.   ordering relations tend to be more strict in morphology

---

[7] Anderson's (1992) arguments for the syntax/morphology in K^{w}ak^w'ala illustrate the dangers of this approach. He begins by declaring that "the relative order of the main constituents of a clause is quite rigid" (referring skeptics to an earlier paper), then proceeds to discuss a number of counter-examples. His conclusion is that the counter-examples are morphological rather than syntactic (which I believe to be appropriate in this case), but the possibility that his initial assumption of rigid order might be wrong and that the counter-examples might have a purely syntactic explanation never seems to have crossed his mind.

[8] One syntactic analysis that can account for both rigidity and freedom simultaneously is the syntactic interpretation of the Pronominal Argument Hypothesis for non-configurational languages, that is, an account where the verb's theta-roles are not discharged morphologically by agreement affixes but by pronominal elements occupying their usual argument positions in a syntactic structure. Within the core sentence (an IP or CP), we would expect the order and adjacency restrictions on elements to be relatively rigid. We would not expect to find the same kind of rigidity outside the core sentence, in positions which are adjoined to IP or CP or in the specifier or complement positions of discourse related projections such as Topic Phrase or Focus Phrase. The most important problem remaining for such an analysis is to motivate why full noun phrases and material of similar bulk cannot surface in the rigid neighbourhood but *must* end up outside the core sentence. Whether or not this problem has been successfully solved by proposals such as Baker (1996) or Russell and Reinholtz (1997), it should be clear that a purely syntactic account is at least conceivable, and that different behaviour at different scales can no longer be taken as iron-clad evidence of a division of labour between syntax and morphology.

b. adjacency constraints tend to be more strict in morphology

c. morphology tends to allow more irregularities than syntax

The crucial point to note is that most of these criteria, including the ones based on the "if it looks like syntax, it's really morphology" assumption, have to be phrased as *tendencies*. Treating them as anything stronger than tendencies will have the result of eviscerating the subject matter of syntax, leaving it little more than the ability to perform unconstrained apposition of words.[9]

But if the boundary between the syntactic and morphological modules is merely a matter of tendencies, the immediate empirical consequence is that we are unable to assign confidently any particular phenomenon of a language to either syntax or morphology. Since there are uncontroversially syntactic phenomena that go against one or two or three of these tendencies and to that extent appear somewhat morphological, there is no *a priori* theoretical reason why there should not be a syntactic phenomenon that goes against most or even all of the tendencies. To put it bluntly, since syntax has all the resources it needs to imitate morphology, there is no certain method to distinguish between morphology and syntax imitating morphology.

The problem is compounded further if we take seriously some of the more extreme abilities claimed for the "word-internal" structure of polysynthetic languages. Payne (1990) gives the following examples from the South American languages, Capanahua and Yagua:

(6)   *Capanahua*

a. pi-yama-ma-catsihqu-yama-hi
    eat-NEG-cause-desiderative-NEG-present
    'He$_i$ doesn't want to make him$_j$ not eat it.'

b. pi-yama-ma-catsihqu-ma-yama-hi
    eat-NEG-cause-desiderative-cause-NEG-present
    'He$_i$ doesn't want him$_j$ to make him$_k$ not eat it.'

(7)   *Yagua*

a. sa-júnay-muuy-tániy-níí Mario
    3sg-cry-completive-cause-3sg
    'He makes Mario stop crying.'

b. sa-júnay-tániy-muuy-níí Mario
    3sg-cry-cause-completive-3sg
    'He stops making Mario cry.'

---

[9] Heath (1984) actually comes quite close to this in his grammar of the Australian polysynthetic language, Nunggubuyu.

In these examples, we see the morphological systems of the languages performing complex manipulations of scope and coreference relationships. In other words, in addition to syntax having all the power it needs to imitate morphology, it seems that polysynthetic morphology has all the power it needs to imitate syntax.[10]

This ability of morphology and syntax to imitate each other might cause one to doubt whether there is really a difference between them. But it is not my purpose here to argue that there is no morphological component separate from syntax. My point is that there are no firm grounds that one can use in practice for assigning a division of labour between the two components. Given a string of formatives in a sentence [a b c d e f g h i j], there are no clear criteria — discovery procedures, if you will — which will lead a linguist to the conclusion that the sequence [d-e-f-g] forms a morphological word, as opposed to, say, [b-c-d-e-f-g-h]. Indeed, there are probably a few different groupings which are equally justifiable by the criteria in (5). What the criteria for morphology do allow a linguist to do is to decide on word-boundaries first and to argue for them after the fact.

