

When Zero Isn't There

Author(s): Marianne Mithun

Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society (1986), pp. 195-211

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

When Zero Isn't There

Marianne Mithun
University of California, Santa Barbara

Zero forms of morphemes have often been assumed to be simply arbitrary formal alternatives within paradigms. More astute observers have noticed that zero markers do not appear at random points (Benveniste 1966). If one number marker is zero, for example, it is more likely to be the singular than the dual or plural. If one nominal case marker is zero, it is more likely to be the absolutive than the ergative. If one person marker is zero, it is more likely to be the third than the first or second.

Yet there is intriguing evidence that unmarked third persons may not necessarily be even zero. Compare the paradigms below from Mohawk, an Iroquoian language of Quebec, Selayarese, an Austronesian language of Indonesia, and Lakhota, a Siouan language of South Dakota.¹

1)	Mohawk	Selayarese	Lakhota
'I am sick'	<i>wakenonhwáktani</i>	<i>bambaŋa</i>	<i>mak^húže</i>
'you are sick'	<i>sanonhwáktani</i>	<i>bambaŋko</i>	<i>nik^húže</i>
'he is sick'	<i>rononhwáktani</i>	<i>bambaŋi</i>	<i>k^húže</i>

When asked how to say 'I am sick', Mohawk, Selayarese, and Lakhota speakers all reply with no hesitation, supplying the forms in the first line above. Asked how to say 'you are sick', they happily provide the forms in the second line. When asked how to say 'he is sick', the Mohawk and Selayarese speakers cheerfully give the third forms. The Lakhota speakers, however, look puzzled and ask, "Well who do you mean?" Although most Mohawk speakers cannot segment their language morphologically, they know that the pronoun 'he' is in their verb somewhere. Selayarese speakers know that there is a satisfactory equivalent to 'he' in their verb, even though there is no gender or number distinction. Lakhota speakers, on the other hand, know that there is no 'he' in theirs. Similar comments from speakers of a number of other languages indicate that this is by no means an isolated phenomenon.

The nature of pronominal affixes and clitics like those above has usually been clear to grammarians working with languages that have them. Such affixes have sometimes been interpreted by others, however, as simple copies of nominal features, mere 'agreement' markers. (For further discussion of this issue see Jelinek 1984, Mithun 1985, Van Valin 1985.) The reasons for this are understandable.

In languages like English, pronouns are normally in complementary distribution with full noun phrases.

- 2)a The boy is sick.
b He is sick.

In languages like Mohawk, Selayarese, and Lakhota, the pronominal affixes appear whether separate noun phrases are present or not. The affixes match the noun phrases in person, and, if distinguished, number, gender, and/or case. Compare the verbs in the Mohawk sentences below.

- 3)a *Rononhwáktani* *ne* *raksà: ?a.*
 3.M.SG.PAT-sick-STATIVE the 3.M.SG-child-DIM
 'The boy is sick.'
- b *Rononhwáktani.*
 'He is sick.'
- 4)a *Iakononhwáktani* *ne* *ieksà: ?a.*
 3.F.SG.PAT-sick-STATIVE the 3.F.SG-child.DIM
 'The girl is sick.'
- b *Iakononhwáktani.*
 'She is sick.'
- 5)a *Rotinonhwáktani* *ne* *ratiksa ?okòn: ?a.*
 3.M.PL.PAT-sick-STATIVE the 3.M.PL-child-DISTR
 'The boys are sick.'
- b *Rotinonhwáktani.*
 'They are sick.'

Separate pronouns do not appear unless they signal a contrast.

Like pronominal affixes, inflectional endings on verbs in many languages reflect the person, number, and/or gender of an argument. They are thus termed 'agreement markers'. Like pronominal affixes, they appear with all verbs, even when separate noun phrases are present in the clause. In the Macedonian sentence below, for example, the third person verb matches its third person subject.

- 6)a *Bratučedka* *mi* *dojde.*
 cousin-FEM my came-3.SG
 'My cousin came.'

In some languages with agreement markers, verbs can also constitute grammatical utterances in themselves. In such languages, sometimes called 'pro drop' languages, pronouns are not normally used unless special emphasis is intended. Thus the Macedonian sentences in b) and c) would usually occur without the pronouns in parentheses.

- b (*Taa*) *dojde.* c (*Jas*) *dojdov.*
 (she) came-3.SG (I) came-1.SG
 '(S)he came.' 'I came.'

The agreement markers do facilitate the interpretation of clauses without nominals: they reflect the person and number of the subject. They differ from pronominal affixes in several ways, however.

Agreement markers, in the narrow sense, normally match just one argument, usually the subject. Pronominal affixes, by contrast, refer to all core arguments of the verb. (Agent and patient pronouns are fused in the Mohawk and Lakhota examples.)

