

# Some Lexicalist Remarks on Incorporation Phenomena

Stephen R. Anderson\*  
Yale University

16 June, 2000

## 1 Introduction: Two Approaches to Noun Incorporation

An important theme in the study of language during the twentieth century has been the balance between the grammar of words (morphology) and that of phrases and sentences (Syntax). Some writers, such as Morris Swadesh (1939), have seemed to treat this as a merely terminological issue, referring to (much of) morphology simply as “internal syntax.” More recently, some (e.g. Selkirk 1982, Williams 1989) have argued explicitly that the mechanisms necessary to describe the internal organization of words are a direct extension of syntactic principles. Others, in contrast (cf. among others Anderson 1992, chap. 2; Zwicky 1992) have argued that there are substantive differences between morphological and syntactic principles which have important implications for the organization of linguistic knowledge.

This discussion has largely centered on the analysis of a small number of constructions for which both syntactic and (more purely) morphological accounts seem to be available. If the syntactic account can be substantiated in these cases, the more general conclusion seems warranted that syntactic mechanisms are (or at least can be) responsible for combining pieces into complex words. If, on the other hand, there are reasons to prefer the autonomously morphological analysis in such circumstances (which are, after all, the best candidates for syntactic word formation *prima facie*), that conclusion casts doubt on the more general perspective from which structurally autonomous elements are manipulated syntactically to form unitary words, a notion that pervades much recent work.

---

\*The work represented here was supported in part by grants number SBR-9514682 and SBR-9876456 from the US National Science Foundation to Yale University; an earlier version appeared as Anderson 1997. This paper has benefited from discussion in seminars at Yale during the past several years, and at the Lexicon in Focus conference in Wuppertal on 19 August, 1998. I am especially grateful to the members of my seminar on Clitics at Yale during 1996-97 (David Harrison, Lizanne Kaiser, Matt Richardson and Jennifer Vanloon), some of whom also participated in a seminar devoted to a prepublication version of Baker 1995 that was kindly provided to us by Mark Baker. The comments and criticisms received on these occasions have been very helpful. I have also benefited from the insightful comments of two referees for the current volume.

The topic of this paper is one widely discussed instance of such a morphological structure that looks as if it might be formed within the syntax: Noun Incorporation. This is a feature of the grammar of a wide range of languages, including (as a semi-random, quite small sample) Mohawk, Chuckchee, Southern Tiwa, Classical Nahuatl, and many others. Noun Incorporation has attracted attention for quite some time: grossly, it is a construction in which, on the surface, a verb stem and a Noun stem constitute a single word, with the Noun stem representing an argument of the Verb. The most typical incorporated Nouns correspond to the direct object of a transitive Verb, as in the examples in (1) below.

(1) Southern Paiute (Sapir 1911, p. 263)

- a. qām'ú- yaai- num- puḡa'  
 jackrabbit- hunt- usitative- remote past  
 He used to hunt jackrabbits
- b. cū'q'uc<sup>u</sup> qām'ú- v<sup>ax</sup>qa- q'a'  
 one jackrabbit- kill- COMP (SS)  
 Having killed one jackrabbit,...

In some languages, the underlying object/derived subject of an unaccusative Verb can incorporate; in other languages, unergative subjects can incorporate too.<sup>1</sup> These possibilities are illustrated in (2) below

(2) Chukchee (Polinsky 1990)

- a. ŋeyk-ək ʔəl-ə-lg-ə-gʔi  
 hill-loc snow-evid-melt-evid-aor3sg  
 On the hill, the snow melted
- b. ətlon ŋinqe-et-ə-lʔet-gʔe  
 he child-intens-evid-come-aor3sg  
 He got many children

In a few cases, the incorporated Noun appears to be some sort of adjunct, like an instrumental or a locative.

- (3) a. (Huahtla) Nahuatl (Merlan 1976):  
 yaʔ ki-kočillo-teteʔki panci  
 3sg 3sg:it-knife-cut bread  
 He cut the bread with a knife
- b. Chukchee (Skorik 1948, *apud* Spencer 1995):  
 gətg-əlqət-gʔe walwəŋən  
 lake-go-3sgS raven:abs  
 Raven went to the lake

---

<sup>1</sup>Polinsky 1990 provides arguments for the 'unaccusative' vs. 'unergative' contrast in the Verbs of these examples.

Noun Incorporation has been an important topic for a long time. In the early years of this century, it was argued by some to constitute a particularly characteristic feature of North American languages. Kroeber (1909), in an attempt to rebut that notion, argued that in fact there was no such thing as Noun Incorporation, but in a classic paper Sapir (1911) provided the first reasoned analysis of Noun Incorporation in a variety of languages.

There are two basic views of the nature of Noun Incorporation, and both have been around for a very long time. One of them, which we can refer to as the “syntactic” story, treats Noun Incorporation as a syntactic process, by which an argument of the Verb (or at least part of that argument) is actually moved from its syntactic A-position to adjoin to the Verb (see e.g. Mardirussian 1975). In favor of this analysis *a priori* is the natural account it offers of how the incorporated stem comes to fill the semantic ( $\theta$ ) role of a corresponding unincorporated Noun. Proponents of this view in the recent literature include Jerry Sadock (1980, 1986), who argued, from the apparent need to form words in the syntax by moving a Noun into the same word as a Verb, that the Lexicalist prohibition against syntactic manipulation of the internal form of words must be wrong. Actually, the construction discussed by Sadock in Eskimo is one (the formation of ‘de-nominal Verbs’) that several other authors (Sapir 1911; Mithun 1984, 1986; Gerds 1997) have wanted to distinguish from Noun Incorporation, but as he puts the issues they are quite similar, regardless of terminology.<sup>2</sup>

A conceptually similar (but technically very different) position is developed by Mark Baker, who has explored it at some length in one book (Baker 1987) and made considerable use of it in another (Baker 1995). Baker also maintains that Nouns move in the syntax to take up their incorporated positions inside of Verbs. But Baker’s theory of the kind of movement involved goes much farther, and in fact is the basis of the rather more general notion of ‘head movement’ in syntax. This, in turn, is at the technical heart of much contemporary theorizing about clause structure, because it is head movement that is central to the notion that the inflectional content of a clause is composed of a large number of functional heads that move around syntactically. If head movement were not an established notion, much of the ‘split-INFL’ account of clause structure would have rather less plausibility; and if Noun Incorporation were to turn out not to involve syntactic movement, the original (and empirically richest) support for head movement would disappear.

