The Syntactic Structure of Evidentiality in Shipibo
José Camacho, José A. Elías-Ulloa, Rutgers University
jcamacho@rutgers.edu, joselias@eden.rutgers.edu

Introduction

Over the past two decades, generative linguists have investigated whether there is a universal template for clausal structure and, if so, what categories are part of that template and how they are ordered. Thus, researchers dating at least back to Pollock (1989) have debated what types of functional categories are necessary to account for specific syntactic distributions in several languages, as well as what their hierarchical order might be (see also Baker (1988), Belletti (1990), Rizzi (1997), Cinque (1999), Cinque (2002) among many others). Perhaps the most explicit and well-documented proposal in this line of research is Cinque (1999), who hypothesizes a very rich universal clausal template based on a detailed cross-linguistic study of adverbs and heads. This research program postulates a very rigid, cross-linguistic order of functional categories.

In this paper, I will focus on the higher part of Cinque’s clausal structure, in particular on projections related to point-of-view (PoV). For this area, Cinque (1999) has proposed the hierarchy in (1). More recently, Speas (2004), building on Cinque’s analysis, has proposed the structure in (2) to account for the cross-linguistic distribution of evidentiality-related phenomena and for the correspondence between evidentiality and logophoricity.

(1) \[ \text{Mood}_{\text{speechact}} \rightarrow \text{Mood}_{\text{evaluative}} \rightarrow \text{Mood}_{\text{evidential}} \rightarrow \text{Mod}_{\text{epistemic}} \rightarrow \text{T(past)} \rightarrow \text{T(Future)} \rightarrow \text{Mood}_{\text{irrealis}} \rightarrow \ldots \]

(2) \[ \text{Speech Act Phrase} > \text{Evaluative Phrase} > \text{Evidentiality Phrase} > \text{Epistemic Phrase} \]

I will contrast the predictions these two lines of analysis make for the evidentiality system of Shipibo (a Panoan language from Eastern Peru), and point out some of the shortcomings they face, as well as possible alternatives that maintain the basic insights but avoid the shortcomings.

1 The Structure of Point of View

1.1 The syntactic representation of evidentiality

Speas (2004) argues against a pragmatic analysis of evidentiality and in favor of a syntactic representation of these morphemes. Her arguments involve, on the one hand, the restricted and systematic nature of the information conveyed by these morphemes, and on the other, the fact that they closely interact with other inflectional heads, such as tense and aspect. As an example of the first type of evidence, consider the type of information on the source of information a fairly complex evidential system may convey: visual, inferential, reported. Variation may occur in several places: whether visual is sensory (including other senses), or whether sensory is divided into visual and non-visual, whether there are one or two types of inferential info, etc. (cf. Aikhenvald & Dixon (2003, chapt. 1) and Willett (1988)). However, as Speas points out (2004, p. 257), evidential systems do not make reference to potentially salient sources of information in many cultures, such as experience reported by loved one; divine revelation; legal edict; parental advice (“Momism”); Heartfelt intuition (“gut feeling”); learned through trial and error; Teachings of prominent elder/authority. If evidentiality were content related to pragmatics, one would expect some of these sources to surface as sources.

1
Regarding interaction with other inflectional categories, consider the case of Qiang, a Tibeto-Burman language (cf. LaPolla (2003)). As LaPolla suggests, in Qiang, the inferential/mirative suffix /-k/ takes the inferential sense when the action involved is an activity (cf. (3a)), but if the reported event is a state or the result of some action, then the meaning is mirative, cf. (3b) (exs. (9a-b) from LaPolla 2003):

(3) a. the: \( zdzyta; \) fia-q\( \bar{a} \)-k
3SG Chengdu+LOC OR-go-INF

‘He went to Chengdu. (Used in a situation where the speaker knew the person was supposed to go to Chengdu, but wasn’t sure when, and then saw the person’s luggage gone, so assumed he had left for Chengdu. /-k/ could not be used if the speaker saw the person leave).’