7)	Mohawk	Selayarese	Lakhota
' <u>I</u> know <u>you</u> '	<i>ko<u>n</u>ienté:ri</i>	<i>ku<u>is</u>sekko</i>	<i>slo<u>l</u>č^hfyé</i>
' <u>you</u> know <u>me</u> '	<i>sk<u>i</u>enté:ri</i>	<i>mu<u>is</u>seʔa</i>	<i>slo<u>l</u>máyaye</i>

Although free pronouns are pragmatically marked in both 'pro drop' languages and those with pronominal affixes, they do not appear under identical conditions in the two. The frequency of free pronouns in 'pro drop' languages varies slightly from language to language, and from speaker to speaker, but they function in such languages in ways not normally found in languages with bound pronouns. As one Spanish speaker related experiences from her childhood, for example, she explained that her grandfather had worked in a penny arcade. When he came home, he would count the coins he earned in tips and put them in little piles on the table. The speaker continued:

8) <i>entonces</i>	<i>cuando</i>	<i>yo</i>	<i>al</i>	<i>ver</i>	<i>todo</i>	<i>ese</i>	<i>dinero,</i>
then	when	<u>I</u>	to-the	see-to	all-M	this-M	money
'so upon seeing all this money,							

me *ponia* ... *mala,*
 myself put-P-IMP bad-F
 I would get into a terrible state,

The free pronoun *yo* 'I' appears with the shift in topic from the grandfather to the narrator.

The tale then continued uninterrupted by free first person pronouns, as the speaker described how she stole three coins, ran out and bought chocolate, then ate it, but neglected to wash her face. Her grandmother noticed the theft, and went to tell her mother about it, stopping to inform neighbors all along the way. The mother replied that she knew nothing about the coins, but that since the little girl was such a glutton, she might have taken them. The speaker continued:

<i>inmediatamente</i>	<i>aparec-i</i>	<i>yo,</i>
immediately	appear-P.PF.1S	<u>I</u>
'Just then I showed up,		

y *me* *vieron*
 and *me* *see-P.PF.3PL*
 and they saw me

con *toda* *la* *boca* *manchada* *de* *chocolate*
 with all-F the-F mouth stain-P.PRT-F of chocolate
 with my whole mouth covered with chocolate.

...yo dije no no,
 I say.P.PF-1SG no no,
 I said no, no,

The free pronouns function here to foreground their referents. The sentence then continued:

yo no he cojido el dinero
 I not AUX-1S take-P.PRT the money
 I didn't take the money ... '

Here, the overt pronoun signals a contrast.

In the Spanish text, as in connected discourse in general, free pronouns are absent more often than they are present. When the pronouns do appear, they provide special emphasis, much like stressed pronouns in English. In this role, they may indicate shifts in topic, simply highlight or foreground their referents, or indicate the focus of a contrast. In languages with pronominal affixes, free pronouns serve only the contrastive function.

Finally, agreement markers appear to differ from pronominal affixes in a subtle but crucial way: they are not referential. They reflect various features of subjects, but they are not subjects themselves. Speakers are well aware of the difference. Whether they have thought about the structure of their language or not, they spontaneously mention 'leaving out' or 'dropping' the pronoun, even when this is the only idiomatic alternative. Macedonian speakers note that *jas* has been left out of a sentence like 6)c above, just as Spanish speakers volunteer that *jo* has been omitted from Spanish equivalents. Speakers of Mohawk, Selayarese, and Lakhota, by contrast, would never say that a pronoun has been omitted from *wakenonhwáktani*, *bambana*, or *mak^húže* respectively. The presence of agreement markers may obviate explicit reference to the subject with each verb, so long as the reference is inferrable, but this does not mean that the agreement marker itself is necessarily referential. Note that in languages such as Japanese and Chinese, pronouns are typically not present if the identity of arguments is inferrable from the situation (Matsumoto p.c., Li and Thomson 1979, 1981, Chen 1984, Tao 1986). Yet verbs contain no markers at all specifying person, number, or gender.

It might be easier to dismiss speaker intuitions about the presence of pronouns if they did not correlate with major syntactic differences. For many languages, a basic, syntactically defined constituent order can be determined, such as SVO, SOV, VOS, etc. Often this basic order is obvious: the majority of clauses exhibit the same order, and those that do not are clearly pragmatically marked. Languages of this type can differ in how rigid their basic order is, according to such factors as case marking on noun phrases. English, for example, with no nominal case marking, has relatively rigid SVO order, while Slavic languages, with rich nominal case systems and the same basic SVO order, exhibit considerable sensitivity to pragmatic context in the reordering of constituents. Nonetheless, in all of these languages, deviation from the basic

order signals a more or less pragmatically marked situation.

In a number of languages, however, the determination of a basic order seems at first impossible. All possible orders occur. Usual tests for basic order fail to yield conclusive results. When speakers are provided with potentially ambiguous sentences, usually out of context, scrambling the words does not resolve the ambiguity, (unless the speakers feel that the order closest to the contact language, often English or Spanish, will clarify matters for the grammarian). Comparisons of statistical frequency introduce more complications than they resolve. In languages of this type, although speakers consider it perfectly grammatical to include both a separate subject noun phrase and a separate object noun phrase in the same clause, such full sentences occur relatively rarely in spontaneous discourse, often in as few as 1-2₀ of clauses. Speakers normally do not introduce several new entities into discourse simultaneously. Yet even within the small set of full clauses, all orders occur. The marginally most frequent order found in a set of highly unusual sentences seems a questionable foundation for a syntactic description.

If sentences in these languages are examined in context, it becomes clear that although all orders occur, their distribution is not random (Mithun 1984, Payne 1984). It is not the syntactic roles of the nominals that determine constituent order: it is the relative importance of all constituents to the discourse at hand: their newsworthiness within a particular context. A word may be newsworthy for several possible reasons: it may establish a new topic, announce a new action, introduce a new object, specify a significant time or place, or indicate the focus of a contrast.

