

**Kansas Working Papers
in
Linguistics**

Studies in Native American Linguistics X

edited by

John Kyle

**Volume 24, Number 2
1999**

Partial funding for this journal is provided by the
Graduate and Professional Association of the University of Kansas

ISSN 1043-3805

© Linguistic Graduate Student Association
University of Kansas, 2000

Kansas Working Papers in Linguistics
Volume 24, Number 2
Studies in Native American Linguistics X
1999

The Lakhota Definite Articles and Topic Marking Traci S. Curl	1
Switch-Reference and the Identification of Small Pro in Hidatsa John P. Boyle	17
Cherokee Clitics: The Word Boundary Problem Marcia Haag	33
Paragraph-Level Switch-Reference Markers in Chickasaw Conversation Cynthia Walker	45
Ojtlán Chinantec Phonology and Morphology Monica Macaulay	71
Glottalization in Nuu-chah-nulth in Optimality Theory Eun-Sook Kim	85
On the Relationship Between Mixe-Zoquean and Uto-Aztecan Søren Wichmann	101

CHEROKEE CLITICS: The Word Boundary Problem

Marcia Haag
University of Oklahoma

Abstract: The problem of identifying Cherokee clitics is complicated by the fact that the prosodic word, marked by the presence of a tonal boundary, may not match the morphological word. Clitics may or may not respect the the word boundary as marked by tone: they may attach outside the boundary tone or carry the tone themselves. The data suggest Cherokee recognizes the morphological word and affix as independent linguistic phenomena, such that simultaneous alignment of morphological with phonological edges is sometimes not realized. Also raised is the question of syntactic boundaries as another parameter in the definition of clitics.

Introduction

In underdescribed, and especially unwritten languages, the problem of deciding on a word boundary is not at all a transparent one.¹ The word boundary figures crucially in the assignment of morphological status—affix, clitic, word. In a language such as Cherokee, where word formation is unusually complex, the question of where a word begins and ends is complicated by tonal phonology and second-position attachment sites. Ultimately, decisions about clitic status will involve a calculus upon three interdependent linguistic units: an attachment site determined by the morphology, functional versus lexical grammatical status determined by the syntax, and prosodic word status determined by the phonology. A morphological primitive 'clitic' is not discernible in Cherokee.

One problem encountered in establishing what a clitic might be is the regular use of this term, especially in Indo-European languages, to refer to the person arguments of verbs, if these are not free-standing pronouns. This usage reaches perhaps its ultimate extension in cases such as Pashto, where person markers, termed clitics, are able even to attach between the nucleus and coda of monosyllabic roots (van der Leenw 1995). In such cases a clitic is defined morphologically as any identifiable piece that can be associated with phi-features and is not also clearly a word, meaning, I believe, prosodically a word.

Dan Everett (1996), in his careful analysis of clitics-as-phi-features, says that ultimately this is a relation in syntax, and not a morphological primitive; in this case the term 'clitic' might serve as short-hand in describing the morphs that encode phi-features, but has no

independent reality. The effects of syntactic consanguinity, in the case of verbs and phi-features, and syntactic distance, in the case of the Cherokee clitics presented here, will prove to be one of crucial elements in determining assignment of a morph to clitic status.

Jerrold Sadock (1995:260) takes an interesting view on clitics when he offers a ‘sociological definition of clitics’, which is “A clitic is an element whose distribution linguists cannot comfortably consign to a single grammatical component.” My analysis of the Cherokee data takes a bolder view in positing that clitics are morphological but not prosodic words.

Right-Edge Morphs

Cherokee morphology cannot easily accept clitics as being synonymous with person markers. These are affixes, unless a clitic has no morphological pertinence at all.

The order of morphs on the Cherokee verb places negators and conditionals, locatives, directionals, the indicator of plural objects, and some aspects before the forms that mark phi-features.

- (1) ʔa yi-wi-d-**iji**-kato-sdi
 not neg-away from speaker-loc-**I**/**them**-look-tns
 I am not looking at them. (from L. Anderson)
- (2) i-n-da-**gwa**-t-sta-n-cl-i
 repeat-simultaneous-fut-1sg-reflex-happen-caus-dative-fit
 It will happen to me. (from L. Jordan)

Leaving the question of pronominals aside, far more interesting is the group of forms that always occur at the right edge of a word, and so fall more easily into the group of what Zwicky (1977) calls simple clitics, morphs that require phonological support from a more adequate host. Three of the forms are question markers; of these at least two have main-clause scope. One of the forms marks potentiality, with at least scope over the verb phrase, and two are focus markers. What is important here is that several of the forms are clearly syntactic functional heads, and so not inside a word, X^0 , lexical projection.