The alternative to this syntactic movement analysis is the “lexical” account of Sapir 1911, Mithun 1984, Anderson 1992 and others, which says that the Noun plus Verb combinations we call ‘Noun Incorporation’ are built in the lexicon, not in the syntax. As with the syntactic analysis, the lexical one has some apparent *prima facie* advantages.

One sort of fact that seems to point to a lexical analysis is that the shape of incorporated nominal elements often differs phonologically from a corresponding free form in unpredictable ways. In some instances, such variation is limited to a small number of suppletive forms: e.g., Mohawk “-*nahskw*- ‘domestic animal’ appears only incorporated, while a semantically equivalent stem *-tshenv* appears only as an independent N[oun]” (Mithun 1984, p. 876). In other languages, however, this phenomenon may be quite pervasive. Thus, in the Australian language Tiwi, “[i]ncorporated forms and free forms are generally not cognate, [. . .] as the free forms which were cognate with the exist-

---

<sup>2</sup>For some discussion of Sadock’s arguments in a context quite similar to that of the present paper, see Anderson 2000.

ing incorporated forms have long since disappeared, owing to the rapid rate of lexical replacement in Australian languages” (Osborne 1974, p. 48). Some representative pairs of incorporated and free forms are given in (4). Note that in some cases, some degree of phonological resemblance can be seen (especially in some body part names), but in others there is no apparent similarity. Also, in some cases a single free form corresponds to more than one incorporated form, while in others the reverse is true.

(4) Tiwi (Osborne 1974):	incorporated	free	gloss
	wuɾati-	-wuɾatiŋa	forehead
	uŋintami-	-jərəŋintamuɾa	nose
	awəri-	-wuɾara	belly
	awəri-	təmuriɸa	navel
	maratiŋa	t̪əraka	(live) wallaby
	ŋiliwant̪ə(ŋə)	t̪əraka	(dead) wallaby
	paŋəlaɾinti-	-taraŋa	buffalo

A language may show semi-systematic relations between free and incorporated forms, where these do not in fact reflect phonological regularities found elsewhere. A particularly interesting instance of this is found in the Munda language Sora and some of its relatives, as documented and analyzed by Zide (1976). In this language, (the free forms of) a number of Nouns can be regarded as derived from verbal or other roots by the addition of prefixes, suffixes or infixes: thus, *əb-ga* ‘feed’ is related to *ər-əb-ga* ‘food (being fed to an infant)’; *koŋ* ‘shave’ is related to *k-ən-oŋ* ‘razor’, etc. In some cases, the combining form of such a Noun can be made by simply removing such added morphological material; thus, the combining form of *k-ən-oŋ* ‘razor’ is *-koŋ*, and that of *gətasi* ‘play’ is *-gəsi* (cf. the verb *gəsi-* ‘to play’ from which the Noun is derived by infixation). This relation has been massively over-generalized and inverted, however; in that a huge number of polysyllabic free form Nouns have incorporated forms produced by removing non-existent prefixes, suffixes and infixed -VC- sequences so as to reduce the form to a canonical -CVC shape. For details, see Zide 1976; our point is simply that the elements appearing in incorporation structures are quite distinct from semantically similar free forms in ways that do not follow from the independent phonology of the language, but rather point to lexical processes.

Just as the phonological form of incorporated and non-incorporated elements may require specifically lexical description, the semantics of an incorporated nominal construction may also differ, again unpredictably, from the interpretation of a corresponding construction in which the nominal is represented by a free NP.

- (5) a. Mohawk (Baker 1995):
- i. tu-s-a-yu-[a]t-háh-a-hkw-e’  
 dup-iter-fact-FsS-srfl-road-pick.up-punc  
 She started her journey (lit.: picked up the road)
  - ii. #tu-s-a-yú-([a]te)-hkw-e’                      ne oháha  
 dup-iter-fact-FsS-srfl-pick.up-punc art road  
 She picked up the road (literal reading only)

b. (Huahtla) Nahuatl (Merlan 1976)

- i. i.  $\emptyset$ -neč-maka- $\emptyset$ -k      paʔtli  
           3sg-1sg-give-pst-sg    medicine  
           He gave me medicine
- ii.  $\emptyset$ -neč-paʔ-maka- $\emptyset$ -k  
           3sg-1sg-medicine-give-pst-sg  
           He doctored me
- ii. i. tesiwitl    weci- $\emptyset$ - $\emptyset$   
           hail      fall-pres-sg  
           (What's falling? — answer:) Hail is falling
- ii. tesiwí-weci- $\emptyset$ - $\emptyset$   
           hail-fall-pres-sg  
           (What's the weather like? — answer:) It's hailing

Neither of these types of idiosyncrasy is to be expected if Noun Incorporation is simply a syntactic process that may (or may not) apply to a fully articulated underlying syntactic structure to move some of its lexical material from one position in the tree to another. Such arguments have not received much attention in the syntactic literature, but this does not mean they are negligible: syntacticians tend to assume that both the phonology and the lexical semantics will take care of themselves, and the possibility that this might not always be the case often carries little weight.

These sorts of idiosyncrasy point away from the syntax in the analysis of incorporation structures, and toward the locus of item-specific information, the lexicon. We must ask, of course, whether such [<sub>V</sub> [<sub>N</sub> stem] stem] combinations are in fact within the scope of the type of regularity found in the lexicon, but the answer is surely in the affirmative. We build such combinations in the lexicon anyway in the case of “synthetic” compounds: *duck hunting* ([<sub>V</sub> [<sub>N</sub> duck] [<sub>V</sub> hunting]) is built from [<sub>N</sub> duck] and [<sub>V</sub> hunt] and has an interpretation in which the Noun is taken as specifying one of the arguments of the associated Verb. While some have argued for a syntactic incorporation analysis of such compounds, the appeal of such a view is limited, since compounds are evidently structurally parallel to non-compound members of basic lexical categories. The possibility of building all lexical compounds in the syntax was tried out in the early days of generative grammar (Lees 1960), but that was really before there was any theory of the lexicon. Nowadays most authors agree that the formation of compounds takes place in the lexicon, not the syntax.