b. the: ct\( \tilde{c} \)imi zdzi-i-k!
3SG heart sick-INF

‘He’s unhappy!’ (just discovered; relatively sure, not a guess)

As LaPolla points out, it is perfectly possible to express an inferential sense about a state or a perfective situation, but in that case, one cannot use the inferential /-k/, but rather an adverbial such as /xsuni/ ‘seem’, or the construction [-tan] or [-lahan] (LaPolla’s ex. (10)):

(4) dzy zge-m-tan yua.
door open-NMLZ-appearance COP

‘It appears the door is open.’/‘Apparently the door is open.’

The interaction between evidential morphemes and aspect in Qiang does not easily lend itself to an obvious pragmatic account, and the fact that states and perfective situations draw the mirative meaning in /-k/ but activities the inferential one does not seem to follow from any universal or principled semantic constraint, given the availability of the alternative in (4), but rather from a language-specific restriction on the syntactic combination of evidentiality and aspect.

Another argument in favor of the need to syntactically project evidential morphemes involves scope interactions among them, as pointed out by Valenzuela (2003), fn. 2. In Shipibo, an inferential can be combined with a reportative, yielding a potentially ambiguous sequence: either the speaker reports someone else’s inference, or s/he expresses an inference from reported evidence. The latter sense is the one attested in texts, and the one given by Valenzuela in her example below (example (1)f, shown below), but she suggests speakers allow for both interpretations. Given the evidentiality hierarchy proposed by Oswalt (1986), Willett (1988), the second interpretation (inference > hearsay) is the one expected.

(5) Ani-ronki i-bira-[a]-i jawen jema.
large-REAL be-INFRA-INCL POSS3 village

‘Her village must be large (from what I have heard).’

1.1.1 The evidentiality hierarchy

Several studies have tried to explore whether propositional source-marking obeys some kind of cross-linguistic tendency. Willett (1988, p. 57), for example, finds in a survey of 38 languages, that evidentiality falls within three categories: attested (which can include subdivisions that include visual, auditory or other sensory information), reported, and inferred. Speas (2004, p. 4), on the other hand, proposes four hierarchical categories: personal experience, direct (sensory) evidence, indirect evidence and reported evidence (hearsay). Aikhenvald & Dixon (2003, p. 3), suggest two broad types of evidential systems: those that state the existence of a source of evidence without specifying it, and those that specify the source. Within the second type, they discuss several subtypes:
Two-distinction systems:
1. Eyewitness and noneyewitness.
2. Nonfirsthand and everything else.
3. Reported and everything else.

Three-distinction systems:
1. Visual, inferred, reported.
2. Visual, nonvisual sensory, inferred.

Four-distinction systems:
1. Visual, nonvisual sensory, inferred, reported.
2. Visual, inferred (2), reported.
3. Nonvisual sensory, inferred (2), reported.
4. Visual, inferred, reported (2).

It seems clear, particularly for three and four-distinction systems, that they are built on Willett’s three basic distinctions, with some possible subdivisions within each of those categories. Speas, following Oswalt (1986), suggests an additional category different from visual-sensory: personal experience. Willett, Oswalt and Speas suggest that there is an underlying hierarchy corresponding to the degree in which the source directly involves the speaker’s evidence. This scale goes from more direct experience to no experience at all (Speas 2004, pg.258):

personal experience >> direct (e.g. sensory) evidence >> indirect evidence >> hearsay

This hierarchy is supposed to constrain the way in which languages encode evidentiality. Thus, we can interpret the classifications presented in (6)-(8) as cuts at different points in the hierarchy, as exemplified below for some of the cases:

Given this hierarchy, one can formalize the fact that evidential markers may encode meanings that are adjacent in the hierarchy (for example direct and indirect evidence), but not those do not encode two non-adjacent meanings (direct evidence and hearsay vs. indirect evidence). Note that the meanings encoded in the hierarchy in (9) are linguistically constrained in both the type of information encoded and the way this information is grammaticalized.

