10

Prosodic Conditions
on Anaphora and Clitics in Jakaltek

Judith Aissen

In studies which reach back more than twenty years, Colette Craig described certain disjoint reference effects in Jakaltek, a Mayan language spoken in the highlands of Guatemala (Craig 1977), and one of the rigid VSO languages in the family. These effects are induced by the presence of an overt pronominal element, as in (1), where the two instances of *ix* can only refer to different women (the genitive follows the head in Jakaltek).[1,2]

(1) Xkolwa *ix*   yinh [smi'   *ix*].
     helped NCL(she)   P   her.mother   NCL(her)
     'She$_i$ helped her$_{j/*i}$ mother.' [158]

The co-referential interpretation is expressed by omitting the second instance of the pronoun (this omission is represented throughout by Ø, glossed PRO). In fact, this is the only interpretation of (2).

(2) Xkolwa *ix*   yinh [smi'   Ø].
     helped NCL(she)   P   her.mother   PRO
     'She$_i$ helped her$_{i/*j}$ mother.' [158]

Craig considered at length the conditions which determine the distribution of the overt pronominal and Ø, and established that the two elements are in something approaching complementary distribution. With respect to examples like (1) and (2), Craig's key observation was that the covert form of the pronoun is required when it is *preceded* in some local domain by an antecedent. Conversely, the overt form of the pronoun is required when it is not. As a consequence, the first pronoun in (2) cannot be omitted.

In line with early transformational approaches, Craig conceived of the alternation between the overt and covert pronouns in terms of deletion. For her, the Ø form arose only at a late point in the derivation, as a consequence of deletion. Deletion was possible only if the pronoun was preceded by its

185

antecedent within some local domain, a domain which Craig characterized in terms of syntactic boundaries. The facts were reinterpreted by Hoekstra (1989) within the Binding theory. For Hoekstra, both the overt and covert forms were present throughout the derivation, but they were subject to different clauses of the Binding theory. Essentially, the overt form was analyzed as a pronoun, hence subject to condition B, while the covert form was analyzed as an anaphor, hence subject to condition A. Hoekstra argued more generally that the level at which the binding conditions hold should be parameterized, with surface structure the relevant level in Jakaltek. This conclusion was forced on him by his framework, which included the assumption that precedence is relevant only at superficial levels of representation. Both Craig and Hoekstra, then, attributed the fact that precedence was the critical relation between the pronoun and its antecedent, (rather than, say, c-command) to the fact that the alternation between Ø and the overt pronominals was determined at a superficial level of syntactic representation. It is true that if (surface) precedence is the relevant relation, then the conditions must hold at a very superficial level. But the fact that c-command apparently plays no role remains unexplained.

In the present chapter, I suggest that precedence plays the key role in licensing the Ø pronoun because the relevant conditions refer to prosodic structure, not syntactic structure. In prosodic structure, precedence is a meaningful relation between elements, while c-command is not. Support for the prosodic account will come from the fact that the domain in which Ø must find its antecedent likewise corresponds to a prosodic constituent, the intonational phrase. If this analysis is correct, it bears on the nature of the connection between phonology and syntax/semantics. The licensing conditions on Ø require access to prosodic structure; but at the same time, Ø is an anaphor, and as such, its presence restricts both the form and interpretation of the sentence in which it occurs. The licensing of Ø therefore appears to require some communication between the phonology and the level at which anaphoric relations are represented. For a conception of grammar like the Minimalist Program in which PF and LF are isolated (Chomsky 1995), such facts are problematic.

This work is based on the material and analyses of Jakaltek presented in Day (1973a), Datz (1980), and especially Craig (1977), which contains most of the analysis and data on which this chapter draws. It also draws on previous analyses of Jakaltek pronouns by Foley and Van Valin (1984), Hoekstra (1989), and Trechsel (1995).

1. Jakaltek Pronouns

Unlike most Mayan languages, Jakaltek has overt pronominal elements. These are drawn from a set of "noun classifiers which classify concrete objects and spiritual entities in twenty-four classes" (Craig 1986a:245); see also Day (1973b) and Craig (1977, 1986a, 1986b). Common classifiers include *naj*, used to classify male non-kin, *ix* for female non-kin, *ya'* for respected human, *ho'* for siblings, *no'* for animals, and *ixim* for corn and corn products.

The noun classifiers precede the nouns they classify, but they also occur self-standing with pronominal function[3]:

(3) a. naj Pel   'Peter'
    b. naj winaj   'the man'
    c. naj   'he'

Not all nouns are associated with classifiers. In particular, nouns denoting abstract and other non-concrete elements are not. Elements not associated with a classifier have no overt pronominal form, and pronominal reference back to such elements always involves a null pronoun. Following Wallace (1992), I assume that the classifiers belong to the category Determiner (D), where D may select an NP complement (Abney 1987). This makes sense of their capacity to both occur with a nominal complement and without. The other property of the pronominal classifiers which is relevant here is that they can be omitted. That is, reference back to an established discourse referent is sometimes accomplished with the overt classifier, and sometimes with Ø.