Of course, compounds like *duck hunting* involve not only a relation between the Noun and the argument structure of the Verb, but one which is like the one we find in Noun Incorporation constructions. Indeed, in both cases the Noun most typically corresponds to an argument of the Verb that would fill the syntactic direct object position and the  $\theta$ -role of THEME. This is not always the case, however. In compounds like *earthquake*, *sunrise*, *landslide*, etc., the Noun apparently represents the subject of an intransitive Verb. In these examples, we might invoke the Unaccusative hypothesis and say that the argument in question represents an underlying Direct Object. Such an account is less plausible, though, for examples such as *crybaby*, *flashlight*, *workman*, *playboy hand laundry*, *cottage industry*, etc. that seem to involve a Noun specifying the agentive argument of an unergative Verb.

What does seem constant about all of these cases is the fact that the associated Noun consistently corresponds to the  $\theta$ -role of THEME. Other compounds, though, correspond to other non-thematic NI types: e.g. *hand laundry*, *cottage industry*. Once we admit the possibility that all of these verb-argument(/adjunct) relations can be established by a lexical rule, as they must be if we are to deal with true compounds, the initial motivation for a syntactic account of NI disappears.

Sapir (1911) was probably the first to propose that Noun Incorporation constructions are actually instances of lexical compounding. Since the possibility of a ‘syntactic’ analysis in the modern sense wasn’t really open to him, though, his arguments for this position came mostly from the formal consideration that incorporated structures involve a combination of two stems, like compounds, together with phonological and semantic idiosyncrasies of the sort adduced above.

## 2 Some (Ultimately Neutral) Arguments

Noun Incorporation, then, is a construction type with some overall properties (the combination in a single word of verbal and nominal stems, where the nominal generally supplies information about one of a small number of thematic types from the argument frame of the verb, and where the whole functions as a verb), and whose analysis can be approached from either of two perspectives (syntactic or lexical) each of which seems to have some natural conceptual affinities with the construction’s basic character. The question to be addressed is which of these (if either; or perhaps some combination of the two) is correct. This may seem like an issue of limited import, but in fact its resolution has major implications for linguistic theory.

Baker’s work has focused for some time on an account of syntactic incorporation and its generalization from the core cases of Noun Incorporation to other head-movements. It is quite important for him, then, that syntactic Noun Incorporation be at least possible (even if other sorts of Noun Incorporation exist too). On the other hand, for those who maintain that “the syntax neither manipulates nor has access to the internal form of words” (Anderson 1992, p. 84), as at least some versions of the Lexicalist Hypothesis require, it is quite important not to allow syntactic rules to put words together in this way. The choice of analyses for Noun Incorporation constructions thus has a good deal of importance for the theory of how morphology and syntax are related. And as noted above, Baker’s theory of head-movement (in support of which Noun Incorporation is the most obvious empirical domain) is central to the articulated-INFL account of clause structure common to much current syntax, so any challenge to the syntactic nature of Noun Incorporation undercuts this view, at least indirectly.

A choice between the two accounts does not seem to follow from the basic descriptive properties of Noun Incorporation. Both views seem to be able to describe the fact that the Noun stem supplies content associated with a thematic position in the argument structure of the Verb. As far as the limitations on which positions can be involved, the lexicalist view accommodates these thematic restrictions rather straightforwardly: lexical rules often refer to the relation of THEME, and compounding in particular does so. If Noun Incorporation is simply a form of Noun-Verb compounding, this is exactly what we would expect. As we have seen, the incorporated Noun is not always a THEME: it is sometimes a locative or an instrumental, as in (3). But this is again

quite parallel to the facts about compounds, and supports the view that there is a single regularity at work in the two cases.

The syntactic view derives this result in a way that is, at a minimum, rather less direct: some might find the account in Baker 1995, sec. 7.3 of why only the Direct Object position is accessible somewhat tortuous, but for the sake of discussion let us assume that such an account is at least possible. Similarly, we will assume that the phonological and semantic idiosyncrasies referred to above can find a home within the syntactic account: idioms, for instance, provide a clear precedent for the assignment of non-compositional interpretations to syntactically complex structures, and it is at least possible that suppletion and other non-phonological variation occurs as a function of syntactically derived environments. If we are to choose between the two theories of Noun Incorporation under consideration, it will probably be on the basis of detailed study, not simply as a consequence of the basic descriptive regularities that define the construction.

Baker proposes to do just this. His primary evidence that Noun Incorporation is syntactic comes from Mohawk, for which an analysis is developed at length in Baker 1995. He suggests that a lexical Noun can be generated (as the exhaustive content of a NP) in argument position, and then moved to adjoin to the governing Verb. Why does this movement take place? According to an argument that he supports in great detail, overt NP's are not licensed in A-positions in Mohawk, but only appear in adjunct positions, where they serve as appositive expressions (a position similar to, but not identical with, the position of Jelinek 1984). If such a [<sub>NP</sub>[Noun]] were generated in an A position, it would thus be ill-formed. In order to get  $\theta$ -marked, such a NP has to be co-indexed with an element in the Verb (a key sub-part of what Baker calls the **Morphological Visibility Condition**, or MVC). If that element were Agreement, the Agreement would (on Baker's hypothesis) absorb the case assigning property of the Verb, so the overt NP would still be ill formed. But if we move the Noun to adjoin it to the Verb, the MVC is satisfied, and the otherwise case-less NP no longer has phonetic content, and no well-formedness conditions are violated.

The Morphological visibility Condition is a parametrically determined characteristic of certain languages (called "Polysynthetic" by Baker, in a usage that deviates somewhat from the traditional sense of this term). We may ask, naturally, which languages have (Syntactic) Noun Incorporation? Baker's answer is: those that have to. That is, in Mohawk, incorporation is forced, as above. In English, on the other hand, as in most languages, the MVC does not hold, and so movement is not forced. But on minimalist assumptions, if you do not have to do something, you have to not do it. As a result, Noun Incorporation is impossible in English, since it is not forced.

Now in fact many languages have constructions that look like Noun Incorporation (i.e., cases where Noun plus Verb together seem to form a single word, and the Noun is interpreted as an argument of the Verb). They differ quite a bit from one another, though. For one thing, in most Noun Incorporation languages, the incorporated Noun is always interpreted as indefinite and/or generic. This is comparable to the interpretation of Nouns in (English) lexical compounds.

- (6) a. She's a truck-driver, which is why she has back problems.
- b.\*She's a truck driver, which is why it's parked over there.

If you say *She's a truck-driver*, that means she drives trucks, not (just) some specific truck. The

*truck* in this compound is not available for anaphoric reference, since it is necessarily generic (or at least referentially under-determined).