The question of clitic status in Cherokee turns on the notion of the Cherokee word boundary, and for that we need to look at the prosody. The status of tone or pitch distinctions in Cherokee remains unclear. The best synopsis I can give to date is that there seems to be a tone system, that is perhaps evolving from a pitch system (Wright 1996), but that contrastive tone seems to have syntactic and discourse functions, rather than lexemic contrast (Scancarelli 1987). However, there are at least a few words that are lexemically contrasted

by tone alone (based on work with Keith Johnson). I would have to say that no one has thoroughly explained the tone system, but what all researchers agree on is something called the boundary tone, after Lindsey (1985).

In the following sentence, each of the strings divided into words, except for *hest*, the negative imperative, has a high tone on its final syllable. For this example only, this nuclear vowel is marked with an accent *´*. This is predictable, and has been described as a ‘falling superhigh’ by Feeling and Pulte (1975) (unless preceded by a superhigh, in which case it is a normal high). Lindsey calls this a ‘boundary high tone with optional upstep’. Wright (1996:12) calls it ‘high tone’, treating it as underlying high with phonetic upstep, and says it remains a question if it is associated late in derivation or underlyingly specified high.

- (3) Hest gige ijanuwesti ana:galiba ale: ahya:dagwalosgi?a
 don’t red 2pl/wear lightning and thunders
 Don’t wear red when there is lightning and thunder.

The question is whether the boundary tone has real significance as a prosodic marker of the end of the prosodic word. If it did, we might be able to make inferences about the status of morphs and their relative position with respect to this overt prosodic word boundary. In earlier work (Haag 1997), which used Anderson’s 1995 work as a launching position, I showed how the question marker *sko* reliably attaches to the first morphological word of a clause, Wackernagel-style, with a concomitant appearance of the boundary tone, including lengthening of deficient vowels, on the host supporting *sko*. My hypothesis was that Cherokee morphs of the *sko* type, were functional heads requiring attachment to the right edge of the first morphological word on the left syntactic boundary, with phonological repair of the host if necessary. This meant that a too-light morphological word would be boosted in weight and given the word-marking boundary tone.

In (4), the question marker *sko* appears after the first morphological word, *kahwi* ‘coffee’, which is also a prosodic word, minimally two moras. The final syllable of *kahwi* carries the boundary tone, and *sko* has a low tone.

- (4) Kahwi-sko jaduli.
 coffee-Q 2sg-want
 Do you want coffee?

In (5), the definite demonstrative *na* is not a prosodically adequate word, having only one mora, but as the first morphological word, its right edge serves as the adjunction site for *sko*. In this case, the vowel is lengthened, and the boundary tone assigned to the final mora of *naa*. *sko* has a low tone.

- (5) Naa-sko anaskaya ajalagi
 def-Q man Cherokee
 Is that man Cherokee?
 na → naa

Sko also attaches at the right boundary of the first word of word-word compounds:

- (6) Anije-sko-yusti (anijeyusti 'green')
 sprouts-Q-like
 Is it green?

When I tried to generalize this solution over the set of edge-final morphs, I ran into widespread lack of conformity. (The way I selected the set of 'edge-final morphs' was from the pedagogical literature. These are all termed suffixes no matter what their function or semantics because of the impression of the speakers that they are attached to the string in front of them.)

The most important contrast involves two question markers, *ske* from above, and *ke*, a disjunction marker. *Ke* is often used as a question marker, because if no contrast is provided lexically, it may be interpreted as 'or not?'

The following two sentences are interchangeable in speech.

- (7) Janti-sko na askaya ajalake y-igi ajada-ke
 you-know-Q that man Cherokee if-be Choctaw-disj
 Do you know if that man is Cherokee or Choctaw?
- (8) Janti-ke na askaya ajalake y-igi ajada-ke
 you-know-disj that man Cherokee if-be Choctaw-disj
 Do you know (or not) if that man is Cherokee or Choctaw?

