Studies in Native American Languages IV

*Kansas Working Papers in Linguistics*

*Volume 10, number 2, 1985*

# FOREWORD

It is indeed gratifying to recognize the degree of acceptance the *Kansas Working Papers in Linguistics* has come to enjoy, and this is especially true for the series of *Studies in Native American Languages*. Even before the call for papers went out in the fall, we had received inquiries from prospective contributors, and the response to the call itself was remarkable in quality as well as diversity.

This year the *KWPL* marks its first decade of existence, and we are publishing two numbers. Number one is devoted to theoretical issues, general linguistcs and old-world languages, while number two is the fourth in the *Studies in Native American Languages* series. This number includes articles representing seven different language families from all over North America (Uto-Aztecan, Muskogean, Yuman, Siouan, Otomanguan, Athabaskan and Algic), and a great deal of original scholarship.

We wish to thank the contributors, both those whose papers appear in this volume, and those whose papers we did not include. We also wish to thank the faculty of the Linguistics department of the University of Kansas for their support and encouragement for the *KWPL* throughout the year.

# INTERNALLY HEADED RELATIVE CLAUSES IN CHOCTAW

George A. Broadwell

## 0. Introduction

In Choctaw, as in several other American Indian languages, relative clauses may occur without apparent external heads[1,2]. The NP which is the semantic head of the clause occurs within its relative clauses, as in the following examples:

(1) [ John-at    ofi aaipa nota-ma    aa-pisa-tok-ma]
    John-NOM   dog table under-ACC  LOC-see-PAST-DS

    homma-atok.
    red-PAST.

    'The dog John saw under the table was red.'
  or 'The table John saw the dog under was red.'
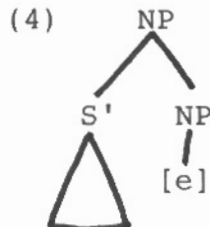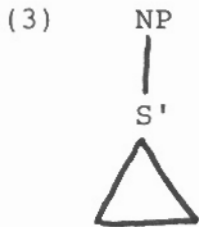
(2)     Joyce-at [John-at ofi  aaipa nota-ma    aa-pisa-tok-ma]
        Joyce-NOM John-NOM dog table under-ACC LOC-see-PAST-DS

        chopa-tok.
        buy-PAST.

        'Joyce bought the dog John saw under the table.'
      or 'Joyce bought the table John saw the dog under.'

Let us call relative clauses of this type internally headed relative clauses (IHRCs).

What is the syntactic structure of such RCs? Since the clause as a whole serves as an NP at S-structure, one proposal is that they have a structure like (3) or (4).[3]

(3)   NP         (4)     NP                    NP
      |                 /  \          or      /  \
      S'              S'    NP             NP      S'
     /\              /\     |              |      /\
    /  \            /  \   [e]            [e]    /  \
   /____\          /____\                       /____\
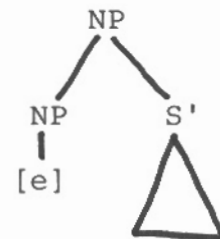
In both structures, some coindexation is assumed. In (3), the internal head is directly coindexed with the NP dominating S'. In (4), the internal head is coindexed with an empty NP position. An important similarity between the structures is that the IHRC is dominated by NP in both cases. Since a verb such as 'buy' in (2) above selects an NP object, it is often assumed that an NP must dominate the IHRC to satisfy the verb's categorial requirements.

An important concept underlying this assumption is the Projection Principle (Chomsky 1981). Informally, this principle requires lexical subcategorization to be satisfied at all three levels of grammar (D-Str, S-Str, and LF). If we believe that sucategorization frames specify both the theta-roles and the categories of their arguments, then in its strong form, the Projection Principle specifies that the categorial requirements of lexical items be satisfied at all levels of grammar.

However, Pesetsky (1982) has suggested that this is not a necessary or useful requirement for the grammar. It is possible to restrict the Projection Principle to a requirement on theta-role assignment. Categorial selection might then apply only at LF, allowing constituents to occur as complements at D-structure or S-structure even if the constituent is not of the subcategorized category.

Alternately, there might be some theory of complementation that will allow us to predict the category of an argument from its theta-role. This possibility is suggested in Grimshaw (1979). Under either assumption, it seems possible to separate categorial selection from theta-role assignment, and to restrict the Projection Principle to the latter.

In this paper I will explore the possibility that there is no NP node dominating IHRCs at S-structure. I will assume that Move-alpha applies to S-structure representation and produces LF structures which satisfy categorial selection. This rule generates a different LF representation for each possible reading of the sentences in (1) and (2). The argument consists in showing that an S-structure analysis of IHRCs as S-bars is superior to either (3) or (4).