The following excerpt from Datz (1980:332-3) illustrates some of these properties. The protagonist (Ramon) is named at the start, in topic position. Subsequent references are either by the noun classifier *naj*, used for non-kin males, or by the null element Ø.

(4)   Naj Ramon, kaw xtxumtxun   sk'ul   *naj*   tzet
      NCL Ramon, very thought   his.stomach   NCL$_{(his)}$ how

      chu   yelkanh *naj;*   xtxumniloj   *naj*   yijb'antoj
      could   flee   NCL$_{(he)}$he.thought   NCL$_{(he)}$ he.accompany

      Ø   heb'   ix   yeskinahil   skayehal   yatut   Ø,
      PRO   PL   NCL$_{(them)}$ its.corner   its.street   his.house   PRO

      mach xin   yohtajoj   heb'   ix   yorona   tzetet lanhan
      not   they.know   PL   NCL 'llorona'   what   was

      xtxumtxun   sk'ul   *naj*   yinh   Ø...
      thinking   his.stomach   NCL$_{(his)}$ in.him   PRO
      'Ramon, he wondered how he could escape; he decided to
      accompany them to the corner of the street of his house
      without the "lloronas" knowing what he was thinking. . .'

2. Problems with a Binding Account

Examples like (1)–(2), repeated below as (5a,b), illustrate the fact that the overt pronominal classifier is used when a pronoun has no local antecedent, and that Ø

is used when it does. (In the examples which follow, italicized expressions are to be understood as co-indexed.)

(5) a. *Xkolwa   *ix*   yinh   [smi'   *ix*].
    helped   NCL$_{(she)}$ P   her.mother   NCL$_{(her)}$
    'She$_i$ helped her*$_i$ mother.' [158]

   b. Xkolwa$_i$ *ix*   yinh   [smi'   Ø].
    helped   NCL$_{(she)}$ P   her.mother   PRO
    'She$_i$ helped her$_i$/*$_j$ mother.' [158]

Following Reinhart (1983), and more specifically Trechsel (1995), I will assume here that the disjoint reference associated with the overt pronominal classifier is a pragmatic inference which follows from the speaker's failure to use the Ø form where s/he could have. I focus then on the licensing conditions associated with Ø. With Hoekstra (1989), we could say that Ø is an anaphor, subject to (some version of) condition A. Assuming that the domain for the binding conditions is at least the clause, (5b) is grammatical because Ø is bound; ((5a) is ungrammatical with co-reference because Ø could have been used, and was not). However, while Ø is an anaphor, in the sense that it must be anteceded locally, its properties do not follow from condition A.

The first problem is that the notion of "bound" and "free" which is relevant here does not involve c-command, but precedence; that is, the overt form cannot be *preceded* by a co-indexed element within the relevant domain. That precedence, rather than c-command, is the operative notion is suggested by examples like (6)–(8), in which the second pronoun is not c-commanded by the preceding co-indexed element, yet Ø must be used if co-reference is intended. In (6), the antecedent is a genitive within the subject; examples (7)–(8) involve preposing of a focused prepositional phrase,[4] with the antecedent a genitive within the object of that preposition:

(6)   Xkolwa [yunin   *ix*]   yinh   [s-mi'   Ø/*ix].
    helped her.child NCL$_{(her)}$ P   her.mother   PRO /NCL$_{(her)}$
    'Her$_i$ child helped her$_i$ mother.' [161]

(7)   [Boj smam   *naj*]   xtoyi Ø/*naj.
    with his.father NCL$_{(his)}$ went   PRO /NCL$_{(he)}$
    'It's with his$_i$ father that he$_i$ went.' [163]

(8)   [Sat [stx'at   *naj*]]   xhwayi Ø/*naj.
    on   his.bed NCL$_{(his)}$ sleeps   PRO /NCL$_{(he)}$
    'It's on his$_i$ bed that he$_i$ sleeps.' [161]

However, since genitives are known to take scope over the structure
c-commanded by the NP or PP in which they are contained (Reinhart 1987),
examples like (6)–(8) are not necessarily compelling. These might still be
amenable to an account of Ø as bound variable anaphora. But there are examples
in which the antecedent is much more deeply embedded, e.g., within a relative
clause on a preposed object[5]:

(9)   [Ixim ixim k'ochb'il yu *ix*]   xitoj   Ø   yinh molino.
      NCL corn   shelled   by NCL$_{(her)}$ she.took PRO   to   mill
      'It is the corn that she$_i$ had shelled that she$_i$ took to the mill.' [167]