While *ske* and *ke* clearly have main clause scope, and would seem to be the same kind of syntactic functor, they do not have the same way of attaching to their hosts.

Looking at the phonology in examples (9) and (10), *ske* has a low tone, and always appears after the boundary tone. But *ke* seems to coincide with the boundary tone.

Also, *ske* provokes lengthening of the final vowel in *janti*, to produce a low-high contour, but the final vowel of *janti* is phonetically both short and high when *ke* is attached.

(9) LB -L
Jantii-sko

(10) H-B
Janti-ke

The situation with *ke* seems to be what Lindsey is referring to when he says (1985:139)

The clitics, like underlying final vowels, are associated with a H% boundary tone, so that H%'s may be "stacked"; there is some evidence (including Durbin Feeling, pc.) that Upstep applies iteratively, creating a gradual increase in pitch over a series of clitics. This question, however, requires further study.

My observation is that there are several clitics that are not associated with boundary tone, and I have never heard iterative upstep in any speech, though it may well occur. But Lindsey's suggestion can't be the general case, as will become clearer.

In the following example, we see that it is not the local phonology that accounts for the difference in the *ska/-ke* attachment. When *ska* is clipped to just the -s, the way it is usually pronounced by today's speakers, thus losing an association site for its low tone, the final vowel of the host is still lengthened, and a boundary tone appears. The abstraction of *ska* triggers the same phonological response that the full form does.

(11) LB
Jantii-s

In the minimal contrast, that of a phonologically deficient host (I have been using the definite demonstrative marker *na*), both *ska* and *ke* trigger lengthening of the vowel of the host, but *ska* triggers a rising contour, while *ke* may trigger phonetic anticipatory high tone, with *ke* itself providing the boundary high tone, consistent with what we have seen.

(12) LB- L
na-a-ska

(13) LL-B
na-a-ke

So *ke* may coincide with the boundary, but it does not affix itself to *na*: a word of the prosodic form *nake* is a well-formed Cherokee word. Observationally, when *ke* attaches to an adequate host, it simply forms the boundary tone, without imposing distortions upon the host, but when that host is not adequate, it must be repaired *before* attachment. *Ke* would seem to be seeking a morphological word as an attachment site; it itself cannot be counted in the prosody.

Comparison of Edge Morphs

I have selected seven edge morphs for comparison, and noted that they have different tones, and the same tone in whatever environment I've heard them in. If Cherokee is a tone language, this is certainly what we'd expect.

	<u>morph</u>	<u>tone</u>	<u>meaning</u>
	hΛ	high falling	'but (contrastive)'
	gwi	high falling	'still'
(14)	ju	high falling	'rhetorical Q'
	na	low	'and as for (focus)'
	sko	low	'yes/no Q'
	ke	boundary	'disjunction, Q'
	le	boundary	'potentiality'

Let me point out that *na* definite is distinct from clitic *na*, which is a focus marker. They form a common pronominal: *naa-na* 'that one', which has the same prosody and tone pattern as *naa-sko*.

In Fig. 1, the output of the pitch traces of *nih-na-hΛ* 'but what about you?'; *nih-na* 'and as for you'; and the clipped version of the latter *nii-na*, I show an example of serial clitics, *na* and *hΛ*. We see that *na* occurs after the boundary tone, has a low tone, and is lengthened itself when *hΛ* is attached. *hΛ* has a high falling tone.

In the two other versions, including the clipped version of *nih*, where the final syllable is deleted, the remaining syllable is lengthened, and a contour tone applied to it so that it will have a boundary tone. Notice from the spectrogram that *na* is short unless it has to serve as an attachment site for a clitic.

This last point is interesting: if the host is prosodically adequate (*nih-na* has three syllables and three moras), why does *na* have to be lengthened so that *hΛ* can be attached? This suggests that the prosody has 'started over' with *na*, but omitted a boundary tone.