Note that (3), as formulated above, violates the endocentricity requirement of X-bar theory (Jackendoff 1977, Stowell 1981), since NP has no head, only a complement. Although we will reconsider (3) somewhat later, at this point let us rule this structure out as ill-formed. The main alternative is then (4), where the IHRC has an empty head. I will refer to a RC with a structure like that of (4) as an empty-headed RC. The empty head analysis is contrasted with an LF movement hypothesis.

There are several empirical differences between the theories. On the basis of evidence from subjacency and switch reference, I will conclude that an LF movement analysis is superior for Choctaw.

## 1.  Subjacency

In languages with headed RCs, like English, elements inside a relative clause cannot be moved.  This is due to the subjacency condition, which prohibits movement over more than one bounding node, where bounding nodes are defined as NP and S.[4]  Choctaw is an exception to the general prohibition against extraction from RCs.[5,6]

(5)  Katommako [ [John-ato ofi __ aa-pisa-tok-ma]
     Where        John-NOM  dog   LOC-see-PAST-DS

     homma-atok] ?
     red-PAST

     'Where [the dog [John saw __] was red] ? '

(6)  Katommah [Joyce-at [John-at ofi __ aa-pisa-tok-ma]
     Where      Joyce-NOM John-NOM dog  LOC-see-PAST-DS

     chopa-tok] ?
     buy-PAST

     'Where [Joyce bought the dog [John saw __]] ? '

If the IHRCs have a structure such as that in (4), wh-movement of elements from within the RC is predicted to be ungrammatical by subjacency.  If, on the other hand, the S-structure category of an IHRC is S-bar, then no subjacency effects are predicted, since elements within the RC are dominated by a single bounding node.  The predictions of the LF movement analysis are correct.

One alternate explanation of the apparent subjacency violations is that Choctaw allows zero resumptive pronouns, so that the gaps in (5) and (6) are pro rather than wh-trace.  However, Choctaw obeys certain other island constraints which show that a zero resumptive pronoun strategy is not available.  Consider the following example:

(7a.) Ak-ithaano [S'Bonnie-t  ahi    honi-kma
      1sgNeg-know    Bonnie-NOM potato boil-IRR=SUB/DS

      nanatoka]
      whether

      'I don't know whether Bonnie boiled the potatoes'[7]

(b) ? * Natah-o ak-ithaano [$_S$'Bonnie-t ___ honi-kma
     what-ACC 1sgNeg-know   Bonnie-NOM   boil-IRR=SUB/DS

nanatoka]?
whether

'*What don't I know whether Bonnie boiled? '

(c) ? * Katah-oosh ak-ithaano [$_S$' ___ ahi honi-kma
      who-NOM   1sgNeg-know        potato boil-IRR=SUB/DS

nanatoka ]?
whether

'*Who don't I know whether boiled potatoes? '

The ungrammaticality of (7b) and (7c) is due to the Wh-Island
Condition, a sub-case of Subjacency (but see footnote 7). If
Choctaw had a zero-resumptive pronoun strategy, then presumably
it would apply in the examples above, and render them gramma-
tical. Since it does not, we may conclude that Choctaw does not
employ zero resumptive pronouns. Thus the movement facts discus-
sed above constitute evidence for the LF movement analysis. This
analysis provides an intuitively satisfying explanation of the
relation between extraction from relative clauses and the non-
occurence of overt heads in these same RCs.


## 2. Switch reference

Choctaw has a well developed system of switch reference
(SR), which indicates whether two clauses have the same or dif-
ferent subjects (SS or DS). Consider the following examples:

(8)   Kah sa-nna-hatokoosh, iskali ittahobbi-li-tok.
      car 1sgII-want-because/SS, money save-1sgI-PAST

      'Because I wanted a car, I saved money.'

(9)   Kah banna-hatoko, iskali ittahobbi-li-tok.
      car want-because/DS, money save-1sgI-PAST

      'Because he wanted a car, I saved money.'

The SR marker is suffixed to the complementizer of the lower
clause, and it shows whether the matrix subject is the same as
the embedded subject.

