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 Citics: The Word Boundary Problem
Marcia Haag ......................................................... 33

Paragraph-Level Switch-Reference Markers in Chickasaw Conversation
Cynthia Walker ...................................................... 45

Ojitlán Chinantec Phonology and Morphology
Monica Macaulay .................................................. 71

Glottalization in Nun-chah-nulth in Optimality Theory
Eun-Sook Kim ......................................................... 85

On the Relationship Between Mixe-Zoquean and Uto-Aztecan
Søren Wichmann ................................................... 101
SWITCH-REFERENCE AND THE IDENTIFICATION OF SMALL PRO IN HIDATSA

John P. Boyle
University of Chicago

Abstract: The purpose of this paper is to explore the Hidatsa switch-reference (SR) system and determine the relationship it has with the null category pro. I focus on two morphemes: uik (SS) and niak (DS) whose roles have been misidentified (Robhnert 1955, Jones 1984). I show that the morphemes are better analyzed in SR markers. This analysis reveals a system where the morphemes function as coordinating conjunctions that also contain the properties of a SR system. In this role, they act as lexical heads of coordinator phrases. Hidatsa often omits all agreement markers leaving only the predication and the SR markers. I argue these SR markers are sufficient to identify pro. This paper demonstrates an additional system (the SR system) that have not been examined before with respect to the role it plays in recovering information to identify null categories. This paper also shows that the SR morphemes are textual coordinations that serve to identify pro under government.

This paper has two sections. The first section is devoted to exploring and clarifying the switch-reference system that exists within the narrative speech style in Hidatsa. The second section attempts to determine exactly what relationship exists between the Switch-Reference tracking system and the empty category pro. Hidatsa is spoken by between 150 and 500 people living in or near the Ft. Berthold Indian Reservation in North Dakota. Hidatsa is a member of the Missouri River branch of the Siouan language family. Its closest relative is Crow in Southeastern Montana. Like the other members of the Siouan language family, Hidatsa is typologically a S-O-V language.

The Switch-Reference System in Hidatsa

Until recently the switch-reference system that exists in Hidatsa had not been correctly analyzed. Hidatsa, like many Native American languages, has been studied by very few people. Washington Matthews (1877) published a short grammar and dictionary; Robert Lowie (1939) published five Hidatsa texts with additional grammatical notes and a partial morphological

Kansas Working Papers in Linguistics 24-2, pp. 17-32
analysis provided by Z. Harris and C. F. Voegelin; Florence M. Rohanet published three articles in DIAL (1955) which listed various affixes and stems as well as a phonological analysis; Hubert Matthews, who wrote his dissertation on Hidatsa, compiled an early Transformational Grammar analysis of Hidatsa syntax (1965); since this time few articles dealing with Hidatsa have appeared. Wesley Jones did field work on Hidatsa in the 1980s but most of his material remains unpublished. The data that was used for this paper is taken from the Traditional Narrative Texts collected by Robert Lowie and published in 1930.

In his 1987 article on switch-reference in Crow, Randolph Gracey suggested that a system similar to that of Crow might exist in Hidatsa. I believe that such a system exists, and the syntactic analysis of this system is what this article will examine. Hidatsa actually has two switch-reference systems: one that is used in the conversational style of speech and one that is used in the traditional narrative style of speech. Since four of the five texts collected by Lowie are in the Narrative style this article will examine the morphemes used in that style of speech. The only difference in the two styles with regards to the switch-reference system is the form of the Different Subject morpheme (-wa in the conversational style and -ruk in the traditional narrative style). My analysis will work for the conversational style as well.

Switch-reference is a device for referential tracking where one of a set of morphemes is affixed (usually suffixed) onto a verb in order to indicate something about the identity of a noun phrase (Haiman and Mann 1983: ix). These markers help to track the identity of a subject from one clause to the immediately following clause. In Hidatsa, these morphemes are affixed onto the clause-final, but not matrix-final verb5. The switch-reference morphemes are -ak (same subject, henceforth SS) and -ruk (different subject, henceforth DS). Traditionally the -ak has been analyzed as a verb-final, but not sentence-final marker (Lowie 1939: 187). The -ruk was analyzed by Washington Matthews as “an adverb of future time, that is suffixed to subjoined verbs, to denote doubt or condition in regard to future time, and is therefore equivalent to a sign of the subjunctive mode in the future tense” (Matthews 1877: 105). With respect to the conditional use of -ruk, this is a fair analysis. Harris and Voegelin first noticed the importance of these morphemes with regards to their noun phrase tracking characteristics. In text 1, they provide a footnote which states:

5 -ruk verb-final used of actions occurring at the same time as the sentence-final verb, apparently when the actor of the verb is not identical with the actor of the sentence-final verb. Verb-final -ruk
is probably not a contraction of -ru and -ak, for aside from difficulties of phonology, verbs in -ak have identical rather than non identical actors in respect to sentence-final verbs” (Lowie 1939: 189).

Clearly Harris and Voegelin could see that these morphemes were keeping track of the actors with regards to the action, but they did not appear to realize how pervasive it was throughout the system. Switch-Reference had not yet been explored as a grammatical issue, and as a result it was not something for which they were looking. Throughout the footnotes of the texts, Harris and Voegelin occasionally mention that -ak is the marker for a different actor from the actor of the sentence-final verb. This error in assuming that the point of reference is the sentence-final verb rather than the verb in the following clause causes them to never see the larger discourse implications, these being a switch-reference system that tracks the actors in order to keep referential ambiguity to a minimum.

Florence Robinett (1955: 173) analyzes -ruk as -ra + -k, with -ru being a type of spatio-temporal locative and -k being a subordinator/conjunction. A. W. Jones concurs with this interpretation. Further, he states that the -k is a subordinator/conjunction in both the -ak and the -ruk morphemes (Jones 1984: 8-10). Robinett and Jones both analyze -ra as a locative wherever it occurs. The locative morpheme -ru (on, in, at) clearly exists in Hidatsa:

(1) məʔis ʔaːkaa-ru on top of the rock
the rock its top-at
məʔis ʔiʃẹki-ru in front of the rock
the rock its front-at
məʔis ʔamii-ru beside the rock
the rock its side-at (Jones 1984: 12)

I believe that this morpheme is different than the -ruk DS morpheme. In every instance in Jones (1984) involving the -ruk morpheme, there is a change in actor within the sentence. All of his examples use the narrative style of speech. It is only when the -ru is affixed without the -k that the locative sense is clearly evident. These examples are in the conversational style of speech. I believe that two separate morphemes -ru and -ruk exist in Hidatsa. As stated above, there are two styles of speech present in Hidatsa, the narrative and the conversational. In the narrative style
-rak is a switch-reference marker, which indicates a different subject with regards to two different predicated. In the conversational style, the -(r)ak morpheme serves as a temporal/conditional marker and the different subject marker for the switch-reference system -wa. Since until recently, the switch-reference system that exists in Hitada has not been discovered, the different roles that the morpheme -rak plays within the language had not been completely explored. This confusion also exists in Crow between the switch-reference DS marker -dak and the temporal/conditional marker -dak. In Crow -dak can serve as a temporal/conditional marker in both the narrative style and the conversational style of speech. When -dak is used as a temporal/conditional marker of subordinate clauses in the narrative style of speech it is outside the switch-reference system (Gracey, personal communication). This homophony seems to also occur in Hitada.

Robinett and Jones' interpretation does account for the subordinator/conjunction characteristics of -k. My analysis however, is that within the switch-reference system, k is not the -k alone that acts in this manner, but both of the entire switch-reference morphemes. Washington Matthews states that the word ba is the word used to conjoin two sentences. However, he adds that other words are used in joining words and sentences (Matthews 1877: 121). In the Lowie texts, ba is always translated as 'again', and never 'and'. In fact the Hitada equivalent of 'and' does not seem to be manifest in any morpheme, unless the -k serves this purpose. There is no separate word, which serves this purpose. In the Lowie texts, the translation 'and' is given almost every time either the -sk (SS) or -ruk (DS) morphemes are used. Jones (1984) argued that the coordinator aspect of these morphemes is -k, that the -(r)u- is a locative, and that the -a- is phonemic. This coordinator aspect of -k can be seen in the following Hitada example:

(2) mléékaatí kiyawasaka hitáa-
buffaloes and horses are fast

Here, the -k is clearly used as a conjunction. I will show that it is best to analyze the -ak and -ruk morphemes in the narrative speech style as coordinating conjunctions that additionally contain the properties of a switch-reference system. In the past an over analysis of the morphemes occurred due to not understanding the differences that exist between the two speech styles. As is often the case in switch-reference systems, the SS/DS morpheme additionally acts as a conjunction (Harman 1983: 107). Given this, separating the -k from the -ak would miss
an important generalization about the properties of the same subject marker. To explain this 
marker as a phonemic variable would destroy the switch-reference system.