Variability of the Boundary Tone

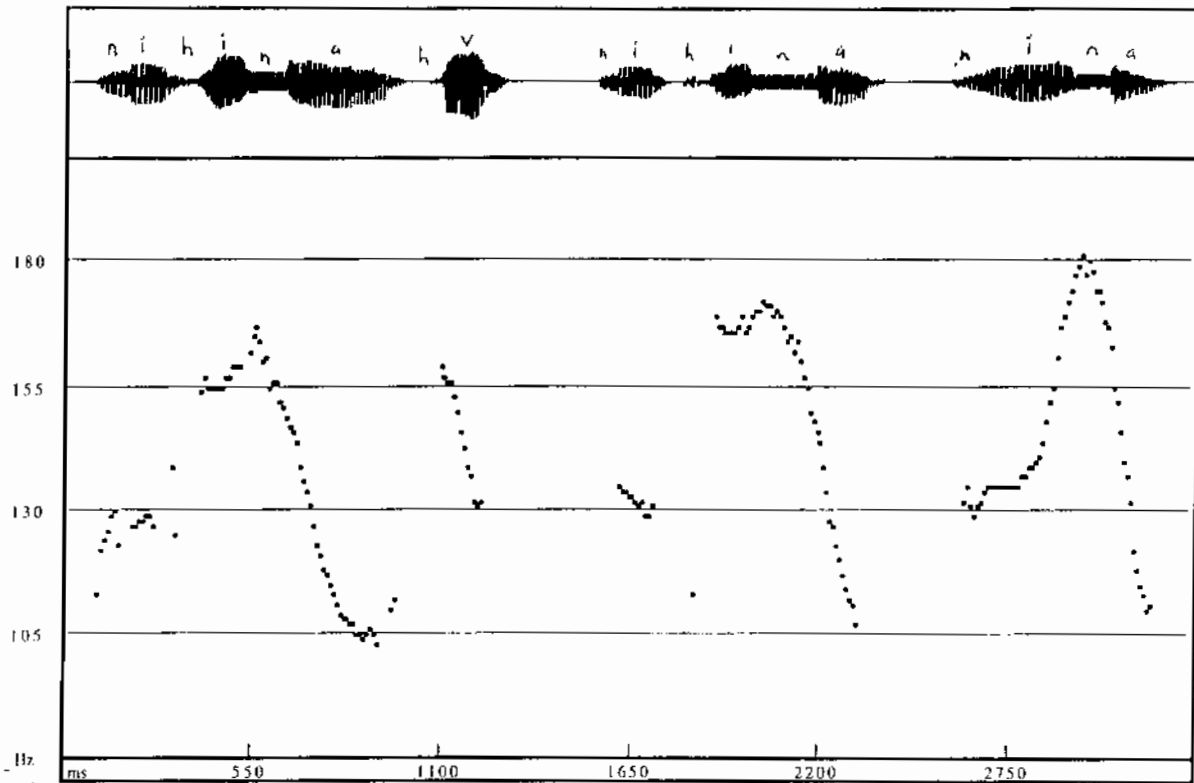


Figure 1: nihi-na-ha and variants (Keith Johnson 1998)

In these two examples of pitch traces, we see two words with high rising boundary tones *dalonige ajulsgile* ‘possibly a yellow flower’ (Fig. 2), and two words with high falling boundary tones *jalagi sagwu* ‘Cherokee One’ (Fig. 3). The difference in these two speeches that immediately strikes me is that the high rising tones are associated with predicative words, and the high falling tones are associated with a name ‘Cherokee One’, the name of a class. In contrast, the speech *na askaya ajalaki* ‘that man is Cherokee’ contains high rising boundary tone on the word for ‘Cherokee’ *ajalaki*. While this is a small amount of data, it does lend support to the hypothesis that tone has a syntactic function (Scancarelli 1987).

Moreover, a number of words do not have boundary tones, as we saw with *hest* in example (3), or have a high tone elsewhere in the word. In particular, a number of derivatives have final boundary tone in a base word, but high tone moved leftward in a class of derivatives involving the suffix *ʔi*. In these examples, boundary tone or displaced boundary tone is marked with *ʔ*.

- (15) taliné ‘two’ talinéʔi ‘second’
 sadusiné ‘eleven’ sadusinéʔi ‘eleventh’