In the face of this apparent indifference to hierarchical structure, there are two
possible conclusions. One is that the structural relation between Ø and its binder
is characterized in syntactic terms, but the relevant notion is precedence rather
than c-command; the other is that this relation is not characterized in syntactic
terms. Hoekstra (1989) takes the first tack. Because his discussion is embedded
in a framework in which precedence is a property only of superficial levels of
syntactic structure, Hoekstra concludes that the Binding theory is parameterized
as to the level at which it holds, and that in Jakaltek, it holds at S-structure.[6]

As support for this interpretation of the facts, Hoekstra (1989) observes that
reconstruction is irrelevant to the distribution of the pronominal classifiers and
Ø (see also Craig 1977:163). As illustration, consider (10) which shows that if
co-reference is intended, the first instance of the pronoun must be overt and the
second Ø. Thus, of the four possible outcomes, only one is grammatical, that in
which the first pronoun is overt and second is covert:[7]

(10) Yinh molino xitoj   *ix/*Ø*   [ixim k'ochb'il yu *Ø/*ix*].
      to   mill   2s.took NCL$_{(she)}$/PRO corn shelled   by PRO/NCL$_{(her)}$
      'To the mill she$_i$ took the corn that she$_i$ had shelled.' [167]

However, if the object is fronted for focus, the pronoun that must be covert in
(10) comes to be first, and must now be overt; the pronoun that must be overt in
(10) comes to be second and may now be covert, as in (9). If the conditions
determining the distribution of Ø could be satisfied through reconstruction, the
version in (11), where Ø is followed by *ix,* would be grammatical.

(11) *[Ixim ixim k'ochb'il yu Ø]   xitoj   *ix*   yinh molino.
       NCL corn   shelled   by PRO she.took NCL$_{(she)}$ to   mill
      *'It is the corn that she$_i$ had shelled that she$_i$ took to the mill.'

But (11) is impossible. The fact that it is provides further corroboration that
what is relevant here is surface precedence, and shows that any Binding theory
account must be parameterized to a very superficial level of structure, per
Hoekstra. So far, then, it seems possible to sustain a syntactic account of the
distribution of Ø, as long as reference to precedence is permitted. However,

when we turn to the question of the *domain* within which Ø is licensed, the syntactic account becomes less plausible.

Within the typology of Reuland and Koster (1991), Ø would not qualify as a short-range anaphor. Unlike the reflexive anaphor in Jakaltek, which is strictly clause-bounded (and must be anteceded by the subject) (Craig 1977:217, 362), Ø can be bound from outside the minimal clause, and across an intervening subject[8]:

(12)   Xil   *ix*   [$_{VP}$ hawilni   Ø].
   she.saw NCL$_{(she)}$ you.look PRO
   'She$_i$ saw you looking at her$_i$.' [168]

(13)   Xil   *ix*   [hawatx'en   skamixh   Ø].
   she.saw NCL$_{(she)}$ you.make   her.blouse PRO
   'She$_i$ saw you make her$_i$ blouse.' [169]

However, it might qualify as a medium-range anaphor. Reuland and Koster (1991) propose that medium-range anaphors satisfy two conditions: First, the domain within which they are bound coincides with the minimal domain of tense (or inflection) containing the anaphor; and second, the binder is a subject. However, neither of these properties holds for Jakaltek Ø. With respect to domain, there is an interesting range of cases in which it looks likes the domain within which Ø is bound does coincide with tense (cf. Foley and Van Valin 1984, esp. p. 298, table 12). The examples in (12)–(13), which allow binding across a subject, involve tenseless complements, while CP complements and adjuncts, which are inflected for tense, are barriers to binding of Ø:

(14)   Chal   *naj*   [$_{CP}$ chub'il chuluj   *naj/*Ø*].
   he.says NCL$_{(he)}$ that   will.come   NCL$_{(he)}$/PRO
   'He$_i$ says that he$_i$ will come.' [172]

(15)   Chin=tzotel tet *ix*   an [yunhe sta'wi *ix*   wet   an].
   I.talk   to NCL$_{(her)}$ 1S so.that answers NCL$_{(she)}$ to.me 1S
   'I will talk to her$_i$ so that she$_i$ will answer me.' [280]

But setting the binding domain for Ø as the minimal tensed domain faces a serious obstacle: That Ø within a tensed relative clause can be bound from outside the relative clause:

(16)   Xa'   *ix*   hune' kamixhe$_i$ [stz'isa   Ø   t$_i$]
   she.gave NCL$_{(she)}$ a   shirt   she.sewed PRO   t$_i$

   tet   snoh   Ø.
   to   her.brother PRO
   'She$_j$ gave a shirt that she$_j$ had sewed to her$_j$ brother.' [165]

(17)   Mat yohtajoj   *ix*   naj_i   [xmaqni   t_i Ø   yul parke].
       not she.knows NCL_(she) NCL_(he) WH.hit   t_i PRO in   park.
       'She_j doesn't know the man that hit her_j in the park.' [165]

(18)   Xchiwa   sk'ul   *ya'*   yinh ni'an unin_i
       was.angry her.heart NCL_(her) at   little child

       [xpohnitoj   t_i   [sxih   Ø]].
       broke   t_i   her.pot PRO
       'She_j was angry at the child who broke her_j pot.'   [166]

Examples like these indicate that the minimal domain of tense is not the correct characterization of the domain within which Ø must be bound.