This conjunction aspect of switch-reference markers can be seen in Mundan, a Siouan 
language distantly related to Hidatsu. The switch-reference morphemes in Mundan had also been 
analyzed as coordinating or aspectual/subordinating suffixes (Kennard 1936; Hollow 1970; 
Carter 1991). Mixco (1997) has shown convincingly that the two morphemes, -ni and its 
phonetic variation -nj (SS), and -ak (DS) are indeed switch-reference markers. In Mundan, 
another function the SS morpheme possesses is as a coordinating conjunction. In this Mundan 
example the same subject morpheme -ni is used as a coordinator between noun phrases:

(3) kach'/ka-ni macku/chapka-ni' minisa? 
chokecherries and bull berries and willows (Kennard 1936: 24)

This unification of functions is not limited to Hidatsu and Mundan within the Siouan 
language family. Graczyk (1987) states that the morpheme -dak serves as a coordinate 
conjunction for noun phrases:

(4) Peter-sh-dak John-nak Mary-sh-dak Ammalapâśhiku countless-dûn-ak 
Peter-DEF-and John-and Mary-DEF-and Billings to-go.PL-DELC2 
Peter, John, and Mary went to Billings.

This -dak, however, behaves differently phonologically than the switch-reference -dak and the 
conditional -dak, in that it may never be accented, while the others can bear accent. 
Synchronically this would appear to be a different morpheme (Graczyk personal 
communication). However, Haiman (1983) argues that not only is it a common function of 
switch-reference markers to be conjunctions, but often that is their ultimate origin as well.

The Syntactic Analysis

As can be seen from the above data, the morphemes in the Mundan and Crow systems 
have the additional ability to conjoin noun phrases and are glossed in English as 'and'. I will 
analyze the switch-reference markers in Hidatsu in the same manner. These switch-reference 
markers not only act to track agent/subject they also serve as clause coordinators. In this role, 
they act as lexical heads of coordinator phrases.
Consider the following data, which exemplifies some interesting properties of switch-reference examples 5-8):

5. waŋ-šiŋ waː-te-ɾük ʃaːk-haːwiɾiɾh aʃki rü-te-kho-ak tʃu-haːɾ-ak tʃu-te-waɾe-dək
   5. Day-Sun already-go-DS First-Worker (him)self already-know-SS lift-up-NE
   5. After Day-Sun had gone First Worker, who now know how it was done, stood up. *
   5. When the Day-Sun had already gone, First Worker himself, knowing now how it’s done, stood up (they say). (1-10)

6. ʃaːk-kaɾ-ak tan-ʃe-ak ak aʃk-kaɾ-ak wiɾa-waː greeting ak aŋk-te-ɾəp ak rə-ɾe-ʃe-dək
   6. SC go-SS wing-by holding-SS bring-SS woods-inside carrying-in-have-SS go-NE
   6. Then holding him (goose chief) by the wing, he (First Worker) carried him inside the woods.
   6. Then going, holding him (the goose chief) by the wing, bringing him over, carrying him in the woods, he went (they say). (2-5)

7. ʃaːk-raɾiɾ-awhaɾiɾ-ɾək ʃaɾ-te-waɾe-dək
   7. Then when First Worker twitched, Spotted Tail ran away again.
   7. Then when he [First Worker] gave a start, Spotted Tail, running away, went (they say). (3-34)

8. aʃk-kiɾiɾ-ɾək kəɾ-te-waɾe-dək
   8. SC jump-DS run-SS go-NE
   8. Then First Worker jumped, and Spotted Tail ran away.
   8. Then when he [First Worker] jumped at him, he [Spotted Tail], running away, went (they say). (2-37)