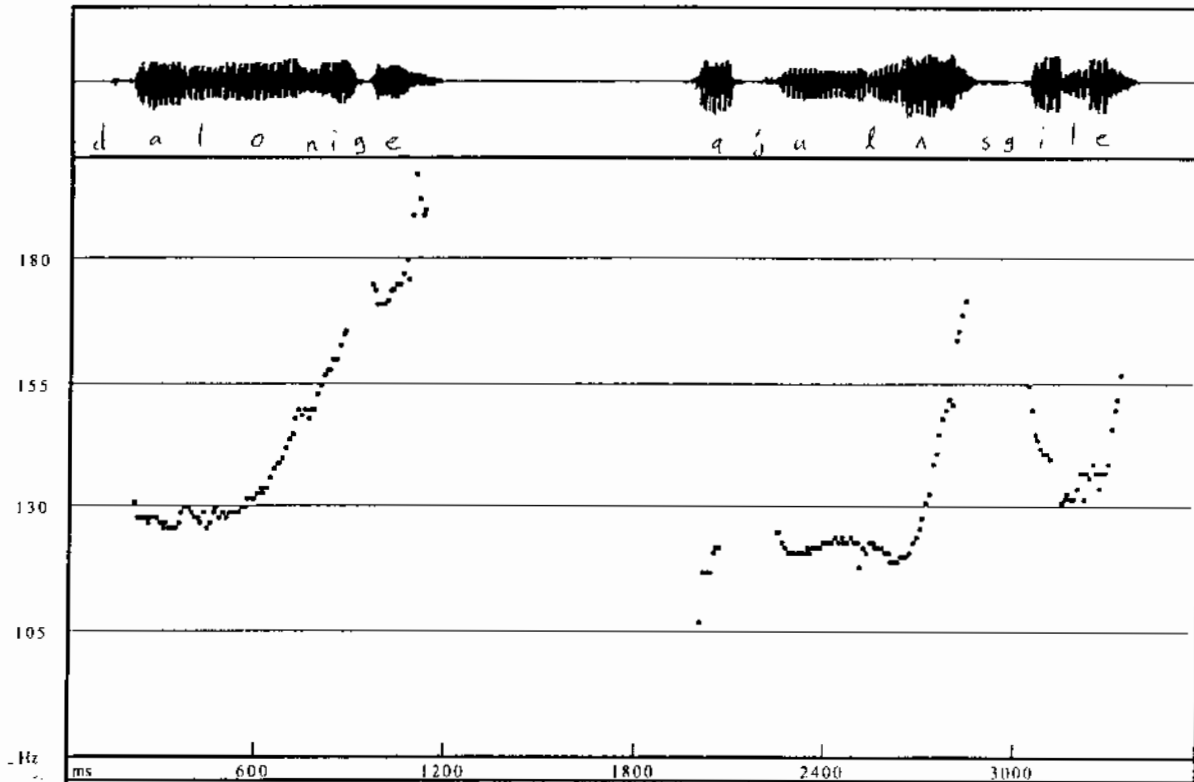


Figure 2: high rising boundary tones (Keith Johnson 1998)

In (15), the derivation of the ordinal numbers from the cardinals, the *ʔi* suffix behaves like the *ske* question marker: it appears outside a boundary tone, yet the effect is seemingly derivational, that is, it is a word-level, X^0 level, effect.

Besides the regular derivation with *-ʔi*, other lexicalizations occur where high tone placement distinguishes the forms: that is, where a boundary tone contrasts with a left-displaced high tone. In the unmarked form, the gloss is of the form 'the X one', and the boundary tone occurs to the left of *ʔi*. In the lexicalized form, the gloss is unpredictable, and the boundary tone occurs on *ʔi*, the final syllable.²

The derivations with *ʔi* above are troubling because they cast doubt on the conclusion that a boundary tone marks a prosodic word boundary. In my previous work (Haag 1997) I had concluded that Cherokee is sensitive to the difference between the morphological affix and the morphological word, and that the latter exists independently of the prosodic word. Attachment sites for clitics were the major evidence for this conclusion. However, the derivations with the *ʔi* morph above weaken the predictive power of the phonology even further: It may be that the tone system, with its presumptive underlying final high tone on lexemes,