The relationship between the relative importance of constituents and their order has been noticed before. On the basis of the Indo-European languages they observed, Firbas and his Prague School colleagues postulated that information is ordered within clauses according to an increasing degree of 'communicative dynamism', 'the degree to which a sentence element contributes to the development of the communication, to which, as it were, it pushes the communication forward' (1972:78). Speakers begin from an established point of departure, the theme or topic, and move toward newer, increasingly important information, the rheme or comment.

In languages with purely pragmatic constituent ordering, however, the order is the reverse of that observed by Firbas in Indo-European. The more newsworthy or important a constituent is to the message at hand, the earlier it appears in the clause. Speakers specify the main point of their message, then fill in details, in decreasing order of importance. New topics precede presentative verbs. Significant shifts in time or scene precede subsequent events. (Speakers can assume that if no topic or scene switch is signalled at the outset, these remain the same. There is no need to reset the frame for every clause.) Significant actions are announced before predictable topics, or incidental adverbials.

Because there is no basic, syntactically defined order in these languages, no order is more pragmatically marked than any other. This does not mean that all possible orders necessarily

occur with exactly equal statistical frequency. Languages seldom distribute important information equally among all types of constituents. Objects might precede verbs somewhat more often than they follow them, for example, if free object noun phrases more often introduce new information than the verbs they appear with. This would not mean that OV was a pragmatically neutral order.

Languages with purely pragmatically based constituent order share several characteristics. In such languages, connected discourse exhibits an especially high proportion of verbs. The verbs may be morphologically quite complex. They often incorporate nouns. Particularly interesting is the fact that they also always have bound pronominal affixes or clitics.²

Mohawk is a language of this type. Constituent order is pragmatically determined. Words are arranged in descending order of newsworthiness. Natural discourse consists predominantly of verbs, which are relatively complex morphologically. Noun incorporation is highly productive. As was seen above, pronominal prefixes refer to the agents and/or patients of all verbs.

An example of this can be seen in the passage below from a Mohawk legend. A man has gone out fishing with friends, become separated from them, and wandered around by himself all night. Then the scene shifts to the next morning and the friends.

- 9) *O:nen ki iá:ken' wa'io:rhen'ne'*
 now then QUOT PAST-N-dawn-PUNCTUAL
 'Now then, the next morning,

ronnatén:ro?s teionkwè:take wahiátkétsko?
 M.PL.PAT-friend-DISTR DU-person-number PAST-M.DU.AG-RFL-rise-PUNCT
 his two friends got up.'

The time shift is most important here, so it appears first. Next in importance is the shift in characters, so that the subject precedes the predicate.

The friends notice that the man is missing, and go looking for him. They find him floating in the river. Then we are told:

- 10) *O:nen wahiáttoke? ki ronnateñ:ro ...*
 now PAST-M.DU.AGT-RFL-point-PCT this M.PL-friend
 'His two friends realized ...'

Here the friends represent a continuing topic, so their realizing is more newsworthy than their identification. (The friends were reidentified overtly only because our attention had been momentarily focussed on the floating man.) The predicate thus precedes the subject. Neither of these orders, SV or VS, is pragmatically neutral, in the sense that both are pragmatically meaningful. The same is true of all other possible orders involving other types of constituents.

The pragmatic rather than syntactic ordering of constituents is no accident. Compare again the Mohawk and English versions of 9).

- 9) ...ronnateh:ro's teionkwè:take wahi^hatkétsko'
 ...his friends two people they two got up
 ...'his two friends got up.'

In the English version, the subject of the verb 'get up' is the noun phrase 'his two friends'. In the Mohawk sentence, the nuclear clause relations are established within the verb itself. The agent of the verb *-atketsko-* ('get up') is the pronominal prefix *-hi-*, ('they masculine dual'). The external nominals further identify the agent argument, but they are appositional to the primary argument, the pronominal prefix.

The same is true of sentences with separate noun phrases identifying patients: the nuclear clause relations are specified within the verb.

- 11) Tsiá:ta niionkwè:take ne: ken nithotiioñ:sa
 seven people the this young men

ne kwáh rati'sátste's tentehsheia'teñhawe'.
 the most they are strong you will bring them back.
 'Bring back seven of the strongest young men.'

External nominals may further identify the arguments, but these are appositional adjuncts. It is not surprising that these external adjuncts fail to be governed by nuclear syntactic considerations. They are coreferent with the nuclear arguments, but they are not primary arguments themselves.

Lakhota also has obligatory pronominal prefixes on verbs, referring to agents and/or patients, as was seen above. Yet constituent order in Lakhota is clearly syntactically based: it is basically SOV, with modifications reflecting pragmatically marked circumstances. Lakhota speakers, like others, do not normally introduce both a new subject and a new object into the same intonation unit or clause, but when noun phrases representing several major arguments do appear, the pragmatically neutral order is SOV. In the narrative cited below, the speaker had already mentioned that he and his friends had put on a pow wow.

- 12) Lak^hóta ikčéáč^ha he wač^hípiwā ūkáyapi.
 Indian ordinary that pow wow we made
 'We were just ordinary Indians that put up the pow wow.'