In contrast, in Mohawk, an incorporated Noun can refer to something that is referentially specific or definite.

(7) (Mohawk, cf. Baker 1995, p. 288)

Thetáre' wa'-ke-nakt-a hnínu'  
 yesterday FACT-1sS-bed-∅-buy-PUNC  
 I bought a bed yesterday

Í-k-ehr-e' Uwári ʌ-ye-núhwe'-ne'  
 ∅-isS-think-IMPF Mary FUT-FsS-like-PUNC  
 I think Mary will like it (the bed)

This is just what would be expected if the incorporated Noun comes from an NP in an A-position, since such an NP can perfectly well be specific or definite.

If the lexical analysis is to remain viable, it must do something to accommodate this possibility. It looks like the right thing to do is to allow the Verb plus Noun structure that constitutes a lexically compound Verb to take arguments in the thematic position that is (also) specified by the incorporated Noun stem. The most common view to be found in the literature probably is based on the assumption that when a Noun is compounded with a Verb, the Noun satisfies (or ‘saturates’) the corresponding argument in the Verb’s argument structure, with the Noun also being interpreted generically or indefinitely. Assuming that Nouns themselves have an external  $\theta$ -role to discharge (the “R-role,” connected with the Noun’s possibility of referring), this generic interpretation corresponds to a certain sort of binding of the Noun’s own external  $\theta$ -role, along with the corresponding variable in the interpretation of the Verb. We might say that the relevant compounding Operation, like many others, introduces an appropriate generic operator binding the logical variable corresponding to the Noun’s R-role and the verbal  $\theta$ -role identified with it.

To accommodate the facts of languages where an incorporated Noun is potentially referential, we need to extend this account. Let us say that in some languages, at least, Noun-Verb compounding is an Operation that “unifies” the semantics of the Noun with the argument position of the Verb, but without saturating the argument itself. That is, in such a language, *fish-catch* is a transitive Verb meaning “X catches Y, Y a fish.” This is essentially the nature of what Rosen (1989) has called “classifier incorporation,” and which others have proposed for at least some incorporation structures (e.g. Di Sciullo & Williams 1987, Spencer 1995; a similar view was presented orally in Anderson 1985). On this view, the Verb in (8a) would have an associated semantics something like (8b).

(8) a.  $[_V [_N \text{fish}]] [_V \text{catch}]] = \text{“X catches Y, Y a fish.”}$

b. *fish* + *catch*  $\Rightarrow$  *fish-catch*

$\left[ \begin{array}{c} \overline{\text{R}} \\ \text{FISH} \end{array} \right]$  CATCH  $\left[ \begin{array}{c} \overline{\text{Agent}} \\ \overline{\text{Theme}} \end{array} \right]$  CATCH  $\left[ \begin{array}{c} \overline{\text{Agent}} \\ \overline{\text{Theme}(\text{R})} \\ \text{FISH} \end{array} \right]$



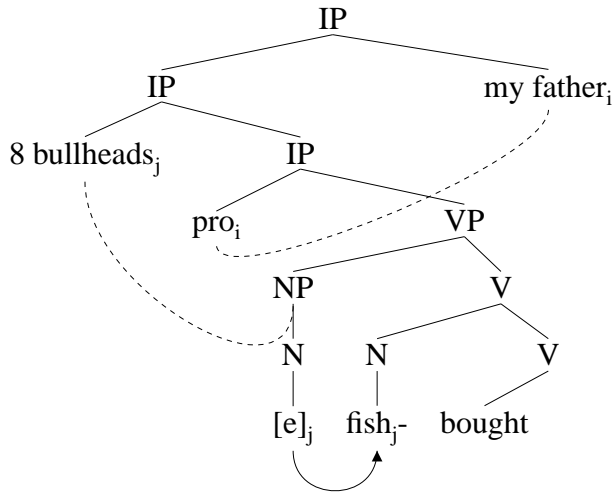
This Verb is still syntactically transitive; and while its THEME argument is specified as having the properties of a FISH, it is not logically or referentially bound, and so the Verb still takes an argument (perhaps just *pro*) which can bear independent reference. In this picture, Mohawk differs typologically from other languages, including those in which incorporated arguments are always necessarily generic, not in having syntactic (*vs.* lexical) Noun Incorporation, but rather in the fact that the lexical Noun Incorporation rule does not saturate (or bind, logically or referentially) the argument position. The semantic operations here are of a sort that we need independently, to deal with lexical operations of compounding and also with phenomena like the lexical suffixes of Wakashan and Salish languages (see Anderson 1992, Bach 1998), and the de-nominal Verbs of e.g. Eskimo (constructions often conflated with Noun Incorporation: see Gerdts 1997).

There is another typological difference among languages with respect to Noun Incorporation. In some languages an incorporated NP can be doubled by an external NP whose content also specifies the corresponding A-position. There are potentially two cases here: first, where the external NP consists only of modifier material, as in *I a new (one) bed-bought*,<sup>3</sup> and secondly, where there is a head N, too, as in *John six bullheads fish-caught*. The first case is straightforward for Baker's original head-movement account: you just say that the head alone moves, leaving any modifiers, determiners, etc. *in situ*. The second case was much more problematic for the analysis of Baker 1987, of course, because how could the head have moved if it is still in its original position? His response was to say that in these sentences, the doubling NP is actually an adjunct, related to a separate (phonetically null) NP in argument position. And of course, on the analysis presented in Baker 1995, that is claimed to be the case for all overt NP's. (9) below provides a schematic analysis in these terms of the relevant parts of a Mohawk sentence taken from Mithun 1984, p. 870. In this structure dotted lines indicate anaphoric links, and the solid line indicates a movement relation.

---

<sup>3</sup>An anonymous reader points out that the lexical and syntactic accounts make different predictions here in one possible case. Consider a language like English, with a required overt head (e.g., *one*) in this construction. Baker's approach seems to predict that such a language could nonetheless have headless NP's just in case the head was incorporated, since in syntactic terms, the phrase would have a head as required. The lexical approach taken here, in contrast, predicts that an overt head will be required, if at all, even in incorporation structures. I know of no data that bears on this issue, but simply record the prediction here.