Nor is it the case that the antecedent for Ø must be a subject. A number of earlier examples show this for cases in which Ø and its antecedent are clause-mates (e.g., 6–9). But even when they are not clause-mates, the antecedent need not be subject. Example (18), where the antecedent is grammatically genitive, is one case; example (19) is another:

(19)   Xkin=b'ey   wila' [yatut_i   *naj*
       I.went   I.see   his.house NCL_(his)

       [swatx'e   [sk'ahol Ø] t_i]].
       he.made   his.son   PRO
       'I went to see his_j house that his_j son made.' [166]

In sum, the relation between Ø and its antecedent is constrained by none of the properties which typically constrain the relation between a medium-range anaphor and its binder—the antecedent need not c-command Ø, the antecedent need not be a subject, and the two may be members of separate tensed domains. Taken together, these facts suggest that a binding condition which is defined over syntactic structure is not correct. In the following section, I argue that there is a straightforward *prosodic* characterization of the domain in which Ø must be bound.

3.   The Intonational Phrase as 'Binding' Domain

The basic proposal is that Ø must find an antecedent within the intonational phrase that contains it.[9]

(20)   CONDITION ON Ø:
       The anaphor Ø must be co-indexed with a nominal which precedes it within the same intonational phrase.

In Aissen (1992), an algorithm for deriving intonational phrasing from phrase structure in Jakaltek was proposed to account for a different set of facts, discussed below in section 4. This algorithm depended crucially on the fact that CP complements obligatorily extrapose in Jakaltek (Craig 1977:248). Thus, normal order is VSOX, but when O is a CP, it follows adverbials:

(21)   Xal naj tet anma   yul parke ewi
        said he to people in park yesterday

        [chub'il chim huluj naj presidente konhob'].
        that   may come the president village
        'He said to the people in the park yesterday that the president
        may come to the village.'

In contrast, relative clauses and tenseless complements do not extrapose:

(22)   Xitij   [naj   ah   hoyom [x'apni yet   qani
        he.brought   NCL(he) from T.S.   arrived when last night

        boj   sk'ahol Ø]] ixim.
        with his.son PRO corn
        'The [man] from Todos Santos [who came last night with his
        son] brought corn.' [194]

(23)   X'okkanh [ha=loq'ni ha=cheh]   yinh ha=k'ul.
        entered   you.buy your.horse in   your.stomach
        'You decided to buy yourself a horse.' [255]

I will not review the algorithm proposed in the earlier article except to say that it involved alignment of the right edge of an ungoverned maximal projection with the right edge of an intonational phrase. Under the phrase-structural assumptions of that article (which still seem reasonable), extraposed clauses and topics were not governed. What that algorithm achieved, and what must be achieved by any such algorithm, is a division of the sentence into (potentially) three spans, each corresponding to a separate intonational phrase: the topic, the body of the clause (containing the focus, core arguments, relative clauses, and tenseless complements), and what we might term the tail, consisting of any (extraposed) CP complement or subordinate adverbial clause.

(24)

TOPIC	BODY	TAIL (Extraposed CP)

Day's description of Jakaltek (1973) provides phonological evidence that each of these spans corresponds to a separate prosodic constituent. He distinguishes several forms of juncture, including one which corresponds to what I call *intonational phrase*:

The juncture ≠ sets off some major constituents from the remainder of the sentence, . . . Its occurrence with some constituents is optional. The more important the syntactic separation, the more likely that it will be marked by ≠. At some syntactic breaks ≠ is accompanied by a contour. [p. 20]

Relevant here are those instances of ≠ that are accompanied by contour (contour is indicated phonetically by a change in pitch on the last stressed syllable before the juncture). It is precisely those breaks that are accompanied by contour that mark the boundaries of the domain relevant to the distribution of the overt pronoun and its covert counterpart.

Day's description supports the spans that are identified in (24). It makes clear that topics and extraposed and adjunct clauses constitute their own intonational phrases (pp. 22–23, 87–88, 103), while relative clauses, tenseless complements, and preverbal focus do not (pp. 72, 84–87, 89). All three of the latter constructions are subsumed in larger intonational phrases.