Finer (1984) presents an analysis of SR in terms of binding
theory. Briefly, he assumes that the SR marker originates in the
AGR of the embedded clause and is moved to the COMP position.
Since SR originates in AGR, it has the agreement features of the

verb. In particular, it is coindexed with the subject of the embedded clause. Finer proposes that SS is an anaphor and DS is a pronominal. Their governing category is then the matrix S. By standard assumptions, the subject of the matrix clause c-commands the SR marker in COMP. If the matrix subject agrees in features with the SR marker, then the matrix subject binds it. Note that the matrix object is not a potential binder for the SR marker since it is in VP, and thus fails to c-command it.[8] By the binding theory, anaphors must be bound in their governing category, while pronominals must be free in their governing category. Thus the SS switch reference marker must be bound by (coindexed with) the matrix subject. On the other hand, the DS switch reference marker cannot be bound by the matrix subject; therefore the two clauses must have different subjects. For detailed discussion and justification of these proposals see Finer (1984).

Now let us reexamine (2), repeated as (10) below.

(10)  Joyce-at [John-at ofi aaipa nota-ma    aa-pisa-tok-ma]
      Joyce-NOM John-NOM dog table under-ACC LOC-see-PAST-DS

      chopa-tok.
      buy-PAST.


      'Joyce bought the dog John saw under the table.'
   or 'Joyce bought the table John saw the dog under.'


Contrast this with (11):


(11) Joyce-at [ofi aaipa nota-ma    aa-pisa-tok-mat]
     Joyce-NOM dog table under-ACC LOC-see-PAST-SS

     chopa-tok.
     buy-PAST.

     'Joyce$_i$ bought the dog that she$_i$ saw under the table.'
  or 'Joyce$_i$ bought the table that she$_i$ saw the dog under.'


These sentences are explicable under the binding theory account of switch reference without modification if the S-structure category of an IHRC is S-bar. Under this account, the SR marker in COMP is c-commanded by the matrix subject, and the governing category is the matrix S, just as in (8) and (9). The LF movement hypothesis would then claim that there is no syntactic difference between relative clauses and complement clauses at S-structure.

An analysis which posits an empty head for the IHRC will encounter several problems with these examples, however. If the relative clause is dominated by NP, then this will be the governing category for the anaphor in COMP[9]. Since the matrix subject

is outside the governing category, it cannot bind the SS anaphor. However, the empty NP head <u>will</u> c-command the anaphor, and serve as a local binder.  The empty head analysis will then predict that SR in relative clauses should mark whether the head is coreferent with the subject of the relative clause.  Yet this is not the case.  SR in relative clauses indicates the coreference between matrix subject and embedded subject.  A test case is (10) above.  If the head controlled SR, then we would expect DS switch reference, since 'dog' (or 'table') is non-coreferent with 'Joyce'.  In fact, it is the matrix subject that controls SR, so the RC receives SS marking because 'Joyce' is the subject of both clauses.

The empty head analysis makes incorrect predictions about SR in Choctaw, and is therefore less adequate than the LF movement analysis.  On the basis of this evidence and the subjacency effects discussed above, I conclude that the LF movement analysis is descriptively and explanatorily superior.

## 3. Further considerations

Above we dismissed the structure in (3) as ill-formed by X-bar theory.  This assumes that endocentricity is a sort of filter at S-structure.  Suppose, however, that the endocentricity filter applies at LF.  Then (3) will be an acceptable S-structure, so long as movement applies to create a head at LF.

Rizzi (1982:92) suggests that some principle like (11) exists:

11.) $\alpha$ is bounding only if branching.

Since the NP dominating the IHRC is non-branching, then by hypothesis it does not count as a bounding node.  Hence the subjacency effects noted would be accounted for.

What of the SR phenomena?  Note that if there is no head, then the NP cannot be the governing category for an anaphor in COMP because there is no accessible subject within NP.  The governing category is rather the matrix S, and the matrix subject will bind the SR marker in COMP, as needed.

An advantage, perhaps, of this approach is that it allows categorial selection to be satisfied at all levels of grammar. In this respect it is an interesting counterpart to the theory that IHRCs are S-bars at S-structure.  In the latter approach, endocentricity is satisfied at all levels, while categorial selection holds only at LF.  In the structure (3), however, categorial requirements are satisfied at all levels, while endocentricity holds only at LF.  It may well be the case that there is little or no empirical basis for distinguishing these hypotheses due to general properties of non-branching dominance (Lasnik and Kupin 1977).  It seems clear, however, that for Choctaw we cannot claim that both categorial requirements and

endocentricity hold at all levels of grammar.


NOTES


1.  Some other languages with this construction are Navajo, Hopi, Washo, Tonkawa, Pomo, the Yuman languages, Lakhota, and the other Muskogean languages.

2.  It may be the case that all RCs in Choctaw can be analysed as IHRCs, but the evidence is not conclusive.

3.  (3) is suggested by Gorbet (1976); (4) by Finer (1984).