This difference in ordering strategies is not completely unexpected. Recall that only first and second person arguments are represented by overt pronominal prefixes in Lakhota. These are just the arguments that are rarely represented by noun phrases, since first and second person are inherently established by the discourse context. When asked for third person verb forms out of context, speakers prefer to supply a noun phrase establishing the identity of the third person, or at least a demonstrative.

- 13) *špamáyaye* 'you burned me'
 hé špayáye 'you burned him' ('that you.AGT burn')
 hé španíye 'he burned you' ('that you.PAT burn')
hé wíya hé špáye 'he burned her' ('that woman that burn')

Arguments need not be respecified in every clause when their identity is clear from preceding discourse. A zero prefix is not sufficient in itself, however, to establish reference.

The absence of third person pronominal prefixes in Lakhota signals a crucial difference in the organization of relationships between verbs and their arguments. In Lakhota, the third person agents and patients of verbs are not some zero morphemes invisibly prefixed to verbs. They are external noun phrases or demonstratives. It is not surprising that the relative order of major constituents should mirror this syntactic relationship. Pronominal systems like that in Lakhota, whereby only first and second persons are referenced by verbal affixes or clitics, are common among languages throughout the world. Constituent order in such languages is syntactically based, often SOV, mirroring the syntactic relationships binding verbs and their noun phrase arguments.

The existence of non-zero third person pronominal affixes or clitics is not sufficient, however, to guarantee that constituent ordering in a language will be purely pragmatically based. The Mayan languages, for example, exhibit many of the features characteristic of pragmatically ordered languages. Verbs predominate in spontaneous discourse. Du Bois (1985) has shown for example, that in Sacapultec, a Mayan language of Guatemala, nearly 60% of clauses in connected speech contain only one core noun phrase, and nearly 40% contain no core noun phrase at all. Mayan verbs can exhibit a relatively high degree of polysynthesis, and many show a type of compounding akin to incorporation. They contain pronominal affixes referring to all three persons.

Yet word order in these languages is clearly syntactically based. Most Mayan languages show a syntactically defined, pragmatically unmarked VOS order, a few show VSO, and perhaps one has SVO (Kaufman p.c.). The rigidity of the basic order varies somewhat from one language to another. Nominal constituents may be fronted for focus, or to indicate a new or contrastive topic, but the resulting orders are pragmatically marked, as indicated by certain grammatical markers or intonation breaks separating fronted constituents from the remainder of the verb-initial clause.

A careful look at the forms of the third person markers is revealing. Mayan verbs contain pronominal affixes that operate on an ergative basis: subjects of transitive verbs are referred to by ergative affixes, while objects of transitive verbs and subjects of intransitives are referred to by absolutive affixes, at least in the first and second person. Compare the Sacapultec verbs below. The first prefix refers to the absolutive, the second to the ergative.

- 14) *in-warnaq* 'I was asleep'
 at-warnaq 'you were asleep'

<i>in-a:-čixmax</i>	'you have cared for me'
<i>in-ra-čixmax</i>	'he has cared for me'
<i>at-n-čixmax</i>	'I have cared for you'
<i>at-ra-čixmax</i>	'he has cared for you'

When the absolutive is a third person, there is no overt pronominal affix on the verb. This has traditionally been described with a zero third person absolutive marker.

<i>Ø-warnaq</i>	'he was asleep'
<i>Ø-ni-čimax</i>	'I have cared for him'
<i>Ø-a:-čimax</i>	'you have cared for him'

In light of the above discussion, it might be more appropriate to recognize that the pronoun simply is not there.

As Du Bois has pointed out, absolutives constitute more than an accidental, hybrid category. The objects of transitive verbs and the subjects of intransitives are exactly the arguments normally used to introduce new entities into discourse. New entities, especially third persons, are necessarily introduced by full noun phrases, rather than pronouns. Thus full noun phrases are especially likely to appear as absolutive arguments. In most of the Mayan languages, absolutive noun phrases typically follow the verb, both in transitive sentences (VOS) and intransitive sentences (VS). Their postverbal position appears to be syntactically based, mirroring the syntactic relationship between verbs and their nominal arguments. (As in other languages, overt reference to absolutives is not obligatory in every clause, providing the identity of the referent is clear from context.)

Ergative arguments, the subjects of transitives, rarely represent new entities. They are typically continuing topics, so their only mention is pronominal. It is not surprising that established entities should be systematically backgrounded by morphologization. The ergative noun phrases that do appear often reiterate the identity of the subject in potentially confusing situations, as in 'afterthought' constructions, or they provide the focus of a contrast. Such noun phrases are often separated intonationally from the rest of the clause, the verb-object nucleus. This suggests that in at least some Mayan languages with basic VO order, the primary ergative arguments of verbs may be pronominal affixes, while the absolutive arguments are separate noun phrases (or understood from context).

Although this situation is relatively common among ergative languages, it is not restricted to them. In Daga, a Papuan language of New Guinea, bound pronominal suffixes refer to subjects and objects (Murane 1974). Subject pronouns are fused with tense markers, and show no zero morphemes. The set of direct object suffixes, however, contains one zero member: third person. The language has syntactically determined SOV order.

Yet even a full set of non-zero third person pronominal affixes is not sufficient to guarantee pragmatically based ordering in a

language. In Selayarese, the Austronesian language cited earlier, ergative arguments are referred to by pronominal prefixes on verbs, and absolutes by clitics suffixed to the first word of the verbal phrase. There are no 'zero' morphemes, although third persons are not distinguished for number or gender.