How these three spans might correspond to the syntactic representation is shown in figure 1, following the phrase-structural assumptions of Aissen (1992). The topic (see Jelinek, this volume, for a different view of topic and focus) is adjoined to the root node; the focus occupies specifier of IP, and extraposed CPs are adjoined to VP.[10]

Figure 1
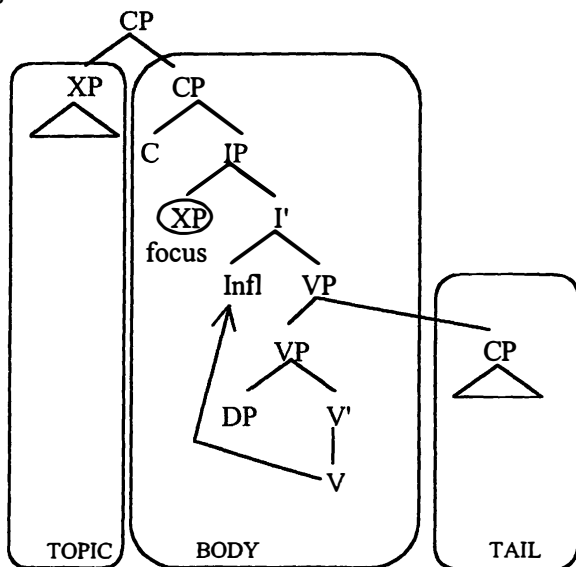
The division of the sentence into its intonational phrases provides the right structures for characterizing the distribution of Ø. Within a simple sentence, reference back to an antecedent within the same sentence can always be accomplished by Ø, as in examples (2) and (5)–(8). The Ø form is possible

because a simple sentence falls under a single intonational phrase. Within a relative clause or a tenseless (non-extraposed) complement, reference back to an antecedent outside that clause may also be accomplished by Ø, as in examples (12)–(13), (16)–(18). Again, this is true because such structures are part of larger intonational phrases, and as long as the antecedent is contained within that larger intonational phrase, reference back to it by Ø will be possible. Since extraposed and adjunct CPs constitute their own intonational phrases, reference back to an antecedent outside the CP always requires the overt pronominal; see (14)–(15).

There are differences in the prosodic relation of the topic and focus to the rest of the clause, and these differences generate a set of predictions concerning Ø. As noted above, the topic constitutes its own intonational phrase, while the focus is subsumed into a larger intonational phrase. Under the prosodic account developed here, (20) predicts that Ø within the body of the clause can be anteceded by a nominal within the focus. That this is true is shown by examples (8)–(9), repeated below:

(25)   [Sat [stx'at   *naj*]]$_j$ xhwayi Ø   t$_j$.
        on his.bed NCL$_{(his)}$ sleeps   PRO
        'It's on his$_i$ bed that he$_i$ sleeps.' [161]

(26)   [Ixim ixim k'ochb'il yu *ix*]$_j$ xitoj   Ø   t$_j$ yinh molino.
        NCL   corn   shelled   by NCL$_{(her)}$ she.took PRO   to   mill
        'It is the corn that was shelled by her$_i$ that she$_i$ took to the mill.' [167]

However, since the topic is contained in a separate intonational phrase from what follows, Ø within the body of the clause will never be anteceded by the topic. Accordingly, the topic is always resumed by the appropriate overt pronominal classifier (if one exists), as in (27)–(28):

(27)   *Naj pel*   [smaq   *naj/*Ø*   ix].
        NCL Peter   he.hit   NCL$_{(he)}$/ PRO   NCL$_{(her)}$
        'Peter$_i$, he$_i$ hit her.' [12]

(28)   *Wal naj tz'um   wex   xin*, [xal   ya' mame tet *naj*   ta...]
        but NCL leather   pants ENC he.said NCL father to   NCL$_{(him)}$ that...
        'But Leather Pants$_i$, the father said to him$_i$ that ...' [Datz 1980, 374]

I conclude, then, that it is possible to define the domain within which Ø must be bound as the intonational phrase—i.e. in prosodic terms. Can this domain be defined in syntactic terms? One would expect that it might be possible, since intonational phrasing is determined by syntactic structure. One possibility, suggested by the structure in figure 1, would be to say that Ø must be bound within the lowest CP that dominates it. This is Trechsel's proposal (1995), and it corresponds roughly to an update of Craig's original proposal, which was stated in terms of syntactic boundaries. Such an account makes most of the same predictions as the prosodic account. The two accounts might make

different predictions about relative clauses, depending on whether they are CPs or not. If they are, then the syntactic account makes the wrong prediction (see examples 16–18), while the prosodic account makes the right prediction. Relative clauses in Jakaltek contain no overt relative pronoun or complementizer, so it is not clear whether they are CPs or not; Trechsel (1995) assumes they are not.