In the literature, it is proposed that a language that employs a pro-drop system has to have some way to recover the information that is lost by not having an overt NP. Originally, this was thought to be done through a rich INFL. The grammatical features of the subject could be recovered from the features of INFL, specifically from ARG, in languages with a complex verbal inflection system (Rizzi 1982: 119). Hidatsa has person and number agreement but in situations like those in the above examples part of or all of this agreement is lost. This happens with
particular frequency in clauses with 3rd person references. In Hidatsa, third person pronoun agreement is optional and these pronouns are often dropped in speech. This means that Hidatsa does not seem to provide enough agreement morphology to recover the information that is required for pro. Huang (1984) offers Chinese as a counter example to the rich NFL hypothesis. Chinese, Huang argues, has no AGR specification. He claims that the basic topological generalization concerning pro in subject position is that it is possible only in languages with strong AGR (i.e., languages with complex verbal morphology) or in languages which have no AGR at all. The former he labels as “medium” languages and the latter he labels as “cool” languages. He argues that only in these types of languages it is possible to have pro (Huang 1984: 557). This seems also to be counter to the above data. Hidatsa has agreement morphemes. However, there is not always verb-subject agreement. Hidatsa, like many Native American languages, have a series of pronominal affixes. Although I would argue that the first and second pronominal prefixes are the syntactic arguments rather than agreement markers (as per Jelinek 1984), for the sake of simplicity I will refer to the pronominal affixes as agreement markers. The reason for this is that in the above examples, we are really concerned with the phonologically null third person. I will assume (as per Graczyk 1991: 88) that in Hidatsa, these zero markers are vacuous agreement markers. In addition, also following Graczyk 1991, I will assume that lexical noun phrases in Hidatsa function as genuine syntactic arguments rather than adjuncts or appositives referential with the zero pronominal affix. This agreement in the third person markers would seem to make Hidatsa a “hot” language, which is the same category into which English falls.

In Hidatsa, we seem to have a language that has the NFL system that is needed to be able to be a pro-drop language, however Hidatsa also drops these agreement markers that are so important to analyzing the information that the language is trying to convey. How does a speaker of Hidatsa then understand the reference that goes with each predication? I believe that while this can be done with gender, number, and/or person agreement morphemes, Hidatsa can also employ only the switch-reference marker morphemes. These markers are in a unique position to govern the clause that follows them. This position, which we will examine below, might seem unnatural to some. Hidatsa is a head-marking and generally head-final language. While many languages that are structurally similar to Hidatsa are head-final with regards to all their constituents, we will see below that Hidatsa is not this judged. It must be kept in mind that X’ Theory does not predict directionality of any given language. It is well documented that many languages show mixed directionality (cf. V-final languages like German and Dutch are head-initial for Ns and Cs, and even in strongly head-final languages like Turkish, we find
The definition of government, m-command, and the ECP are assumed based on Chomsky's Barriers (1986)\textsuperscript{41}. It is my claim that Coor antecedent-governs pro in SPEC of IP. Although if we follow Chomsky (1986) carefully this is not true. Chomsky claims that IP is a Blocking Category (BC) for pro (because IP is not L-marked and it dominated pro). Therefore CoorP is a barrier for pro since CoorP immediately dominates IP (where IP is a BC for pro). However, this is a technical problem. When Chomsky stated this he wasn't considering coordinate syntax, and certainly not the X' respecting variety that I'm proposing here. In fact, it is clear that the presence of Coor and its projections shouldn't count as BCs at all, given examples like: 'I want [her to be arrested] and [him to be set free]'\textsuperscript{7}. In this example, if Case assignment happens under government, we must be able to ignore the CoorP projections. In addition, we could also redefine barrierhood to exempt CoorP and although this is a more brute force method it is also an answer to our problem. Given these arguments, I will now assume that there is no problem with antecedent-government. It should also be remembered that the SR markers are not true antecedents. I am not claiming that that pro *must* be identified by a SR marker, only that it *can* be. This is crucial with regards to the DS marker. The DS marker has the semantics of "the subject of the following clause is different from the subject of the preceding clause". This isn't actually enough to indicate the reference of the second subject (it only eliminates one alternative even though they have been co-indexed with the preceding subject for convenience). Finer (1984, 1985) has suggested that the same subject SR marker can be analyzed as an A'-anaphor and that the different subject SR marker can be analyzed as an A-pronominal. In addition, we can coin the SR markers as we would pronouns and anaphoric elements. This should allow the speaker to extract enough information to identify pro.