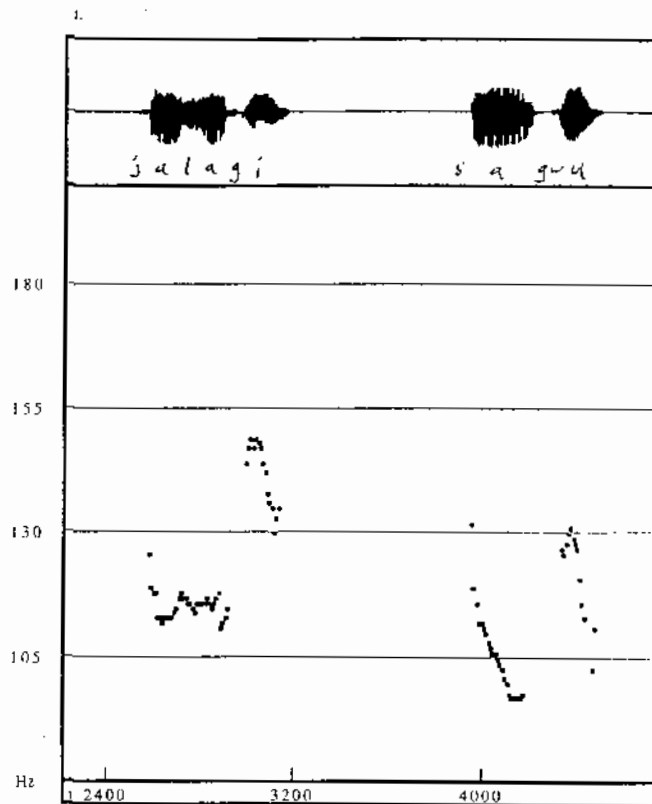


Figure 3: high falling boundary tones (Keith Johnson 1998)

will not distinguish the edge of a lexical head from the edge of a functional head (the case of *ke* and *le*. And it will not distinguish derivation from functional heads (the case of *ʔi*).

However, I'm investigating another hypothesis, which is that *ʔi* is actually a relativizer, so that these seeming derivations are actually of the form: 'the one that...' Even in the case of the ordinal numbers, this could as well be analyzed as *taline* 'number two', and *talineʔi* 'the one that is number two'. If this is the case, we can better preserve the argument that the boundary tone marks the morphological word, though every morphological word is not inevitably marked with a boundary tone. Because Cherokee is an endangered language with ever-fewer speakers, the data is more prone to frozen forms of formerly productive rules.

A Calculus for Clitics

Deciding on clitic status for the Cherokee candidates, and by implication those of other languages, depends first on whether the morph properly occurs outside the word. Second, the morph must have some constrained, obligatory position that is also prosodically deficient. This is because a morph inside a word boundary is generally supposable to be an affix, and prosodically adequate forms are presumably words. Thus, the question must be decided

based of what kinds of grammatical features are proper to words and what are proper to the phrase-level syntax. Beard (1992) and Anderson (1992) have done extensive work on this topic: a number of syntactic categories clearly mark higher-ordered functional categories; these would include complementizers, negation nodes, question markers, and the like. Other grammatical features are far more problematic in that they tend to be included in the inflection, or morphological exponence, of a word: these include, to use the category verb as an example, person markers (phi-features), aspect, tense, and voice. Beard's paradigm, which I agree with, assigns phi-features (dominated by Agree) and tense to phrase-level syntactic categories, while aspect and transitivity (I have termed this property Valency (Haag 1996), a verb's obligatory selection of its arguments) are lexical features. The distinction between lexical and phrasal grammatical features is important in arriving at a decision about what can be outside the morphological word.

The Cherokee data I have selected here, particularly the *sko-ke* contrast, make clear examples because they mark a high-ordered functional category (the affirmative or negative question) and are not arguably lexical features. (No one would argue that questions of any kind grammatically inhere in any of the lexical categories.)

In the case of the Cherokee clitics, the interaction of three linguistic domains, morphology, syntax, and phonology proceeds along these lines:

- (1) The morphology 'decides', or governs the phenomenon of the occurrence of, what entities are morphological words, and that attachments are made on the right boundary, but not the left, of the morphological word (and the first morphological word of the clause, in the case of *sko* and *ke*).
- (2) The syntax decides the level of the morph, whether it is a syntactic category above word level, or a word-level grammatical feature.
- (3) The phonology decides locally on the viability of the host-plus-morph as a prosodic unit. Weight compensation in the form of vowel lengthening of the host may be necessary, but the phonology cannot reject a morphological word and its attachment site because of phonological inadequacy.
- (4) If an attached morph marks a syntactic category above word-level, that morph is a clitic. If it is a lexical grammatical feature or categorial property, it is an affix.

By this calculus, *ke* is a clitic even though it co-occurs with the boundary tone and so could not be said to be 'outside' the word phonologically, as is *sko*. This is because *ke*, as a question marker, cannot be a word-level feature and so cannot be an affix. On the other hand, the syntax can show that these morphs are outside the word, but it cannot show they are not words themselves. It is the fact that they impinge on the phonology of the left-lying morphological word that raises the question in the first place.