Speas goes one step further and proposes that the evidentiality hierarchy in (9) corresponds to a syntactic hierarchy, the one already quoted in (2), above, which in turn is essentially Cinque’s hierarchy. In order to capture why only four PoV types seem to be attested cross-linguistically, Speas suggests that PoV projections have an implicit subject argument subject to local binding.1 According to Speas, coreferential relations among those pronominal categories yield precisely four possible combinations, illustrated in (11), where each pro is referred to using some mnemonic term.

---

1The details of this system are not clearly spelled out; if pros must be locally bound, in the case of a disjointly indexed pro, (as in the case of perceiver in (11b)), it is not clear what the binder is.
There is an alternative way of achieving the same result: if one assumes that only one of the PoV heads can be present or filled at a time, then one predicts exactly the four PoV types in (9). Either analysis can account for languages in which evidentiality morphemes are in strict complementary distribution, but not for languages like Shipibo, where it is possible to have two evidentials in the same clause, as I will now illustrate.\(^2\)

2 The evidential paradigm in Shipibo

The evidential paradigm in Shipibo raises some interesting questions about the structure of PoV projections, both in that language and in general. There are two clear-cut evidential morphemes: -ra ‘direct evidence, and -ronki ‘reportative’. There is also a potential second reportative, -ki, whose syntax is slightly different from the other two. Additionally, there are two inferentials: -mein/-main (in free distribution) ‘speculative’ and -bira ‘inferential’. Valenzuela suggests that -ra/-ronki and -bira/-mein belong to two separate layers of evidentiality, but, as we will see below, their syntactic distribution overlaps substantially. Each morpheme is illustrated below ((12a) is from our fieldwork, (12b) from Loriot, Lauriault, & Day (1993, p. 23), glosses are mine, (13) from Valenzuela 2003, p. 44 and p. 47 respectively).


‘I live with my brothers and my parents there.’

b. Moatian-ronki Yoashico Inka rete-kan-a iki long ago-REP Stingy Inca kill-PL-PP AUX

‘Long ago, they killed the Stingy Inca’ (it is said).

(13) a. Mi-n bake pi-kas-bira-[a]i, oin-we! 2-GEN child:ABS eat-DES-INF-INF-INC see-IMP

‘Probably your child is hungry, go see her!’

b. Tso-a-mein i-ti iki? who-ABS-SPEC* be-INF AUX

‘Who could it be?’

The first thing that distinguishes -ra/-ronki from -bira/-mein is that -ra is in complementary distribution with it -ronki, and -bira with -mein, but, according to Valenzuela (2003, p. 34), combinations of -bira and -ra are possible and frequent, whereas combinations of -mein and -ra are less common. Additionally, -bira/-mein can coappear with -ronki, as illustrated in (14) (from Valenzuela 2003).

(14) a. Ani-ra i-bira-[a]i jawen jema. large-DIR.EV be-INF-INC POSS3 village:ABS

‘Her village must be large.’ (because it has a secondary school)

\(^2\)The same argument can be made for Qiang.
b. Ani-ronki i-bira-ai jawen jema.
large-REP be-INFN-INC POSS3 village:ABS
‘Her village must be large.’ (from what I have heard)

c. Ani-mein(-ronki) iki jawen jema.
large-SPECL-REP COP POSS3 village:ABS
‘Perhaps her village is large.’ (from what I have heard)

There are additional differences and similarities between the different PoV morphemes: -ra, -ronki and -mein are usually second position clitics: they appear after the first syntactic phrase in the clause, excluding certain adjoined material, as illustrated in (15) (first sentence from a narration of a story taken from our fieldwork, the other two from Valenzuela (2003, p. 41 and 34)). In the first example, the NP including the comitative complement is treated as the first syntactic constituent, in the second case, the subordinate verb kepentanan is the first constituent, with -ra attached to it. When the first word is a finite, main verb, then -ra, -ronki and -mein appear inside the verbal complex, possibly followed only by the number and the aspect morphemes, as seen in (16) (first example from our fieldwork, second one from Valenzuela (2003, p. 36), third one from Loriot et al. 1993, p. 241).