Now consider the four example sentences (5 - 8) again in more detail (I repeat here example 5):

5. [wa.pi-wiri3 water-re-ruk] [reikha:wa:bire iheki] water - e:heke - o:he] [pro in:hi:wa:re]

5. Day-Sun already-go-DS First-Worker (him.self already-know-SS lift up-NE

5. After Day-Sun had gone First Worker, who now knew how it was done, stood up.

5. When the Day-Sun had already gone, First Worker himself, knowing now [how it's done], stood up (they say).
In example 5, we see that the first two coordinating clauses have overt subjects. The switch-reference markers signal the subject of each of the clauses. In the above structure, we can see that the SS/DS morpheme antecedent-governs the following Coordinate Phrase or Infection Phrase, assuming that IP is not a barrier to government, and that the Coor P, a conjunct of two IPs, shares that same lack of barrihood. While it may seem that Ross’ 1967 Coordinate Structure Constraint (CSC) implies that each member of a coordinate structure is a barrier to government, in fact does not. If this were the case, it would rule out ATB movement in such examples as “Who did you say [[Bill invited t] and [Sue talked to t]]?” In point of fact, Chomsky 1986 does not try to make the CSC follow from the barriers system, and for the most part later investigators (e.g. Mann and Johannessen) have followed in this, analyzing the CSC as essentially semantic.

According to the Extended Projection Principle, SPEC of IP must be filled. In the above example we see that all but the most embedded SPEC of IP is filled with a lexically null NP. The most embedded IP is like all of the IPs below in the following example (6):
In this sentence there are five coordinating clauses with no overt agent performing the action. The reference is assigned from the discourse. In this sentence from Text II, First-Worker, the Hidatsa trickster, has the goose chief and is carrying him into the woods to eat. In these coordinating clauses we can see that although the SS morpheme is bound to a clause final but not matrix final verb it actually governs the phrase which follows it. The switch-reference
morpheme tells the listener what the subject of the next phrase will be. We can see that the SS morpheme governs pro in the subject of the following Inflection Phrase, given that Coor P and IP are not barriers.

This analysis holds true for the DS morpheme as well (seen in example 7).

7. [harlik pro ruč ih-awahke-ruk] [cįstupšiši karé-(a)k] [pro rē-wuse]]
7. SC twitch-there-DS Spotted Tail run-SS go/NE
7. Then when First Worker switched, Spotted Tail ran away again.
7. Then when he [First Worker] gave a start, Spotted Tail, running away, went (they say).

(7)

```
CP
  
  Adv P
  
  harlik
  
  Coor P
  
  I
  
  -ware
  
  pro
  
  Coor
  
  Coor P
  
  IP
  
  ruk
  
  I
  
  -(a)k
  
  cįstupšiši
  
  VP I
  
  karé
  
  VP I
  
  re:
```

Here we can see that pro plays the same role for the DS switch-reference marker as it does for the SS switch-reference marker. All of the grammatical relationships function the same way. The only difference is the subject reference. In this sentence First Worker is sleeping. He is the actor and subject of the first clause. Spotted Tail sneaks up on First Worker, but First Worker switches in his sleep. This causes Spotted Tail to panic and run away. The DS morpheme -ruk signals the listener that the subject in the upcoming clause is about to change and will not be the same as the previous clause. This is the listener's only grammatical cue that the subject of both clauses is not First Worker, which could be assumed from discourse, or Spotted Tail, which is the only overt NP in the sentence.
Given this information we can see that PRO is not reasonable to postulate in the SPEC of IP since PRO must be ungoverned. The empty category also can not be trace, since it has no antecedent in an A position. In Hidatsa, I is filled with [+ten, +arg]. This enables INFL to govern SPEC of IP. Therefore, it is evident that this position is licensed and filled with pro. Further more, SPEC of IP receives a theta-role from the predicate in IP. In Hidatsa, empty subjects can function as null arguments. This is clearly shown in example 8:

8. [hunuk [pro akčixi-ruk] [pro karč-(a)k] [pro rē-wāre]]
8. SC jump-DS ran-SS go-NE
8. Then First Worker jumped, and Spotted Tail ran away.
8. Then when he [First Worker] jumped at him, he [Spotted Tail], running away, went (they say).