(9)



Tohka niyohserá:ke tsi nahé  
 several so-it-year-numbers so it-goes  
 sha'té:ku nikú:ti rabahbót wahu-tsy-ahní:nu ki rake'níha  
 eight of-them bullhead he-fish-bought this my-father  
 [Several years ago,] my father bought eight bullheads

How does the lexicalist describe Constructions in which an incorporated Noun Stem is doubled by all or part of an NP in argument position? On that view, when an incorporated N is doubled, we can say essentially the same thing as Baker: the A-position is filled by an empty category, and the overt (adjunct) NP forms a chain with that position in the same way as other overt argument expressions. The only structural difference is that if there is an empty category in the A position, as Baker argues for Mohawk, it is an instance of *pro* rather than of trace. Notice that the presence of the empty category itself follows not from anything about Noun Incorporation, but rather from Baker's independent arguments about the position of overt NP's in Mohawk. Without those, we could also just say that the overt NP is in the expected argument position. Since the incorporated N did not originate there, there is no syntactic reason why some other expression might not fill this position.

In the case of the apparently headless NP's, where it looks as if head movement has taken place, we can note that the expressions themselves are well-formed as NP's in the relevant languages. In Mohawk, as in most languages, there is no overt correspondent of English *one*, so the object phrase in *I want a [new] one* consists of just the Adjective *new*. Such a Noun Phrase arises not by movement of its head out of the NP to adjoin to a governing Verb, but rather by the selection of *pro* as head of the NP, with the semantics of the nominal being supplied from context (or Verb-internally, in a case like 8b). On this view the headed and headless cases of doubling NP's fall together, at least in principle.

With this apparatus in place, we can approach a typology of referentiality and 'doubling' within the class of Noun Incorporation languages. Three major classes of construction must be distin-

guished:

(10) Types of ‘doubling’ of incorporated N’s:

- a. No independent reference, no doubling
- b. ‘Classifier’ incorporation (doubling with or without independent reference)
- c. Free reference, no doubling

The first of these is the case in which the Noun is necessarily interpreted as non-specific and/or generic, and where no doubling expression can be present. In this situation, we assume that the compounding Operation (a) identifies the semantic variables corresponding to the Noun’s external  $\theta$ -role and the appropriate verbal argument; and (b) introduces an appropriate logical Operator binding this combined variable, resulting in a semantics like (11). The syntactic sub-categorization of the resulting Verb is reduced by the elimination of the argument position corresponding to this variable. In some languages, the result is that transitive Verbs become intransitive as a result of incorporation. In others, the morphological properties of the derived Verb may remain transitive, although no argument corresponding to the ‘incorporated’ element is present. We assume that this is an instance of the formal dissociation of morphological and syntactic transitivity, as discussed in Anderson 1991.

$$(11) \text{ GEN}_X (\text{CATCH} \left[ \begin{array}{c} \text{Agent} \\ \text{Theme} \\ \text{FISH} \end{array} \right], \left[ \begin{array}{c} X \\ \text{Theme} \\ \text{FISH} \end{array} \right])$$

Free reference with the possibility of doubling (“Classifier incorporation”) is parallel, but lacks the Verb-internal logical Operator binding (and saturating) the relevant argument: in this case, two semantic variables are available for external specification by expressions in A-positions. Of course, since the semantics of the incorporated Noun Stem have been unified with one argument in this representation, the externally Specified expression must be consistent with that: thus, the information provided by the object of (8b) must (at minimum) be consistent with the property that what it designates is FISH.

Free reference but no doubling (as in e.g. Nahuatl or Southern Tiwa) must be the case where the semantics of N unifies with that of a verbal argument position, but the resulting Verb has some property that is incompatible with an overt NP. Phonologically empty *pro* can appear in the A-position, however, and supplies the possibility of independent reference. What could the relevant property excluding overt argument expressions be? Baker’s own account of a language like Southern Tiwa, where doubling is excluded but non-generic reference in association with an incorporated Noun is allowed, is that the condition licensing adjuncts in association with A-positions only allows them to be licensed by *pro*, and not by a trace (of head-movement); but even this stipulation will not suffice within a theory in which incorporated Nouns are produced lexically (and thus correspond to *pro* in A-position, rather than to the trace of head movement).

What account of the difference between doubling and non-doubling languages is available on lexicalist assumptions? Apparently, argument positions corresponding to an incorporated N have some property that is incompatible with the structure of a Verb derived by Noun Incorporation.

One possibility is the following: suppose that Noun Incorporation in a language like Mohawk (which permits external doubling NP's) not only unifies the semantics of the N with that of the corresponding argument within the semantics of the Verb, but also eliminates the Noun's external R  $\theta$ -role (or perhaps merges it with the  $\theta$ -role assigned to that argument). In Southern Tiwa, on the other hand, the incorporated N's R  $\theta$ -role is identified (but not merged) with the  $\theta$ -role of the argument. That means that if the A position contains an expression headed by a Noun that assigns its own R  $\theta$ -role, there is a  $\theta$ -criterion violation.

Another way in which languages may differ requires further investigation, but the outlines of a solution are reasonably clear. Languages that permit doubling may differ in which of the following schematic sentence types they admit:

- (12) a. John fish-caught three trout.  
 b. John fish-caught a fish.  
 c. (\*) John trout-caught a fish.

As far as is known, all languages that permit doubling of an incorporated Noun stem allow sentences such as (12a), where the doubling expression is more specific and detailed than the incorporated Noun alone. Many (though apparently not all) allow sentences like (12b), where the two are essentially synonymous. None however, appear to allow sentences like (12c), where the doubling expression is actually less specific and detailed than the incorporated Noun. These facts appear to result from a requirement (semantic or pragmatic, depending on one's view of where the line between these is to be drawn) that overt expressions be at least minimally informative with respect to the information already provided as part of the Verb's semantics. (12c) violates this because the information provided by the overt NP is actually a proper subset of that already provided by the Verb, while the overt NP in (12a) clearly supplies information above and beyond that present in the Verb. Languages apparently differ on the basis of whether they consider an essentially equivalent expression 'informative' or not (perhaps by virtue of the possibility it introduces of independent referentiality).

One further point should be noted about the properties of doubling NP's. In many cases, these provide information about the nature of the corresponding argument itself: in a Mohawk sentence like (13), for instance, the overt argument expression and the incorporated Noun combine to specify a single referent 'a polka-dotted dress':

- (13) Kanekwarúnyu wa'-k-akya'tawi'tsher-ú:ni  
 it.dotted.dist past-I-dress-make  
 I made a polka-dotted dress

In other cases, however, the external NP may supply information about a possessor of the argument in question, as in (14).