(15) a. Ochiti-betan bake-ra kaja noko-ai
dog-with boy-DIR.EV box find-INC
‘The boy with the dog found a box.’

b. E-a-ronki/-ra tima-nan-iba-ke
1-ABS-REP/DIR.EV punch-REC-PST2-COMPL
‘I had a fight.’

c. Ani-mein iki jawen jema.
large-SPECL COP POSS3 village:ABS
‘Perhaps her village is large (from what I have heard).’

(16) a. Ja-ra-ke westiora bake westiora ochiti kaja-ya
be-DIR.EV-ASP a boy a dog box-with
‘There was a boy and dog with box.’

b. Bi-bain-yama-ra-kan-ke
get-going.TR-NEG-DIR.EV-PL-COMPL
‘They did not take it while going’/‘They did not take it and left.’

c. Ka-main-ke Ikito-ain.
ir-SPECL-CMPL Iquitos-to
‘They must have gone to Iquitos, right?’

Whereas -mein can coappear with -ra and -ronki, the latter two evidentials cannot coexist. Thus, in (15c), -ronki can optionally appear after -mein. The impossibility of having -ra and -ronki in the same clause may be due to semantic reasons, since -ra signals first hand information and -ronki reported information.

It should be fairly clear that the availability of two distinct positions for certain PoV morphemes in Shipibo makes a unified analysis very challenging. If PoV morphemes are generated in the highest functional projection of the clause, and attached to a constituent to their left, then it is very hard to explain why they can also surface inside the verb, preceding at least two other morphemes. Conversely, if they are generated as part of the inflectional functional complex of the verb (as in (16), for example), it is not clear why they can also surface as second position clitics.
By contrast to the other three evidentials, -bira need not appear in second position obligatorily: it can appear within the verbal complex even if the verb is not the first or only constituent of the clause, as shown in (17a), from Loriot et al., p. 120 (my analysis). Additionally, it can also coappear with -ra/-ronki, as shown in that same example. However, in such case, -bira appears on the verb, as opposed to -mein, which is still a second position clitic. According to Valenzuela, -bira can appear with other constituents, which are focused, as in (17b), from Valenzuela, p. 45-46. Finally, -bira has other related meanings, such as ‘more or less’, shown in (18), from Loriot et al., p. 120.

(17) a. Mi-a-ra yoshi-man bi-bira-ke
   2-ABS-DIR.EV demon-ERG catch-INF.R-COMPL
   A demon (must have) caught you’

   b. Mi-on-bira keen-kin-ra mi-a a-ke, mi-a benta-u yatan-ke
   2-INTRST-INF want-SSST-DIR.EV 2-ABS do.TR-COMPL 2-ABS benta-ERG hold-CMPL
   ‘Probably it is because he wants YOU, that the benta took you.’

(18) No-a chonka nete-bira jo-ti iki
    1.PL-ABS ten days-bira return-FUT.INF AUX
    ‘We will return in aproximately ten days.’

One final distributional difference in the PoV paradigm involves -ra and -ronki. Valenzuela observes that inside the verbal complex, “-ronki cannot be followed by the completive aspect marker -ke or the plural -kan unless an auxiliary verb is placed in between (p. 38)”, as illustrated in (19), (her (12c-f), confirmed by one of our native speaker collaborator).

    do.TR-REP-COMPL
    ‘It is said that s/he did (it)./S/he says that s/he did it.’

   b. A-ronki-a-ke.
    do.TR-REP-do.TR-COMPL
    ‘It is said that s/he did (it)./S/he says that s/he did it.’

    do.TR-REP-PL-PP AUX
    ‘(It is said that) they did (it).’