Which leads us to the central question of what really motivates linguists in the initial assignment of "morphological" word boundaries. The motivation is obviously different for each linguist, and not always conscious. For some, it is clearly an acceptance of traditional orthographic word boundaries. For some, it is the expected correlation with the other definitions of word (e.g., PWds, listed items, free forms). For many, though, the choice is probably influenced by a subtle bias toward *expecting* long words.

Suppose you are a linguist working on an undescribed language of the Americas. If your past experience tells you that most American languages have long words, and you know that several related languages have long words, you are much more likely to "find" long words in your new language. This doesn't mean you are starting out with an intention to falsify data or to distort generalizations. But when you come face to face with one of the countless situations where either [d-e-f-g] or [c-d-e-f-g-h] are plausible groupings, a subtle bias can be enough to decide you in favour of the longer one. And, once [c-d-e-f-g-h] is chosen, the criteria of (6) can always be made to justify the choice.

---

[10] It is worth quoting here one of the speculative conclusions Payne draws from her overview:

> A further question raised by these data is: To what extent is the cognitive processing involved in the production of such verbs the same as, or different from the cognitive processing involved in the syntactic production of both simple and complex sentences? For example, is morphology in these languages more akin to syntax in an Indo-European language? It is commonly assumed that the possibilities in word formation are highly restricted and often 'frozen'. Syntax, on the other hand, is creative and allows for recursion. There are rules, but there are few 'frozen forms'. For example, it has been claimed that the notion of 'possible but non-existent word' is relevant for morphology, while 'possible but non-existent sentence' has no relevance for syntax (Scalise 1984, Halle 1973). But in Arawakan languages, where speakers can produce long and novel verb forms quite freely, is there such a thing as a 'possible but non-existent' verb?   (1990: 234)

There is statistical support for the bare facts, if not for my interpretation of unconscious bias. Perkins (1992) has amply documented the descriptive asymmetries in his ambitious (if misguided) attempt to statistically correlate language features with how culturally "advanced" a society is (as described by anthropologists). For example, linguists are more likely to describe the languages spoken by "less advanced" cultures as having bound pronouns (i.e., in the same orthographic word as the verb), and those of "more advanced" cultures as having independent pronouns (i.e., in separate orthographic words). Perkins accepts this statistical correlation as evidence that "less advanced" languages really are more likely to use bound pronouns, in short that language really does correlate with culture. In my view, the findings are real, but reveal nothing more surprising than that the biases of Western linguists correlate with the biases of Western anthropologists.

5. CONCLUSION

The major practical consequence is that we must take a Bloomfieldian rather than a Sapirian stance toward the orthographic words assigned by linguists to polysynthetic languages — treating them not as primary data, but as possibly (perhaps probably) flawed secondary analyses. This has consequences both for the analyses of individual languages and for cross-linguistic comparisons. In analyzing a single language, linguists should show a greater willingness to re-evaluate the segmentations of earlier linguists if that seems to allow for a better account of the language. For example, comparing the analyses of Potawatomi verb morphology by Anderson (1992), Halle and Marantz (1993) and Steele (1995), abstracting as far as possible away from the differences in their frameworks, I believe that the single factor that is most related to the simplicity or complexity of the analysis is the extent to which the authors do *not* try to account for the person prefixes using the same mechanisms as the verb suffixes — in short, the extent to which they reject the traditional Algonquianist orthographic tradition.

Even the most carefully-argued segmentations should not be automatically be accepted as "word" segmentations, where the "word" carries all the theoretical baggage it does from less synthetic languages. Foley's arguments for sub-dividing utterances where he does are no doubt well-founded and represent genuine insights into the structure of Yimas. Representing these breaks with spaces is convenient in discussing the structure of the language and therefore defensible. Calling the stretches between the spaces "words" is less useful. It should certainly not be assumed that Foley's segmentations have any of the properties usually attributed to words by various theories: that they are entirely assembled by the morphological component, that they are assembled syntactically by successive applications of head movement, that their internal structure is syntactically "inaccessible" to outside elements, that they are the appropriate domain for rules of lexical phonology, and so forth.[11]

---

[11] Phillips (1993) illustrates some of the insights that can be gained through revising aspects of Foley's analysis of Yimas.