- 15) *aʔgoloa* 'I play soccer'
aʔgoloko 'you (familiar) play soccer'
aʔgoloi 'he/she/they play soccer'
- kuʔuraŋko* 'I accompany you'
kuʔuraŋi 'I accompany him/her/it/them'
- muʔuraŋa* 'you accompany me'
muʔuraŋi 'you accompany him/her/it/them'
- lauraŋa* 'he/she/it/they accompany me'
lauraŋko 'he/she/it/they accompany you'
lauraŋi 'he/she/it/they accompany him/her/it/them'

Such verbs may constitute complete predications in themselves, or they may appear in longer sentences, with noun phrases coreferent to one or both of their arguments. It is clear that the affixes constitute referential pronouns, rather than simple 'agreement' markers. When a speaker was asked whether the *-i* of the verb *allariʔi* 'he runs', corresponds to the 'he' of the English translation or the '-s', there was not doubt in his mind, even though he had never written nor read his language before. (He is a good English speaker.) He felt strongly that the suffix corresponds to the 'he'.

Although the verbal conjugation contains non-zero pronominal affixes for all persons, Selayarese has a syntactically defined verb-initial constituent order. Any constituent may be fronted for special focus, but this results in a pragmatically marked order.

- 16)a *Lahalliʔi* *sapoŋjo* *iBasoʔ*.
 3.ERG-buy-3.ABS house-the Basoʔ
 'Mr. Basoʔ bought the house.'

- b *iBasoʔ* *lahalliʔi* *sapoŋjo*.
 Basoʔ 3.ERG-buy-3ABS house-the
 'Mr. Basoʔ bought the house.'

Now compare the two sentences below.

- 17)a *Ku-halliʔ-i* *sapo-ŋjo*. b *Mmalliʔ-a* *sapo*.
 1.ERG-buy-3.ABS house-the INTR-buy-1.ABS house
 'I bought the house.' 'I bought a house.'

When the direct object is indefinite, like *sapo* 'house' in 18)b, it is not registered on the verb. The absolutive suffix *-a* refers to the subject 'I', and there is no ergative prefix.

This is a general pattern in Selayarese. Subjects are always marked on the verb by means of referential affixes, but direct objects are referred to in the verb only if definite or in special focus. In light of the foregoing discussion, this is not altogether unexpected. New entities are most likely to be introduced into discourse as indefinite direct object noun phrases. Since they are new, pronominal reference would not be sufficient. Thus the relation between a verb and a first or second person core argument is specified within the verb. The relation between a verb and its third person subject or definite object is also specified within the verb, although an appositional noun phrase may follow to clarify its identity. The relationship between the verb and an indefinite direct object, however, is specified by the order between the verb and the following noun phrase.

This situation, like the others discussed above, is also not restricted to ergative languages. In Amharic, for example, definite specific direct objects are registered on the verb by means of pronominal suffixes, but indefinite ones are not. The language has a relatively rigid verb-final order. (Hetzron 1975, Hudson 1972)

Finally, languages may exhibit still another relationship between pronominal affixes and separate noun phrases. Consider the situation in Abkhaz, a North West Caucasian language described by Hewitt (1979). Bound pronominal affixes refer to three nuclear arguments: absolutes, indirect objects, and ergatives. There are no 'zero' forms in the paradigm: each affix consists of a single consonant followed by an optional vowel. Yet constituent order is syntactically determined: the basic, unmarked order is S-IO-DO-V (Hewitt 1979:103, Hewitt and Khiba 1985).

An examination of the operation of the pronominal system is revealing. Abkhaz has two third person absolute prefixes, *d(ə)-* for singular humans, *y(ə)-* for all other third persons.

18) *də-cə-yt'*

3.HU.SG-go-FINITE
'(S)he went.'

yə-cə-yt'

3.N.HU-go-FINITE
'It/they went.'

yə-sə-t

3.N.HU-to.me-give
'Give it to me.'

Hewitt notes that 'if the referent of the affix *y(ə)-* immediately precedes the verb, then this affix *y(ə)-* disappears' (1979:101). Compare the sentences below.

19)a *a-xác'a də-cə-yt'*

the-man 3.HU.SG-go-FINITE
'The man went.'

a-la (θ)ce-yt'

the-dog (it-)go-FINITE
'The dog went.'

a-xác^oa (θ)ce-yt'

the-men (they-)go-FINITE
'The men went.'

a-š^oq^o'ə (θ-)sə-t
 the-book (it-)to.me-give
 'Give me the book.'

As noted earlier, third person absolutes are the arguments most often used to introduce new entities into discourse, so they are the arguments most likely to be represented by full noun phrases. As would now be expected, the relative order of verbs and immediately preceding absolute noun phrases is syntactically based. The syntactically based order of the constituents reflects their syntactic relationship. What is interesting here is that the language permits primary arguments to be either free noun phrases or bound pronouns. This situation is also not unusual, nor is it restricted to ergative languages. Comrie (1982) describes a similar situation in the Central dialect of Huichol, for example, whereby third person singular objects are not referred to pronominally in the verb when the object noun phrase immediately precedes.