(8)

Here we can see that there is no overt NP and pro fills the SPEC of IP position in both clauses. In addition, and more importantly, there is no agreement morphology on the verbs other than the switch-reference markers. Although I is filled with [+ten, +arg], it is not phonologically overt. This makes Hidatsa a "hot" language like English. In the above sentence the speaker has dropped all other third person references. The only way a Hidatsa speaker can recover the necessary data to determine the reference of the predication is by following the discourse and employing the switch-reference markers. The speaker uses the switch-reference markers to identify pro.
In conclusion, the preceding data shows another way in which information can be retrieved in a pro-drop language. Less common morphology can play an important role in recovering information for empty categories. This paper shows that in Hidatsa, switch-reference morphemes are lexical coordinators that serve to identify pro under government. Switch-reference morphemes can function in the same manner as gender, number, and person with respect to pro. This is a totally unexplored area. This new analysis adds to our knowledge about morphology that can identify pro and it provides new insights for both Universal Grammar and language typology.

NOTES

1 I would like to thank the support I have received from the faculty at the University of Utah. I would especially like to thank Marko Nixas who inspired me to work with Native American Languages. I would also like to thank Edward Rubin who inspired me to work with theoretical syntax. I owe them both a great deal. In addition, I want to thank Randy Gooseck whose suggestions have been both insightful and thought provoking, and Susan Merchant who helped me work out most of the final links. Any mistakes or errors are my own.

2 Since Hidatsa is an SOV language the matrix verb comes at the end of the sentence. These matrix-final verbs do not take a switch-reference marker. The switch-reference markers attach only to the clause-final verbs and indicate whether the subject of the clause that follows it is identical (SS) to or different (DS) from it.

3 Kemmer did not recognize nasalization in his 1936 grammar.

4 DELC = declarative, DEF = definite article, DS = different subject, NE = narrative ending, PL = plural, SC = sentence connective, SS = same subject.

5 This analysis uses a form of grammatical theory from the 1980's known as the Principles and Parameters approach proposed by Chomsky in 1981. This was used because of its focus on government, a grammatical relation. The relationship that the SR morphemes have to their following clauses is the primary relationship that this portion of the paper will explore.

6 These lines are unrelated to each other and do not form a continuing narrative. They are all taken from different parts of the Lowie texts.

7 The Hidatsa data used in this paper is taken directly from the Lowie texts. Any inconsistencies in vowel or consonant length as well as stress placement are reflective of those that exist in the original texts. Although this is a phonological problem I have made four changes to the original text. Hidatsa, like Crow, has a reduction rule where a sequence of three identical vowels is reduced to two. I have made four changes to the original text which make my analysis clearer. In sentence 1 the word e:kac:ta:k "to know-SS" was changed from e:kac:ta:k; in sentence 2 the word km:ca:k was changed from km:ca:k; in sentences 3 and 4 the word km:jic:ta:k "run-SS" was changed from km:jic:ta:k. No further attempts have been made to reconcile or further address any other differences here.
1 The final translation for each line is that provided by Harris and Voegelin. I found there to be rather subtle. The one above is my own and I believe that this translation captures the events better.

2 This is a reference number for the text and line where the example was found. The first number corresponds with the Text number and the second number corresponds to the line number in the text.

3 This word employs the Vocation-iu. Its use is clearly different from the clause-final switch-reference marker. In addition we can not use the verb to have in a special compound formation with another verb which forms a compound verb (Lowe 1939: 205). For this reason this compound is treated as a single verb in the analysis presented here.

4 Government: A governs B iff A is a governor, and A is a-contrastive B, and no barrier intervenes between A and B, where maximal projections are barriers (Chomsky 1986: 88); v-contrastive A v-commands B iff A does not dominate B and every maximal projection that dominates A also dominates B (Chomsky 1986: 88). The ECU: A theta governs B iff A governs B, and A theta-marks B. A antecedent governs B iff A governs B and A is coindexed with B (Chomsky 1986: 17).

REFERENCES