As Sadock suggests, we cannot find a single diagnostic that will pick out clitics as a morphological primitive. In this model, the fact that *ke* does not behave like *sko* is perfectly acceptable: rather than expecting two question markers to have comparable forms, we note that both have an obligatory position, are functional heads, and cannot therefore be affixes; and both are phonologically linked to the word to their left, and so cannot be prosodic words. In the Cherokee data, we are not positively diagnosing clitics, we are binning residue from three interdependent linguistic modules. This implies that different languages will have different categories that appear as clitics, and some languages may have no clitics at all.

The interesting cases remain those in which syntactic categories also show morphological agreement on the lexeme; the best studied is, of course, person marking morphs of verbs. In a solution that permits both syntactic status of phi-features in Agree and attachment inside a prosodic word boundary, Everett (1996) posits the attachment of Agree as either the complement of a (morphological) word, resulting in an affix, or as an adjunct of a prosodic word, resulting in a clitic.

Arguing from the perspective of lexical categorial properties, phi-features that are expressed as verbal affixes can be expected as a result of Valency (or Beard's Transitivity), an inherent verbal property. Since verbs obligatorily marshal the number and roles of their arguments, if person features are morphologically bound to grammatical roles (eg. third person singular accusative), they can be expected to be registered morphologically on the verb, as a consequence of that information being inseparable from such information as the transitivity valence of the verb. Again, person features are not themselves inherently verbal, but agree with those features in an argument that may be overtly specified by a verb.

NOTES

¹ I thank Bobby Joe Blossom, Keith Johnson, Linda Jordan, and Laura Anderson for help with the Cherokee data, and Daniel Everett for helpful comments.

² These data are from Margaret Bender's field notes of North Carolina; I have permission to state the general rule of the data but not to cite them.

REFERENCES

- Anderson, S. 1995. Rules and Constraints in Describing the Morphology of Phrases. In *Papers from the 31st Regional Meeting of the Chicago Linguistic Society. Volume 2: The Parasession on Clitics*, 15–31, Chicago. The Chicago Linguistic Society.
- Anderson, S. R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Beard, R. 1992. *Lexeme-Morpheme Base Morphology: A General Theory of Inflection and Word Formation*.
- Everett, D. 1996. *Why There Are No Clitics: An Alternative Perspective on Pronominal Allomorphy*. Arlington, TX: Summer Institute of Linguistics.
- Feeling, D., and W. Pulte. 1975. *Cherokee-English Dictionary*. Tahlequa, OK: Cherokee Nation of Oklahoma.
- Haag, M. 1996. *Lexical Categories in Choctaw and Universal Grammar*. PhD thesis, State University of New York at Stony Brook.
- Haag, M. 1997. Cherokee Evidence for Separation of the Morphological and Prosodic Word. In *Papers from the 1997 Mid-America Linguistics Conference*, 343–349, Columbia, Missouri. University of Missouri-Columbia.
- Lindsey, G. 1985. *Intonation and Interrogation: Tonal Structure and the Expression of a pragmatic Function in English and Other Languages*. PhD thesis, University of California at Los Angeles.
- Sadock, J. 1995. A Multi-Hierarchy View of Clitics. In *CLS 31. Papers from the 31st Regional Meeting of the Chicago Linguistic Society. Volume 2: The Parasession on Clitics*, 258–279, Chicago. Chicago Linguistic Society.
- Scancarelli, J. 1987. *Grammatical Relations and Verb Agreement in Cherokee*. PhD thesis, University of California at Los Angeles.
- van der Leeuw, F. 1995. Cliticization as Alignment to Morphological Slots. In *CLS 31: Papers from the 31st Regional Meeting of the Chicago Linguistic Society, Volume 2: The Parasession on Clitics*, 168–180, Chicago. Chicago Linguistic Society.
- Wright, R. 1996. Tone and Accent in Oklahoma Cherokee. In P. Munro (Ed.), *Cherokee Papers from UCLA*, no. 16 in UCLA Occasional Papers in Linguistics, 11–22. Los Angeles: UCLA.
- Zwicky, A. 1977. *On Clitics*. Indiana University Linguistics Club.