    TR-REP-do-PL-PP AUX
    ‘(It is said that) they did (it).’

(20) summarizes the issues PoV morphemes in Shipibo raise.

(20) a. Second position vs. verb-internal position

   1. Accounting for the two possible positions.

   2. The contrast between -bira (verb-internal evidential) and the other (second position clitics or verb-internal evidentials).

b. Verb-internal position contrasts between -ra and -ronki.

c. Coappearance of evidential morphemes.
2.1 One evidential or two evidential positions?

Let us first propose what would, a priori, be the simplest hypothesis, namely that the second position and the verb-internal position have a unified analysis, namely there is a single evidential position and the different orders are derived by movement. It should immediately become clear that under standard syntactic assumptions, the unified analysis is untenable. If we assume head-to-head movement of verbal morphemes, the derivation of the verb-internal cases (cf. (16) above) implies the hierarchy Aspect (-ke) > number > evidentiality > other heads. Under this analysis, in an example like (16a), ja-ra-ke the verb ja head-moves to -ra (evidentiality) an then to -ke (aspect). The fact that the verb is the first constituent in such cases raises further questions: if the verb has reached that position by head-movement, it must be the case that heads are initial, otherwise we would expect complements to precede the verb, contrary to what happens in (16). Alternatively, it could be that a maximal phrase containing the verb has moved to first position, presumably the specifier of the highest projection in the clause. If the generalization that the evidential morpheme only surfaces verb-externally when the verb is the first or only constituent of the clause is true (as both Valenzuela (2002, p. 36) and Loriot, Lauriault, & Day (1993, p.374) suggest), then examples such as (??), from Valenzuela, p. 37, seem to suggest that a maximal projection has moved to first position, since the verb with the evidential is in second position, followed by a light verb.

Under those same assumptions (head-movement and a single underlying evidential position), second position examples for the evidential (cf. (15)) are impossible to derive: why can the verb skip the evidential projection located lower than aspect and proceed to aspect, yielding, for example tima-nan-iba-ke V-REC-PST2-COMPL ‘punched each other’? In order to derive the contrast between second position and verb-internal position for the evidential we need to postulate either excorporation of -ra, or that the verb can skip a head, violating the Head Movement Constraint.

One alternative would be not to assume obligatory verb-raising. In such a case, we could argue that the second position is the highest projection of the clause (along the lines of Elias-Ulloa (2002) for another Pano language, Capanañua), the evidential heads that projection, and constituents to its right appear in the specifier of it. With this assumption, the morpheme ordering inside the verbal complex would not reveal anything about the hierarchy of functional heads (i.e. Baker’s 1985 Mirror Principle does not hold in these cases), and the morphological selectional restrictions for evidentials would allow them to adjoin to anything to their left or to a very restricted set of inflectional affixes.

A different line of analysis on this distribution argues that there are two evidential positions: the highest one in the clause, and one lower, that reflects the morpheme order inside the verb. This lower position would be dominated by aspect and number, but would dominate other projections, such as negation and tense. While more needs to be said about them, examples such as (14a) above, repeated below, where two evidentials coappear would seem to argue in favor of this possibility.

   large-DIR.EV be-INFRIN-INC poss3 village:ABS
   ‘Her village must be large.’ (because it has a secondary school)

(21) Shee a-ra-kan-ai yapa
    ONOM:frying do.TR-DIR.EV-PL.INC fish:ABS
    ‘Fish is being fried. (I smell it and hear it, but cannot see it)’

Under those same assumptions (head-movement and a single underlying evidential position), second position examples for the evidential (cf. (15)) are impossible to derive: why can the verb skip the evidential projection located lower than aspect and proceed to aspect, yielding, for example tima-nan-iba-ke V-REC-PST2-COMPL ‘punched each other’? In order to derive the contrast between second position and verb-internal position for the evidential we need to postulate either excorporation of -ra, or that the verb can skip a head, violating the Head Movement Constraint.
References