- (14) Rembarrnga (Dixon 1980):  
 tiŋ? paŋa-warnta-na-∅  
 woman 1sg/3pl-track-see-past  
 I saw the women's track

Note that although it is a single track that is referred to, the agreement is plural, due to the fact that it is the track of more than one woman, even though the NP *tiŋ?* ‘women’ is not overtly marked as plural. The overt argument expression, thus, is not simply something with a phonologically null head (as in English *I was looking for a track, and I saw the women’s* [ $\emptyset$ ]). Note that the agreement is with a third person *plural*, showing that ‘women’ is the head of the object NP, not simply a modifier. In such cases (often referred to as ‘possessor ascension,’ especially in the literature of Relational Grammar; or more theory-neutrally as ‘external possession’ in Payne & Barshi 1999), the semantics must be capable of construing an external argument expression as specifying the possessor of an (already partly specified) argument, not the argument itself.

### 3 Some Less Neutral Arguments

By and large, up this point, the syntactic and lexical theories of Noun Incorporation are “tied” in that each can be said to account for roughly the same range of phenomena the other can. Baker is quite fair: he considers the lexical account (though not with all of the details supplied here), and says essentially the same thing. In fact, he says not only that some languages have lexical/morphological Noun Incorporation rather than syntactic Noun Incorporation but that even Mohawk has this, in addition to syntactic incorporation. This makes the theory rather close to unprovable: any fact that appears to argue against the syntactic story is dealt with by saying that in such a case, the incorporation is lexical.

But Baker also discusses some phenomena which he feels argue for the syntactic account over the lexical one. As the syntactic and lexical views have been elaborated above, they converge to a great extent as far as the representations they assume. But there is one difference: for Baker, the empty category present in A-position in association with a Verb that has undergone syntactic Noun Incorporation is a trace, while in the Lexical Noun Incorporation case it is a *pro*. As a result, any way these two possibilities could be teased apart could provide a way of discriminating between the theories. On this basis, Baker 1995, pp. 314-329 offers three arguments that at least some cases of syntactic Noun Incorporation exist.

The first of these concerns agreement. He argues that in general, there is no agreement with the position corresponding to the source of Noun Incorporation. In this respect, he differs explicitly from Postal (1979), who claimed that there IS agreement with an incorporated Noun. For Baker, agreement occurs with a position containing *pro*, while syntactic incorporation ought to leave not a *pro*, but a trace. Absence of agreement with Noun Incorporation Verbs, as opposed to its presence with simple *pro*, would then argue for the kind of syntactic difference he assumes.

To evaluate this argument, we must consider some limitations on Mohawk Noun Incorporation. In fact, it is almost exclusively inanimates that are incorporated. This is particularly interesting in light of the fact that in Mohawk, agreement with an inanimate object is indistinguishable from no agreement at all. Incorporation of animates is generally disfavored, and regarded as pejorative (implying the treatment of a person as an object, for example). Why should this restriction obtain? Baker admits to having no explanation for why incorporation of animates ought to be avoided. One possibility, though, is that it is a consequence of the fact that it is precisely in this case that

the speaker can “fudge” the issue of whether agreement is or is not present, similar to the use of modal constructions in English in those cases where no particular agreement seems right.

- (15) Neither Fred nor I  $\left\{ \begin{array}{l} \text{??am/??is/??are coming} \\ \text{will come} \end{array} \right\}$  to the party.

Other examples of the same sort are provided by Pullum & Zwicky 1986. They note, for example, that in German, the conjunction of Verbs taking dative objects with those taking accusative objects is only felicitous when the object NP is ambiguous between the two cases.

- (16) a. Er findet und hilft Frauen  
 he finds(+ACC) and helps(+DAT) women  
 He finds and helps women
- b.\*Sie findet und hilft Männer(n)  
 she finds(+ACC) and helps(+DAT) men(Dat)  
 She finds and helps men

The tension in Mohawk which is resolved by limiting incorporation to inanimates might be described as follows: where the Verb contains an incorporated N referring to an overt argument NP, that seems like “enough” to a speaker, and so the presence of a separate agreement element seems superfluous. Where the transitive object is inanimate, however, the agreement morphology is ambiguous as to whether it contains a marker referring to the object or not, so no surface discomfort results.

The bottom line, in any case, is that in most cases of Noun Incorporation, there is no overt indication of whether agreement is present with the ‘incorporated’ position or not, because of the formal similarity between **Sbj/InanimateObj** markers and intransitive **Sbj** markers. In addition, as Baker notes, in cases where there is no external ‘doubling’ NP we can always say that the lexical operation of incorporation has constructed an intransitive Verb (by saturating the argument), so the absence of agreement would follow on either the lexical or the syntactic view.

- (17) Tu-t-a-yako-kétoht-e’ ts-e-wir-aháwi  
 DUP-CIS-FACT-FsO-appear-PUNC ITER-FsS-baby-carry/STAT  
 She appeared carrying a baby.

In some cases, though, there is clearly an animate Noun incorporated: either there is no agreement with this Noun, or else it is (exceptionally) treated as if it were inanimate.

- (18) Ra-wir-a-núhwe’-s thíkΛ (owirá’a)  
 MsS-baby-∅-like-hab that (baby)  
 He likes that baby

This is the sort of incorporation structure that is crucial to Baker’s argument, since (on his analysis) the lack of agreement is forced by the fact that the relevant position is occupied, after movement, by a trace. The lexicalist, on the other hand, is led to say that in such a case, the NP

referring to the baby is treated as inanimate. Since (animate) agreement is obligatory in such cases when incorporation has not taken place, Baker concludes that the argument position in a sentence like (18) must be filled with a trace rather than with (inanimate) *pro*.

Baker (1995, p. 335) asserts that “if Mohawks can freely view babies as inanimate entities, then one would expect that they could trigger inanimate agreement even when not incorporated [as in (19b) below -sra], contrary to fact.” This conclusion does not necessarily follow, however. It might well be the case that the semantics of the incorporating stem *-wir-* ‘baby’ differ subtly from those of the independent Noun *owirá’a*, perhaps in being under-specified for animacy.

- (19) a. shako-núhwe’-s (ne owirá’a)  
 MsS/3pO-like-hab art baby  
 He likes them (babies)
- b. \*ra-nuhwe’-s (ne owirá’a)  
 MsS-like-hab art baby  
 He likes them (babies)
- c. ra-wir-a-núhwe’-s  
 MsS-baby-∅-like-hab  
 He likes babies

On the other hand, it seems that with animate incorporated objects, agreement is at least optional.