4.  See Rizzi (1982) for discussion.

5.  There are several variations in the examples which do not affect the argument here.  Both -at and -ato are glossed NOM. Both katommah and katommako are glossed 'where'.  In each case the latter is slightly more emphatic.  Both -tok and -atok are glossed as PAST.  The latter is perhaps more common with adjectival predicates.

6.  Translations for these sentences present a problem, since there is no grammatical English translation which preserves the structure of the Choctaw. An approximate gloss of (5) is 'Where did John see the dog that was red?' (6) has the the approximate gloss 'Where did John see the dog that Joyce bought?' In these glosses, however, the main and subordinate clauses have been reversed, so that they do not indicate the syntax of the Choctaw examples.
     A speaker consulted by an anonymous reviewer for KWPL translated (5) as 'Where was it that the dog John saw there was red?' while (6) was translated as 'Where did Joyce buy the dog that John saw there?' That is, this speaker of Choctaw translated the sentences with the interrogation in the main clause. However, it is clear that the Choctaw sentence involves movement from the subordinate clause, since the locative aa- is prefixed to the subordinate verb.
     Given the ungrammaticality of English sentences parallel to these Choctaw examples, it is perhaps not surprising that speakers vary in their translations.

7.  The gloss for nanatoka is quite problematic.  It occurs after the verb + complementizer in the embedded complement of a small number of verbs, and seems roughly to function as an equivalent of English 'whether'.  It is morphologically nana-toka.  Nana occurs elsewhere as an independent word meaning 'something'; it is the indefinite form of the wh-question word natah 'what? '. Toka is elsewhere a complementizer.  Thus there seems to be a complex COMP of the form [$_{COMP}$ -kma [$_{NP}$ nana [$_{COMP}$ -toka]]].

Alternately, nana might be functioning as an external head for the RC. In this case, this would be the only example of of an externally-headed RC. Note that the position is where one would expect a head in a consistent SOV language. In this case the effects in (6a) might be due to the CNPC, rather than the WHIC. The occurence of an apparent complementizer on a noun is problematic for this analysis, however.

Note that either analysis of nanatoka will support the argument advanced in the body of this paper.

8. Note that this proposal relies crucially on the existence of VP in SOV languages, a position which is somewhat controversial (Saito 1985). I believe that the SR facts in Choctaw provide fairly strong evidence for VP in this language.

9. Assuming that the governing category for an anaphor is the minimal NP or S dominating it. Chomsky (1981) suggests that a governing category for an element a be defined as the minimal category containing a, a governor for a, and an accessiable subject. If we assume that the head of a RC serves as accessible subject for an anaphor in COMP, then the governing category is still the NP.

## ACKNOWLEDGEMENTS

## REFERENCES

Broadwell, George A. 1984. Some issues in Choctaw syntax. MS.

Chomsky, Noam. 1981. Lectures on government and binding. Dordrecht:Foris.

Cline, David M. 1985. Free relatives and LF categorial selection. Unpublished ms, UCLA.

24

Finer, Daniel L. 1984. The formal grammar of switch-reference. PhD Thesis, UMass Amherst.

Gorbet, Larry P. 1974. Relativization and complementation in Diegueño: noun phrases as nouns. PhD Thesis, UCSD.

Grimshaw, J. 1979. Complement selection and the lexicon. Linguistic Inquiry 10.2.

Haiman, John and Pamela Munro (eds). 1983. Switch reference and universal grammar. Amsterdam: John Benjamins.

Jackendoff, Ray. 1977. X-bar syntax: a study of phrase structure. Cambridge, Ma: MIT Press.

Jacobsen, William H. Jr. 1983. Typological and genetic notes on switch-reference systems in North American Indian languages. in Haiman and Munro (1983:151-184).

Lasnik, H. and J. Kupin. 1979. A restrictive theory of transformational grammar. Theoretical Linguistics 4.3.

Munro, Pamela (ed). 1980. Studies of switch-reference. UCLA Papers in Syntax, No. 8.

Munro, Pamela and Lynn Gordon. 1982. Syntactic relations in Western Muskogean: a typological perspective. Lg. 58:81-115.

Pesetsky, David M. 1982. Paths and categories. PhD Thesis, MIT.

Rizzi, Luigi. 1982. Issues in Italian syntax. Dordrecht: Foris.

Saito, Mamoru. 1985. Some asymmetries in Japanese and their theoretical implications. PhD Thesis, MIT.

Stowell, Timothy A. 1981. Origins of phrase structure. PhD Thesis, MIT.