The occurrence of 'zeroes' in pronominal affix systems is not random. If one person marker is 'zero', this is most likely to be the third person. If only some third person pronominal affixes are 'zero', these are most likely to be absolutes, or, somewhat less often, direct objects. If only some third person absolute or direct object pronominal affixes are 'zero', these are most likely to be indefinite. These are exactly the points at which new entities tend to be introduced into discourse, by full noun phrases. In fact, in some languages, pronominal affixes are 'zero' specifically in the presence of adjacent noun phrases. In all such cases, the presence of these 'zeroes' cooccurs with the existence of a syntactically based order between verbs and their core arguments.

It is not difficult to imagine how a language might move from a syntactically based, theme-rheme order, to a purely pragmatically based, rheme-theme order. (For ideas along these lines, see, among others, Payne 1984 and Givon 1976.) Probably the majority of languages with syntactically based word order allow 'left dislocation', whereby constituents in special focus appear at the beginning of the clause. The markedness of the resulting structure varies, of course, from one language to another. A pronoun may appear in the basic position of the focussed constituent or not, depending on the language. Left dislocation is often accompanied by cues that the focussed item is not an integral part of the nuclear clause, such as an intonation break separating it from the rest of the sentence.

A second familiar construction is right dislocation, resulting in what is often termed an 'afterthought'. In such situations, a speaker utters a complete clause, then fills in details or repairs unclear reference by means of a following explanation. Such additions can be true afterthoughts, as a speaker realizes that the identity of key participants might not be clear due to intervening material. Speakers also often produce such constructions quite intentionally, when inclusion of the information within the nuclear

clause would interfere with the impact of the message. Right dislocated elements, like left dislocated ones, are not part of the nuclear clause, as shown by the intonation breaks often separating them, and a lowering of tone and sometimes volume.

In a language with obligatory overt third person pronouns, fronting significant new information for focus, or reiterating or refining given information following the clause, would not interfere with an understanding of the relations among sentence constituents. The constituents themselves remain firmly in place within the nuclear clause. As speakers use left and right dislocation more often, the pragmatic and intonational markedness of such structures can fade, providing increasingly natural strategies for presenting information according to its pragmatic importance. This does not mean that the dislocated items necessarily assume the role of primary reference. As long as the clause-internal pronouns remain overt, there is no reason for them to abdicate their function. Since the dislocated constituents retain their role as appositives, there is also no reason for syntactically based ordering patterns to extend to them.

A number of language families provide evidence of such development. A particularly rich example is provided by Australian aboriginal languages. As Dixon (1980) has shown, all of these languages are ultimately genetically related. They fall into two major typologically distinct groups, termed Pama-Nyungan and non-Pama-Nyungan.

The Pama-Nyungan languages show interesting variation in their pronominal systems. Some have no bound pronominal forms at all, such as Uradhi (Crowley 1983), Diyari (Dixon 1980), Thargari (Klokeid 1969), Guugu Yimidhirr (Haviland 1979), Dhalandji (Dixon 1980), Pitta Pitta (Blake 1979), Gumbaynggir, (Bades 1979), Yaygir (Crowley 1979), Warrgamay, (Dixon 1981), Mpakwithi (Anguthimri) (Crowley 1981), Nyawaygi, (Dixon 1983), and others. Some of these, such as Uradhi and Diyari, have relatively rigid SOV orders, while others, such as Thargari, and Guugu Yimidhirr, show considerable flexibility, facilitated by the presence of case suffixes on nouns. In all of these languages, however, a basic, syntactically defined constituent order can be perceived.

A number of Pama-Nyungan languages are in the process of developing bound pronominal affix or clitic systems. In accordance with the suffixing character of the Pama-Nyungan languages, the pronominal forms are usually enclitics, suffixed loosely to the verb or to the first word of the sentence. Their recent development is clear. Their resemblance to free pronouns is still transparent, and they show little phonological fusion with the words to which they attach. There are even two dialects of a single language, Margany and Gunya, that differ principally in this feature: 'Gunya has a transparent and obviously recent system of pronominal suffixes to the verb, which Margany lacks' (Breen 1981:275).

Some languages in the early stages of this development, such as Watjarri (Douglas 1981), Yankunytjatjara (a dialect of Western Desert) (Goddard 1983), and Kalkatungu (Blake 1979), still permit a choice between enclitic and free pronouns. Douglas detected a

preference for free pronouns in Watjarri conversation. Goddard remarked that in Yankunytjatjara, 'sentences using free pronouns only are not uncommon in normal speech' (1983:120). Blake reports that Kalkatungu speakers may use bound pronouns alone, free pronouns alone, or combine them, although only one bound form may occur in an independent clause. He found more free forms in elicited sentences, but otherwise noted that the free forms tend to indicate a focus or contrast. In all of these languages, there are no overt third person clitics. Although word order can be relatively flexible, due to case suffixes on nouns, it is basically SOV.

Some languages that offer the alternative between free and bound pronouns have overt non-zero third person enclitics. Dixon notes that in Ngiyambaa, 'a pronoun will normally be reduced to a clitic unless it requires special emphasis' (1980:365). There are no free third person pronouns, but the enclitic third person ergative is *-lu/-yu* and the absolutive-dative is *-na*. Third person absolutives can be represented by a noun phrase alone, an enclitic alone, or a combination of both, if the referent is definite. The presence of these non-zero affixes is significant. Dixon states: 'Languages for which it has not proved possible to make statements concerning underlying order include Ngiyambaa' (1980:497).