- (20) a. Uwári ye(-ruwa)-ksta-hser-áhaw-e’ ne rake-’níha  
 Mary FsS(/MsO)-old.person-nom-carry-impf prt my-father  
 Mary is holding my father
- b. Wa’-ke (-hi)-ksta-hser-áhset-e’  
 fact-1sS(/MsO)-old.person-nom-hide-punc  
 I hid the old person (the old man)

And in fact Baker (Baker 1995, p. 336) observes in a footnote that “when the doubling material makes explicit the gender of the argument in question, the Noun Incorporation plus agreement construction is preferred where possible.” Since agreement ought to be impossible in any case of (syntactic) incorporation via movement, this would appear to be a strike against his analysis.

So how does the Lexicalist derive the result that agreement and Noun Incorporation do not generally co-occur? The facts are obviously rather complicated. There appears to be a preference for avoiding a situation in which both overt agreement material and an incorporated Noun refer to the same participant. One way to resolve this tension is to treat the agreed-with position in a Noun Incorporation construction as if it were inanimate, in which case no overt marker appears: that provides the basis for the preference for inanimates as incorporated elements, and the sense that incorporation of animates is somehow pejorative (since the overt agreement pattern would imply treatment of the NP in question as inanimate). If this could be maintained consistently, we could say that Noun Incorporation constructions do indeed have morphosyntactic agreement, even though this has no overt phonological consequence.

Actually, Baker notes that in some languages which are otherwise syntactically like Mohawk (Tanoan languages like Southern Tiwa and Gunwinjguan languages like Mayali), overt agreement does appear with positions that are also associated with an incorporated NP. In Ainu, which he puts in the class of polysynthetic languages (incorrectly, as shown in Kaiser 1997), some dialects have agreement with the NP position associated with Noun Incorporation and some don't. Wherever we find agreement with an incorporated position in a 'polysynthetic' language, that presents a problem for Baker's analysis. On the other hand, the Lexicalist account derives these cases with little difficulty, since they represent plain transitive Verbs.

In fact, quite adequate responses can also be given, apparently, to Baker's other two arguments. One of these derives from the fact that in Mohawk the object position associated with a Noun Incorporation Verb cannot be questioned with a general-purpose question word. For Baker, this follows on the syntactic analysis from the fact that the A-position cannot simultaneously contain a Noun to be incorporated and a question word.

There are two sub-cases to this argument. The first is that of *who* questions.

- (21) Úhka wa'- $\left\{ \begin{array}{l} *ke \\ khe \end{array} \right\}$ ksá-ht-a-ya'k-e'?
- who FACT- $\left\{ \begin{array}{l} *1SS \\ 1SS/FSO \end{array} \right\}$ CHILD-NOM- $\emptyset$ -HIT-PUNC
- Who (a child) did I slap?

In this case, we see that if agreement is overtly present, the sentence is acceptable; while the absence of agreement leads to ungrammaticality. As Baker points out, the Variant with agreement validates the predictions of the Lexicalist account; but he suggests that the impossibility of the alternative without agreement supports the syntactic account, since it would only follow (according to him) from the incompatibility of agreement and movement traces. But this does not in fact follow: the badness of the sentence without agreement (treated on the Lexicalist view as involving inanimate agreement) can be due simply to an evident agreement conflict. That is, the question word 'who' makes it explicit that the object is animate, which conflicts with apparent inanimate agreement.

For *what*-questions the account is a bit different. In (22) there is no animacy conflict between the question word and the incorporated object 'meat' but the sentence is still ungrammatical.

- (22)\*Nahóta wa-ha-'wáhr-a-k-e'?
- what FACT-MsS-meat- $\emptyset$ -eat-PUNC
- What did he meat-eat?

Baker again concludes that the incompatibility of question words with incorporation in such sentences follows from his syntactic analysis, but there is an alternative that is readily available to the Lexicalist: the badness of (22) would also follow from the requirement that an overt argument has to be informative. Since 'what' adds nothing to the semantics of 'meat-eat' that would further specify the properties of its object argument, it fails to meet that requirement. Again, the lexical account accommodates these facts without syntactic movement.



Baker's final argument is based on condition-C effects, and appears to be subject to some of the same objections that apply to his other points. He notes that in sentences like those in (23), coreference between the matrix subject and the argument corresponding to an incorporated noun is at best marginal.

- (23) a. (Sak) ra-tsháni-s tóka  $\Lambda$ -ke-ksá-ht-a-ya'k-e'  
 Sak MsS-fear-HAB maybe FUT-1sS-child-NOM- $\emptyset$ -hit-PUNC  
 He (Sak)<sub>i</sub> is afraid that maybe I will slap the child<sub>j/?i</sub>
- b. Sak í-hr-ehr-e' ks-kst $\Lambda$ -hser-akéras  
 Sak  $\emptyset$ -MsS-think-IMPF NsS-old.person-NOM-stink  
 Sak<sub>i</sub> thinks the old person<sub>j/?i</sub> smells bad
- c. Sak í-hr-ehr-e' a-ke-kst $\Lambda$ -hser-áhset-e'  
 Sak  $\emptyset$ -MsS-think-IMPF OPT-1sS-old.person-NOM-hide-PUNC  
 Sak<sub>i</sub> wants me to hide the old person<sub>j/?i</sub>

Baker interprets the verb forms in the embedded clauses in these examples as involving an incorporated object and intransitive agreement. He suggests that the difficulty of getting coreferential interpretations would follow if the position corresponding to the incorporated noun contained a trace, rather than *pro*, and that trace exhibited condition-C effects (rather than the condition-B effects we would expect of a *pro*).

Another interpretation is quite possible, however. We might say that the agreement with the positions corresponding to incorporated nouns in (23) is actually *neuter*, rather than missing (as is indeed overt in the case of the basic intransitive (23b)). In that case, the difficulty of getting a coreferential reading for these sentences would follow from an apparent gender conflict between the matrix subject (overtly animate and masculine in all cases) and the lower NP, apparently identified as inanimate. This interpretation is strengthened by Baker's example in (24), where appropriate transitive agreement with the 'incorporated' position appears, and where coreference is perfectly acceptable.