Cross-linguistically, linguists must use extreme caution in trying to compare polysynthetic languages with each other or with other languages. Every linguist who has segmented a polysynthetic language into words has applied the criteria of (6) slightly differently, marshalling a different set of largely ad-hoc arguments for the segmentation. We must not assume that the stretches of formatives so defined are commensurable between languages: that the Yimas word defined by displacement of the object pronoun under negation is somehow the same kind of thing as the Menomini word defined by the morphophonological rule of initial change (Bloomfield 1961) or the Takelma word defined by Sapir by its ability to stand as a complete utterance, and so on. Linguists must be sensitive to the problems with the definition of word in polysynthetic languages before holding them up as putative counterexamples to putative linguistic universals and before trying to use them (as Perkins (1992) does) in inducing new universals.

REFERENCES

Anderson, Stephen R. 1992. *A-morphous Morphology*. Cambridge: Cambridge University Press.

Baker, Mark. C. 1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.

Bloomfield, Leonard. 1914. Sentence and word. *Transactions of the American Philological Society* 45: 65-75. [Reprinted in *A Leonard Bloomfield* Anthology, ed. Charles F. Hockett]

Bloomfield, Leonard. 1961. *The Menomini Language*, ed. Charles F. Hockett. New Haven: Yale University Press.

DiSciullo, Anna-Maria & Edwin Williams. 1987. *On the Definition of Word*. Cambridge, Mass: MIT Press.

Dyck, Carrie. 1994. The definition of 'word' in polysynthetic languages. *Proceedings of the 1993 Annual Conference of the Canadian Linguistic Association*, ed. Carrie Dyck, 187-203. Toronto Working Papers in Linguistics, University of Toronto.

Foley, William A. 1991. *The Yimas Language of New Guinea*. Stanford: Stanford University Press.

Haas, Mary R. 1969. Grammar or lexicon?: The American Indian side of the question from Duponceau to Powell. *International Journal of American Linguistics* 35: 239-255.

Halle, Morris. 1973. Prolegomena to a theory of word-formation. *Linguistic Inquiry* 4:3-16.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. *The View From Building 20,* ed. K. Hale & S. J. Keyser, 111-76. Cambridge, Mass.: MIT Press.

Heath, Jeffrey. 1984. *A Functional Grammar of Nunggubuyu*. Canberra: AIAS.

McDonough, Joyce. 1990. *Topics in the phonology & morphology of Navajo verbs*. Doctoral dissertation, University of Massachusetts, Amherst.

Nespor, Marina & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Payne, Doris. 1990. Morphological characteristics of lowland South American languages. *Amazonian linguisitcs: Studies in Lowland South American languages*, ed. Doris Payne, 213-242. Austin: University of Texas Press.

Perkins, Revere D. 1992. *Deixis, Grammar & Culture*. Amsterdam: John Benjamins.

Phillips, Colin. 1993. Conditions on agreement in Yimas. *MIT Working Papers in Linguistics 18: Papers on Case and Agreement 1*, ed. J.D. Bobaljik & C. Phillips, 273-312. Cambridge, Mass: Department of Linguistics and Philosophy, Massachusetts Institute of Technology.

Rice, Keren. 1994. The structure of the Slave (northern Athapaskan) verb. *Phonetics & Phonology 4: Studies in Lexical Phonology*, ed. Sharon Hargus & Ellen Kaisse, 145-171. New York: Academic Press.

Russell, Kevin. to appear. The word in two polysynthetic languages. *Phonological Studies on the Word*, ed. T.A. Hall & U. Kleinhenz. Amsterdam: John Benjamins.

Russell, K. & C. Reinholtz 1997. Nonconfigurationality and the syntax-phonology interface. *Proceedings of the Fifteenth West Coast Conference on Formal Linguistics*, ed. B. Agbayani & S.-W. Tang, 441-56. Center for the Study of Language and Information, Stanford University.

Sandalo, F. 1999. Categorical projection and phrase structure in Kadiwéu. This volume.

Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace & World.

Scalise, S. 1984. *Generative Morphology*. Dordrecht: Foris.

Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology* 3:371- 405.

Steele, S. 1995. Towards a theory of morphological information. *Language* 71:261-309.

Whitney, William Dwight. 1867. *Language & the Study of Language: Twelve Lectures in the Principles of Linguistic Science*. London: N. Trübner.

Zwicky, Arnold. 1992 Some choices in the theory of morphology. *Formal Grammar*, ed. Robert Levine, 327-371. Oxford: Oxford University Press.