In still other Pama-Nyungan languages, obligatory bound pronominal systems have developed with no zero members. In Yukulta, for example, 'a clitic complex or cluster of clitics expressing S, A, O, and O² (direct or indirect object), transitivity and tense-aspect is suffixed to the first constituent of a clause.' (Keen 1983:191). The clitics are obligatory. An examination of texts in Yukulta reveals that the order of constituents is pragmatically based, with words arranged in decreasing order of newsworthiness, or rheme-theme.

Ritharngu, a Pama-Nyungan language, is closely related to central Australian languages (Heath 1980a), but is spoken in Northeast Arnhem Land, among the typologically distinct non-Pama-Nyungan languages. It is suffixing, like its closer relatives, but has developed an elaborate system of bound pronominal enclitics, none of which is zero. Constituents are ordered according to their pragmatic function, in decreasing order of newsworthiness, or from rheme to theme (1980:102-3).

The non-Pama-Nyungan languages, spoken mainly in Northern Australia, are typologically quite different from their Pama-Nyungan relatives. Languages of this type include, among others, Tiwi (Osborne 1974), Ngandi (Heath 1978), Dhuwal (Heath 1980b), Gunwinggu (Oates 1964), Warndarang (Heath 1980c), Nunggubuyu (Heath 1984), and Rembarrnga (McKay 1975). Connected discourse in these languages typically contains a high proportion of verbs, which can be morphologically quite complex, and can function as complete clauses in themselves. Most of the languages lack case marking on nouns, and where it exists, it is thought to be the result of influences from nearby Pama-Nyungan languages (Heath 1978). Significantly, verbs contain tightly fused pronominal prefixes, all of which have overt, non-zero forms. It is clear from the excellent textual materials in these languages that constituent ordering is

pragmatically determined, with constituents arranged in decreasing order of importance. (Mithun 1984).

The recurring coincidence of full sets of overt pronominal affixes and pragmatically based constituent ordering is too pervasive to be an accident. It may be the case that languages with full sets of bound pronouns can retain a basic, neutral syntactically based constituent order indefinitely, either because speakers never overindulged in dislocation, or because of the well known sensitivity of word order to external influences. The opposite does not seem to occur, however. It appears that all languages with purely pragmatically determined, rheme-theme order, establish core grammatical relations within their verbs, between verb stems and overt bound pronouns. In light of the non-random distribution of 'zero' pronouns within paradigms, and of their ramifications within the syntax, it seems impossible to maintain that such 'zero' morphemes are mere arbitrary paradigmatic alternatives. Sometimes they are not even there.

NOTES

¹I am grateful to the following speakers who generously provided the examples from their languages cited here, as well as discussion: Mrs. Margaret Edwards, of Ahkwehsane, Quebec for Mohawk, Mr. Hasan Basri, of South Sulawesi, Indonesia, for Selayarese, Mr. Stanley Redbird of Rosebud, South Dakota, for Lakhota, Dr. Olga Tomić, of Skopje, for Macedonian, Miss Alicia Re, of Madrid, and Mrs. Nelly Tellez Schragger, of Barranquilla, departamento del Atlantico, Columbia, for Spanish. Rene Shekerjian transcribed the Spanish texts and conversations discussed here from tape recordings. I also much appreciate helpful discussions with Jack Du Bois, Orin Gensler, Bill Jacobsen, Terry Kaufmann, and Johanna Nichols.

² Clustering of many of the properties discussed here, in particular, a preponderance of verbs in connected discourse, complex verbal morphology, variable constituent order, and the presence of pronominal affixes, have also been noticed by Desmond Derbyshire among Amazonian languages (1985), and more generally by Johanna Nichols (to appear) among 'head-marking' languages.

REFERENCES

- Benveniste, Emile. 1966. 'Structure des relations de personne dans le verbe'. *Problèmes de linguistique générale I*. 225-36.
- Blake, Barry. 1979. 'Pitta Pitta' In Dixon and Blake, 181-242.
- 1979. *A Kalkatungu Grammar*. Pacific Linguistics Series B, No. 57. Canberra: Australian National University.
- Breen, J. G. 1981. 'Margany and Gunya'. In Dixon and Blake, 272-292.
- Chen, Ping. 1984. 'A discourse analysis of third person zero anaphora in Chinese'. Bloomington: Indiana University Linguistics Club.