However, even if the syntactic account is descriptively adequate, it seems less satisfactory on several fronts than the prosodic account. To judge from Reuland and Koster (1991), setting the binding domain for Ø as the lowest CP which contains it neither aligns Jakaltek with other well-known cases, nor does it make any sense of the fact that this binding is blind to hierarchical structure and sensitive rather to precedence. In contrast, defining the domain prosodically suggests a way to understand the importance of precedence, and the irrelevance of c-command. While hierarchical syntactic structure, as constituted by the c-command relations among elements of various categories, is relevant to the determination of prosodic phrasing, once that phrasing is determined, c-command and category type become irrelevant, and what is left as the audible form of syntactic structure is precisely precedence and phrasing. Thus, characterizing the domain condition on Ø in prosodic terms predicts exactly what is found: The domain corresponds to a prosodic constituent, and the relevant relation is precedence, not c-command. A characterization of the domain in syntactic terms (minimal containing CP) still leaves the irrelevance of c-command without any explanation.

There is further evidence that the domain in which Ø is licensed corresponds to a prosodic constituent: This same domain defines the host for the enclitic *an*, a clitic which occurs optionally in sentences containing a first person pronoun.

4. Intonational Phrase as Host for Encliticization

The clitic *an* is subject to two licensing conditions, one prosodic and one syntactic: It attaches to the right edge of a prosodic domain; and it occurs only in domains which contain a first person pronoun.[11] The intonational phrase is the relevant domain for both conditions. The grammatical function of the licensing pronoun is irrelevant, and *an* may be located far from the position canonically associated with that function. The examples in (29) illustrate *an* licensed by a first person which is, respectively, subject, object, and genitive. Here, the agreement morphology associated with the licensor is italicized; first and second person pronouns usually do not occur overtly, but are cross-referenced on governing heads. Cross-hatch indicates an intonational phrase boundary.

(29) a. Xk*in*=to   hawatut   *an*. #
         I.went   your.house   1ST
         'I went to your house.' [278]

b. Xk*in*=hawil tx'onhb'al   *an*. #
   you.saw.me market   1ST
   'You saw me in the market.' [278]

c. Xkam   *hin*cheh   *an*. #
   died   my.horse   1ST
   'My horse died.' [277]

Craig (1977) defined the distribution of *an* in terms of the same syntactic boundaries which she saw as relevant to the distribution of Ø and the overt classifiers. However, that the position of *an* should be defined phonologically, rather than syntactically, is suggested by Day (1973a:56), who calls *an* a "sentence clitic" which occurs "only before contour". The term "sentence clitic" is a misnomer, since *an* frequently occurs *internal to* what we would term a sentence on syntactic grounds. In Aissen (1992), *an* is analyzed as a clitic which attaches to the right edge of an intonational phrase, subject to the condition that that phrase contain a first person pronoun (the licensor). While *an* is optional when the first person licensor is singular, it is obligatory when the licensor is first person plural exclusive, and impossible when the licensor is first person plural inclusive (Craig 1977:277). Thus, *an* has two functions, one prosodic, the other semantic: It demarcates the right edge of an intonational phrase, and it distinguishes first person plural exclusive from inclusive.

Since simple clauses form a single intonational phrase, *an* occurs sentence-finally in examples like (29). More complex examples confirm that the domain in which *an* must find its licensor coincides exactly with the domain in which the covert anaphor, Ø, must find its antecedent.

In structures containing tenseless complements and relative clauses, *an* occurs at the right edge of the sentence, even when its licensor occurs in the main clause:

(30) *tenseless complement*
      Xil   *wanab'*   [$_{VP}$hawek'   yul kaya] *an*. #
      she.saw my.sister   you.pass   in   street 1ST
      'My sister saw you go by in the street.' [279]

(31) *relative clause*
      *W*ohtaj naj$_i$ [xul   t$_i$ ewi]   *an*. #
      I.know NCL$_{(he)}$ came   yesterday   1ST
      'I know the man who came yesterday.' [279]

By hypothesis, this is true because the whole sentence falls under a single intonational phrase, and *an* is licensed at the right edge of that intonational phrase. Example (32) illustrates rather dramatically the fact that *an* occurs sentence-finally even when its licensor is located within a sentence-internal relative clause:

(32) Xitoj   [ya' komam   [xhmunla yinh   *hin* mam]]
      he.sent NCL older.man works   for   my father

[no'   chech [i'o' ixim   ixim tinanh]] *an.* #
NCL   horse carry NCL   corn today   1ST
'The old man who works for my father sent the horse to carry
the corn today.' [279]

Here again, this complex sentence falls under a single intonational phrase, so
that *an* is prosodically licensed; it is also syntactically licensed, since that
intonational phrase contains a first person pronoun. Examples (30)–(32) parallel
(12)–(13) and (16)–(18) above in which the Ø anaphor within a tenseless
complement or a relative clause finds its antecedent in the matrix clause.
    The facts are different in the case of extraposed complements and adjunct
clauses. When the main clause contains a first person, then any *an* which it
licenses must occur *before* the extraposed complement (33) or adjunct (34),
rather than sentence-finally:

(33) X*w*al tet naj   (*an*) # [CP chub'il ch'ahtoj naj
      I.say to NCL(him) 1ST   that   climb.up NCL(he)

swi' te' nhah] (**an*). #
its.top NCL house 1ST
'I told him to climb on the roof.' [281]

(34) Lanhan *hin*tx'ahni xil qape   (*an*) # [yet   xkach=huli] (**an*). #
      ASP   I.wash   NCL clothes 1ST   when you.came 1ST
      'I was washing clothes when you came.' [280]

These examples fall under two intonational phrases, with the break marked
exactly by the position of *an*. The *prosodic* licensing condition is satisfied by the
ungrammatical examples, but the *syntactic* condition is not, since the second
intonational phrase does not contain a first person. These examples parallel
(14)–(15) above, which show that the Ø anaphor within an extraposed clause
cannot be anteceded by an element within the main clause, since the two are not
within the same intonational phrase.
    As expected, if the antecedent for *an* is within the extraposed clause, then
an occurs sentence-finally:

(35) Xal   naj   # [chubil x'apni *hin* mam *w*atut]   *an.* #
      he.said NCL(he)   that   arrived my father my.home   1ST
      'He said that my father had arrived at my house.' [284]

In this case, both licensing conditions are satisfied. (See also example (15), in
which *an* is licensed twice, in distinct intonational phrases.)

The differences seen earlier between topic and focus are replicated in the distribution of *an*. If the topic contains a first person, then *an* is licensed but occurs between the topic and the remaining sentence. Since this position coincides with an intonational phrase break, *an* is licensed here prosodically:

(36) *W*uxhtaj·   *an* # [sloq'   ho'   no' cheh k'ej'inh tu'].#
      my.brother   1ST   he.brought NCL(he)   NCL horse black   that
      'My brother, he bought that black horse.' [280]

Note that the pronoun which resumes the topic (*ho'*) must be overt, because its antecedent is located in a different intonational phrase.

However, when the focus contains a first person, *an* occurs not immediately following the focus, but at the first intonational phrase break following the focus:

(37) [Ha' [*hin* mam]ᵢ] xal   naj   *an* # [chubil x'apni   tᵢ]. #
      FOC   my father   he.said   NCL(he) 1ST   that   arrived
      'It is my father that he said had arrived.' [283]

Example (37) is particularly interesting because the focus is extracted from the embedded (extraposed clause). Nonetheless, it is contained within the first intonational phrase, not the second, and it is this that determines which intonational phrase hosts cliticization of *an*.[12]

I conclude, then, that the domain which Craig established as relevant to the distribution of the pronominal classifiers and Ø corresponds to a prosodic domain—exactly the domain that determines the position and distribution of *an*. This predicts certain interactions between pronoun realization and *an*. In a sequence of clauses like that of (38), in which the first clause contains a first person and a third person, the presence of *an* between the two clauses signals the edge of an intonational phrase; accordingly, reference in second clause back to the same third person will require an overt classifier.

(38) [....third personⱼ...] *an* # [.....NCLᵢ..]#

Examples (15) and (36) above conform to this prediction. So does the text example in (39), where the pronoun *ix* 'she' in the second clause is anteceded by *ix yajaw kusina*, 'the cook' in the first:

(39) Chi*n*to *w*ala' tet ix   yajaw   kusina-h   *an* #
      I.go   I.talk to NCL mistress kitchen   1ST

      [ta   swatx'e   ix   hawan]. #
      that   she.make NCL(she) your.medicine
      'I'm going to tell the cookᵢ that sheᵢ make medicine for you (lit. your medicine).' [Datz 1980:407]

It appears, then, that the domain in which *an* is licensed coincides with that in which Ø must be bound. While it is of course possible that the two domains accidentally coincide, the position taken here is that they coincide because the distribution of both *an* and Ø is defined over the same prosodic domain. The position and distribution of *an* and the alternation between Ø and the overt pronominal classifiers work together to mark the chunking of discourse into intonational phrases.

5. Conclusion

The proper account of the covert anaphor Ø in Jakaltek requires simultaneous reference to two kinds of information, prosodic and syntactic. The need to refer to prosodic structure in an account of Jakaltek Ø could be satisfied quite directly within a model like the Minimalist program by assuming that Ø arises through deletion of a pronominal classifier at PF. This deletion, which represents a return to Craig's original analysis, might be analogized to the PF deletion which Chomsky envisages for ellipsis (1995:202ff.). However, PF cannot be the only level involved. Deletion depends on co-indexation, and the relevant indices are not present in prosodic structure. They are present at LF, and this implies the need for some communication between PF and LF.

