On the word in Angami Naga*

P. P. Girdhar
Central Institute of Indian Languages
Mysore, India-570006

ABSTRACT

This paper seeks to expound and juxtapose two alternative approaches to linguistic wordhood as applied to Angami Naga, a Tibeto-Burman language of North East India. In the first of these approaches the criteria, which are laid out with ample exemplification, apply across the board, which results in the definition of only one kind of wordhood (viz. lexical wordhood). In the other approach different criteria apply in response to different levels (viz. the phonological, intraphrasal and interphrasal), resulting in a six-way typology of the Angami word. There is furthermore an attempt to define compound lexical units. Compounds, however, do not lend themselves to being analysed in terms of the second approach. Although this paper is basically empirically inclined, theoretical exegesis perfuses data-elucidation all along. We conclude that the disjunct approach explicates the phenomenon of the word in Angami in a more fruitful perspective than the conjunct approach.

* (a) Angami Naga, phonemically /əŋgəmə nágə/ belongs to the Western subgroup of the Naga group of the Tibeto-Burman family of languages. It is spoken in the hilly state of Nagaland in North-East India. Angami is a tone language, but tone is not marked here because it is irrelevant to the central thrust of this article. Intonation does bear on the problem of the word, but we lack sufficient data to take it into consideration here.

(b) Drs. E. Annamalai, D. P. Pattanayak (my linguist-colleagues) and Mr. Vishwanath Mirle (my redoubtable graduate teacher) have given ungrudgingly of their time and acumen in commenting on a first version. My genial Angami friends Khose Sale, Atha Vizol, and Zakielle Iralu, the willing celerity of whose response has been a great goad in my attempt at plumbing the depths of the language, never hemmed or hawed when accosted with a request for data. My daughter Usha helped tie up some loose ends in presentation, and converted my forbiddingly squiggly calligraphy into an incredibly neat manuscript. My devout thanks to and ready absolutions on all of them.

(c) This is a recension of a paper which was presented to the Third International Conference on South Asian Languages and Linguistics, 13-16 January, 1982, Mysore, India. I'm grateful for some piquant, if skimpy, feedback I got there.
0.0 To sift lexical units (words, compounds included) from the more numerous syntactic units (phrases) on the one hand and the less numerous sublexical units (bound forms) on the other, has seldom been an easy task in any language. As it happens, some of the criteria which make for the strainer are necessary without being sufficient, some sufficient without being necessary, none both necessary and sufficient, some neither. Two criteria may prove dissonant to each other in deciding the morphological, lexical or syntactic status of a linguistic structure, in which case one criterion may override the other, or the criterion that holds the stage in some specific cases may itself be overridden by others in certain others, or a single criterion may prove inconclusive in which case criteria apply in concert... In a word, there is about the defining criteria of the word a seemingly messy criss-crossing asymmetry.

0.1 I shall, in this exposition, sketch out and juxtapose two alternative approaches to wordhood, neither of which is admittedly entirely novel. The comparative juxtaposition will be more implicit than explicit as the differences between the two approaches will be quite obvious. I will prepare the ground for such an exercise by first laying out some well-defined and well-known criteria. These will establish under what conditions (a) phonetic material is to be considered sublexical or bound, and (b) adjacent

---

1 In languages where the rung of the word can be motivated in the linguistic-structural ladder, word segmentation is not an easy matter (especially in nonagglutinative languages). Any definitions of the boundaries of the word which are restricted to a gnomic sentence or two rarely turn out to be unimpugnable. There is undoubtedly more to the linguistic unit of word than can be encapsulated in a couple of laconic sentences. Further, it is difficult to square the notion of a language with no well-motivated level of the word with the formidable position that the word is a primordial category that "mediates between linguistic entities (phrase, sentence, text) and (nonlinguistic) ontologic entities," so that prototypically — ontogenetically and phylogenetically — the speaking animal (homo loquens) is a lexical rather than a syntactic animal. Such languages also go against the position that "all regular word-formation rules are word-based." (Aronoff 1976:21)