- Comrie, Bernard. 1982. 'Grammatical relations in Huichol.' In Paul Hopper and Sandra A. Thompson, eds. *Studies in transitivity*. Syntax and Semantics 15. New York: Academic Press. 95-115.
- Crowley, Terry. 1979. 'Yaygir'. In Dixon and Blake, 363-84.
- 1981. 'Mpakwithi'. In Dixon and Blake, 144-195.
- 1983. 'Uradhi'. In Dixon and Blake, 306-428.
- Derbyshire, Desmond C. 1985. 'Morphosyntactic areal characteristics of Amazonian languages.' Paper read at the American Anthropological Association, Washington, D.C.
- Dixon, R.M.W., 1980. *The languages of Australia*. Cambridge: Cambridge University Press.
- 1981. 'Wargamay.' In Dixon and Blake, 1-143.
- 1983. 'Nyawaygi.' In Dixon and Blake, 430-525.
- and Barry J. Blake. 1979. *Handbook of Australian languages*. Canberra: The Australian National University Press. Volume 1.
- 1981. *Handbook of Australian languages*. Canberra: The Australian National University Press. Volume 2.
- 1983. *Handbook of Australian languages*. Canberra: The Australian National University Press. Volume 3.
- Douglas, Wilfrid H. 1981. 'Watjarri'. In Dixon and Blake, 196-271.
- Du Bois, John W. 1985. 'Competing Motivations.' in John Haiman, ed. *Iconicity in Syntax*. Amsterdam: John Benjamins. 343-365.
- Eades, Diana Kelloway. 1976. *The Dharawal and Dhurga languages of the New South Wales south coast*. Canberra: Australian Institute of Aboriginal Studies.
- 1979. 'Gumbaynggir' In Dixon and Blake, 243-361.
- Firbas, Jan. 1972. 'On the interplay of prosodic and non-prosodic means of functional sentence perspective. V. Fried, ed. *The Prague School of linguistics and language teaching*. London: Oxford University Press. 77-94.
- Givón, Talmy. 1976. 'Topic, pronoun, and grammatical agreement'. In Charles N. Li, ed. *Subject and topic*. New York: Academic Press. 149-188.
- Goddard, Cliff. 1983. *A grammar of Yankunytjatjara*. Alice Springs: Institute for Aboriginal Development.
- Haviland, John. 1979. 'Guugu Yimidhirr'. In Dixon and Blake, 26-180.
- Heath, Jeffrey. 1978. *Ngandi grammar, texts, and dictionary*. Canberra: Australian Institute of Aboriginal Studies.
- 1980a. *Basic materials in Ritharngu: grammar, texts, and dictionary*. Pacific Linguistics Series B No. 62. Canberra: Australian Institute of Aboriginal Studies.
- 1980b. *Dhuwal (Arnhem Land) texts on kinship and other subjects with grammatical sketch and dictionary*. Oceania Linguistic Monographs No. 23.
- 1980c. *Basic materials in Warndarang: grammar, texts, and dictionary*. Pacific Linguistics Series B No. 72. Canberra: Australian Institute of Aboriginal Studies.

- 1984. *Functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.
- Hetzron, Robert 1975. 'The presentative movement or why the ideal word order is V.S.O.P.' In Charles N. Li, ed. *Word order and word order change*. Austin: Texas University Press.
- Hewitt, B. G. 1979. *Abkhaz*. Lingua Descriptive Series Vol 2. Amsterdam: North Holland.
- and Z. K. Khiba. 1985. 'Contributions to North West Caucasian Comparative Studies I'. *Studia Caucasica* 6. 26-30.
- Hudson, Grover. 1972. 'Why Amharic is not a VSO language'. *Studies in African linguistics*. 3.1.127-165.
- Hudson, Joyce. 1978. *The core of Walmatjari grammar*. Canberra: Australian Institute of Aboriginal Studies.
- Jelinek, Eloise. 1984. 'Empty categories, case, and configuration-ality'. *Natural Language and Linguistic Theory*. 2.1. 39-76.
- Keen, Sandra. 1983. 'Yukulta.' In Dixon and Blake, 190-304.
- Klokeid, Terry J. 1969. *Thargari phonology and morphology*. Pacific Linguistics Series B No 12. Canberra: The Australian National University.
- Li, Charles N. and Sandra A. Thompson. 1979. 'Third person pronouns in zero-anaphora in Chinese discourse'. In Talmy Givón, ed. *Discourse and syntax*. Syntax and Semantics 12. New York: Academic Press. 311-35.
- 1981. *Mandarin Chinese: a functional reference grammar*. Berkeley: University of California Press.
- McKay, Graham Richard. 1975. *Rembarnga: a language of central Arnhem Land*. Ph.D. dissertation, The Australian National University.
- Mithun, Marianne. 1984. 'Is basic word order universal?' presented at Eugene conference on discourse. To appear in Russell Tomlin, ed. *Coherence and Grounding in Discourse*. Amsterdam: John Benjamins.
- 1985. 'Disagreement: the case of pronominal affixes and nouns'. Georgetown University Round Table Conference on Languages and Linguistics, Washington D. C. To appear in the Proceedings.
- Murane, Elizabeth. 1974. *Daga grammar*. Norman, Oklahoma: Summer Institute of Linguistics.
- Nichols, Johanna. to appear. 'Head-marking and dependent-marking grammar.' *Language*.
- Oates, Lynette Frances. 1964. *A tentative description of the Gunwinggu language*. Oceanic Linguistic Monographs No 10.
- Osborne, C. R. 1974. *The Tiwi language*. Canberra: Australian Institute of Aboriginal Studies.
- Payne, Doris. 1984 ms. 'Information structuring in Papago'.
- Tao, Lian. 1986. 'Clause Linkage and zero anaphora in Mandarin Chinese'. *Davis Working Papers in Linguistics No. 1*. 36-102.
- Van Valin, Robert D. Jr., 1985. 'Case marking and the structure of the Lakhota clause'. In Johanna Nichols and Anthony C. Woodbury, eds. *Grammar inside and outside the clause*. Cambridge: Cambridge University Press. 363-413.