- (24) (Sak) ra-tsháni-s tóka  $\Lambda$ -hi-ksá-ht-a-ya'k-e'  
 Sak MsS-fear-HAB maybe FUT-1sS/MsO-child-NOM- $\emptyset$ -hit-PUNC  
 He (Sak)<sub>i</sub> is afraid that maybe I will slap the child<sub>i/j</sub>

Given the uncertain status of condition-C effects in the current syntactic literature, it would not be wise to put too much weight on this argument in either direction, but it certainly appears that even assuming the disjoint reference principles Baker invokes, the conclusion does not necessarily follow that the position corresponding to an incorporated noun is occupied by a trace, rather than *pro*. Indeed, Baker explicitly suggests that the relevant NP in (24) is *pro* rather than trace.

In summary, the lexicalist account of Noun Incorporation seems entirely viable, even for the phenomena Baker treats as most centrally syntactic.

## 4 Conclusion

So where does this survey of incorporation phenomena leave us with respect to a choice between the two primary views of Noun Incorporation? It appears that even Baker agrees that much Noun Incorporation is in fact lexical, not syntactic, even in the language for which he feels the strongest syntactic case can be made (Mohawk). But in fact, the limited sets of facts for which the syntactic account is said to be necessary can also be accommodated within the lexical account, without invoking mechanisms that have no precedent elsewhere. And that means that a purely lexical account of Noun Incorporation, without syntactic movement, is almost certainly possible. But that, in turn, means that the best putative support for an operation of syntactic head-movement may be non-existent — a conclusion with extensive consequences for many areas of contemporary syntax, especially the split-INFL analysis with its proliferation of functional heads.

## References

- Anderson, Stephen R. 1985. *K<sup>w</sup>ak<sup>w</sup>'ala* morphology. read at Conference on Canadian Native Languages in Theoretical Perspective, SUNY Buffalo.
- Anderson, Stephen R. 1991. Syntactically arbitrary inflectional morphology. *Yearbook of Morphology* 4. 5–19.
- Anderson, Stephen R. 1992. *A–Morphous morphology*. Cambridge: Cambridge University Press.
- Anderson, Stephen R. 1997. Remarks on agreement and incorporation phenomena. *Y.A.L.E. 1: Studies in the morphosyntax of clitics*, ed. by L.Kaiser. 29–44. Dept. of Linguistics, Yale University.
- Anderson, Stephen R. 2000. Lexicalism, incorporated (or incorporation, lexicalized). *Proceedings of the Chicago Linguistic Society*. [to appear].
- Bach, Emmon. 1998. Semantic relations in and outside of complex words. Read at Lexicon in Focus conference, Wuppertal, 19 August, 1998.
- Baker, Mark. 1987. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
- Baker, Mark. 1995. *The polysynthesis parameter*. New York: Oxford University Press.
- Di Sciullo, Anna-Maria & Edwin Williams. 1987. *On the definition of word*. Cambridge MA: M. I. T. Press.
- Dixon, Robert M. W. 1980. *The languages of Australia*. Cambridge: Cambridge University Press.
- Gerdts, Donna B. 1997. Incorporation. *Handbook of morphology*, ed. by A.Spencer & A.Zwicky. 84–100. London: Basil Blackwell.
- Jelinek, Eloise. 1984. Empty categories, case and configurationality. *Natural Language and Linguistic Theory* 3. 39–76.
- Kaiser, Lizanne. 1997. The interaction of noun incorporation and applicative formation in Ainu. *Yearbook of Morphology* 157–178.
- Kroeber, Alfred. 1909. Noun incorporation in American languages. *Proceedings of the sixteenth international congress of Americanists*.
- Lees, Robert B. 1960. *The grammar of English nominalizations*. The Hague: Mouton & Co.
- Mardirussian, Galust. 1975. Noun incorporation in universal grammar. *Proceedings of the Chicago Linguistic Society* 11. 383–389.

- Merlan, Francesca. 1976. Noun incorporation and discourse reference in modern Nahuatl. *International Journal of American Linguistics* **42**. 177–191.
- Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* **60**. 847–894.
- Mithun, Marianne. 1986. On the nature of noun incorporation. *Language* **62**. 32–37.
- Osborne, C. R. 1974. *The Tiwi language*. Canberra: Australian Institute of Aboriginal Studies.
- Payne, Doris L. & Immanuel Barshi, eds. 1999. *External possession*. Amsterdam: Benjamins.
- Polinsky, Maria S. 1990. Subject incorporation: Evidence from Chukchee. *Grammatical relations: A crosstheoretical perspective*, ed. by K.Dziwirek, P.Farrell & E.Mejfas-Bikandi. 349–364. Stanford: Center for the Study of Language and Information.
- Postal, Paul. 1979. *Some syntactic rules of Mohawk*. New York: Garland.
- Pullum, Geoffrey K. & Arnold M. Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language* **62**. 751–773.
- Rosen, Sara Thomas. 1989. Two types of noun incorporation: A lexical analysis. *Language* **65**. 294–317.
- Sadock, Jerrold. 1980. Noun incorporation in Greenlandic: A case of syntactic word formation. *Language* **56**. 300–319.
- Sadock, Jerrold. 1986. Some notes on noun incorporation. *Language* **62**. 19–31.
- Sapir, Edward. 1911. The problem of noun incorporation in American languages. *American Anthropologist* (n.s.) **13**. 250–282.
- Selkirk, Elizabeth. 1982. *The syntax of words*. Cambridge MA: M.I.T. Press.
- Skorik, Pjotr Ja. 1948. *Očerk po sintaksisu čukotskogo jazyka: inkorporacija*. Leningrad: Učpedgiz.
- Spencer, Andrew. 1995. Incorporation in Chukchi. *Language* **71**. 439–489.
- Swadesh, Morris. 1939. Nootka internal syntax. *International Journal of American Linguistics* **9**. 77–102.
- Williams, Edwin S. 1989. Maximal projections in words and phrases. *Alternative conceptions of phrase structure*, ed. by M. R.Baltin & A. S.Kroch. 280–291. Chicago: University of Chicago Press.
- Zide, Arlene R. K. 1976. Nominal combining forms in Sora and Gorum. *Oceanic linguistics special publication* 13. 1259–1294. (Austroasiatic Studies, Part II).
- Zwicky, Arnold M. 1992. Some choices in the theory of morphology. *Formal grammar: Theory and implementation*, ed. by R.Levine. 327–371. Vancouver: Oxford University Press.