2 Although most of the examples that are discussed in this article as compounds were listed under the rubric "word compounding" in my grammar of Angami, there were no heuristics advanced for doing so except the native speaker's intuition — which I dubbed "a rather dubious" criterion (Giridhar 1980:90). Although possibly a necessary criterion, it is never a sufficient one. Intuitional, introspective evidence is typically shored up by some kind of formal-structural evidence in language. Surely, there must be something in what the intuitions are about, which brings them into being in the first place. If the intuitions cannot be buttressed by any formal-structural evidence, one should doubt the authenticity of the intuitions rather than give them the benefit of the doubt. Which heuristics should override which, however, is determined both by metatheoretical and intuitional considerations. Even if the criteria behave erratically, there must presumably be some organizing principle which makes a (compound) lexical unit a (cultural) cognitive reality.
morphemes are to be construed conjunctively, as sub-constructions of a single distributional framework of morphemes (i.e. as a single word), or disjunctively, as constituents of different distributional frameworks (i.e. as different words). These conditions will set "autonomous" units apart from "syntagmatic" units, to use Trnka's terminology (Trnka 1960, quoted in Marchand 1969). In the first approach, these criteria apply indiscriminately across the board, without regard to levels (see Section 1). The underlying principle behind this way of looking at lexical distribution and categorization is "once a word always a word," or "once an affix, always an affix." After noting their interactional behavior, I shall suggest (1.7) a hierarchisation of the heuristics to circumvent the conclusion of the squishiness of the Angami word, a conclusion one is willy-nilly led to by the dissonant character of the criteria broached in our opening paragraph. Angami compoundhood (section 2), however, turns out to be a discrete rather than a squishy notion, because of the mutually agreeable nature of the structural attributes. The further question of deriving the compounds thus arrived at in terms either of transformational or lexical (derivational) rules is outside the scope of this study, as are the questions of morphological word-formation processes and lexemes. Nor will the paper address the theoretically pregnant topics of "word as articulated thought" rather than as a formal unit, and of the relative psychological reality of the two approaches. Changing tack, the analyst can explore an alternative theoretical possibility, viz. that the word is a variable notion and that word divisions could vary in response to linguistic levels, with distinct criteria applying at different levels. In this approach, there is a de novo consideration and definition of the boundaries of the word at each linguistic level. Section 3 expounds this theoretical posture.

1.0 The following characteristics governing the delimitation of the concept of "word" that defines the "upper and lower limits of morphological structure" (Nida 1949:102) in Angami keep simplex words and compounds separate from affixes and syntagms. Criteria 2.1 through 2.4 concern only compounds, i.e. complex words which are derived by combining two or more roots.

1.1. NonInterruptibility

That a linguistic unit which cannot be interrupted is a lexical unit seems a reasonable assumption.
1. unhi - u  
eye  indsg3  
"the eye"  

2. tshũ - lie  
do  imp  
"do"  

Apparent counterexamples are 1a and 2a:  
1a. unhi  keza-u  
big  
"the big eye"  

2a. tshũ  pevi-lie  
well  
"do well"  

It might be argued that keza "big" and pevi "well" interrupt the structures in 1 and 2 respectively, and therefore 1a and 2a are counterexamples to 1 and 2. However, the fact that in Angami inflectional affixes close the phrase — NP or VP (branching leftward or rightward) — buttresses a possible initial hypothesis that 1a and 2a are not the interrupted versions of 1 and 2, and hence are not counterexamples to our statement.4 The question, however, still remains whether 1 and 2 are morphological structures and -u and -lie bound morphemes. Even the

---

3 Abbreviations with their expansions and symbols are: ? = gloss uncertain; 3 = third person (coded in the verb or lexically marked in the subject); = = morpheme boundary (word boundary is represented by space); Δ = the preanalytic, presystematic indeterminacy about (juncture at) linguistic boundaries; ū = lower-high retracted unrounded central vowel, the retraction giving it a cock-eyed quality (inadequately described as a central vowel in Giridhar [op. cit.:7]). This cock-eyed quality (for lay ears, that is) is even more conspicuous in Mao Naga, a sister language; n = palatal nasal; η = velar nasal; ⇒ = "can possibly become"; ⇒ = "cannot possibly become"; Ω = zero; acc = accusative case marker; com = common gender; con = conditional mood marker; cont = continuous participialiser; conts = contrastive particle; conj = conjunctive participialiser; dat = dative case marker; def = definite suffix; dim = diminutive; emph = emphatic particle; ex = expositive mood marker; fem = feminine gender marker; fut = future tense marker; futg = future negative indicative; gen = generic proform; hab = habitual; imp = imperative mood suffix; indsg = individuated singularity marker; loc = locative; masc = masculine gender indicator; nom = nominative case marker; pl = plural marker; prf = perfective aspect marker; prg = progressive aspect marker; prpsv = purposive marker; prs = present tense marker; prt = participle marker; pst = past tense marker; Q = Interrogative marker; qut = quotative; subj = subject; vr = valency role marker which indicates the number and nature of "players". See Giridhar (op. cit.:63-65) for a brief sketch.

4 This of course could be true only of subordinate or attributive phrases, and not of coordinate phrases. In lesūda mu likhuo-u "book(s) and the bag", u closes only the morphological structure that it forms along with likhuo "bag" and not the whole phrase. In fact, as the gloss shows, lesūda in the above phrase could mean "books".