The Jakaltek material thus provides support for the view that the phonology and syntax cannot be entirely isolated from one another, and further, that the interaction is not limited to the construction of prosodic structure on the basis of a syntactic representation. What is required is not wholesale access to the phonology, but just access to prosodic structure, and this is consistent with the suggestion of Zec and Inkelas (1990) that such communication be mediated exactly by prosodic structure.

Acknowledgments

Notes

1. Unless otherwise noted, all page references are to Craig (1977). Throughout, orthography has been changed from original sources to conform with that put forth by Academia de las lenguas mayas (1988). However, word boundaries in the original sources are retained. Unfamiliar symbols may include {tx, tx', x} which represent a series of retroflex consonants, {ch, ch', xh}, a series of palatal consonants, the velar nasal {nh} and the velar fricative {j}. Abbreviations which appear in glosses are: ASP: aspect;

ENC: enclitic; FOC: focus marker; NCL: noun classifier; P: preposition; PL: plural: PRO: Ø anaphor; 1ST: 1st person enclitic. The symbol = links prosodically distinct words which are glossed as a unit.

2. In Jakaltek, verbs are rigidly classified as transitive or intransitive. Transitive verbs agree with both subject and object through a system of agreement which is organized along ergative lines in tensed clauses, and along what has been called "extended ergative" in untensed clauses. That is, the ergative markers are extended to intransitive subjects in untensed clauses. Furthermore, nouns agree with their possessors. Morphological analysis of words is not provided, but nouns and verbs which contain an agreement marker are glossed with the corresponding pronoun (e.g., *smi'* 'her.mother' in (1) and *xitoj* 'she.took' in (9)). In tensed clauses with third person subjects, this distinguishes transitive from intransitive verbs, since third person intransitive subjects are not cross-referenced. Thus, *xitoj* in (9), glossed 'she.took', is transitive, while *xkolwa* in (1), glossed 'helped', is intransitive.

3. The classifiers are glossed throughout by NCL, regardless of function. An English translation is given in parentheses for classifiers in pronominal function.

4. The verbs in (7)–(8) end in the terminal element *-i*, which appears only before contour; if the overt subject pronoun were possible in these examples, *-i* would not appear, yielding *xto* and *xhway*.

5. According to Victor Montejo, the second pronoun in (9) may also be overt (i.e., *ix*).

6. In what follows, I continue to refer to the relation between Ø and the expression which licenses it—e.g., *ix* in 9—as *binding*, and I use the term *antecedent* interchangeably with *binder*.

7. Thanks to Victor Montejo, who confirmed that (10) is ungrammatical if the first pronoun is covert; likewise (11).

8. According to Craig (1977), the Ø form in (13) can be replaced by the overt form of the pronoun without resulting in disjoint reference. Craig attributes this to two factors: The fact that the pronoun is a genitive, and the fact that it is located within a different clause than its antecedent. Some speakers also allow the overt pronoun in (12).

9. There is one regular exception to this which refers to the grammatical function of Ø. Craig (1977) notes that Ø in genitive function can find its antecedent outside the domain which she identifies as relevant to Ø, and the texts of Datz (1980) confirm this. When Ø occurs in object function in the second of two coordinate clauses, it is sometimes omitted if its antecedent has parallel function in the first coordinate clause (Craig 1977:169ff).

10. Extraposed CPs might be adjoined higher.

11. I treat *an* here not as the first person pronoun itself, but as an element which must be licensed by the presence of a first person pronoun somewhere within the relevant domain. Two problems arise if *an* itself is interpreted as the first person (argumental) pronoun. First, *an* may co-occur with the lexeme which is usually considered the first person pronoun, *hayin* 'I'. Second, in domains containing several first person pronouns (e.g., analogues of 'I went to my house'), only one instance of *an* is ever possible.

12. In Aissen (1992), I considered the question whether the constituent to which *an* attaches could be defined syntactically, and concluded that it could not. This conclusion depends crucially on where the extraposed clause attaches.

The Syntax of Verb Initial Languages

Edited by
Andrew Carnie and Eithne Guilfoyle

OXFORD
UNIVERSITY PRESS
2000

Maui
P
295
.S94
2000)

OXFORD
UNIVERSITY PRESS

Oxford   New York
Athens   Auckland   Bangkok   Bogotá   Buenos Aires   Calcutta
Cape Town   Chennai   Dar es Salaam   Delhi   Florence   Hong Kong   Istanbul
Karachi   Kuala Lumpur   Madrid   Melbourne   Mexico City   Mumbai
Nairobi   Paris   São Paulo   Singapore   Taipei   Tokyo   Toronto   Warsaw

and associated companies in
Berlin   Ibadan

1 3 5 7 9 8 6 4 2

Printed in the United States of America
on acid-free paper

Contents