

Areal diffusion and the development of evidentiality

Evidence from Hup*

Patience Epps

Max Planck Institute for Evolutionary Anthropology, Leipzig

Evidentiality is prone to diffusion and has been identified as a diagnostic feature of linguistic areas such as the Vaupés region of the Brazilian Amazon (e.g., Aikhenvald and Dixon 1998). This paper examines the processes by which a complex evidentiality system can develop in a particular language, catalyzed by language contact but fed by language-internal resources. The discussion considers data from Hup, a Vaupés language of the Vaupés-Japurá (Makú) family, and demonstrates that Hup has developed an evidentiality system parallel to those found in the two other unrelated language families of the region. Finally, a reconstruction of an evidentiality distinction for the Vaupés-Japurá family challenges Aikhenvald and Dixon's (1998) claim that evidentiality had two independent points of innovation in northern Amazonia.

1. Introduction and background

Amazonia is well known as a region with a remarkably high diversity of genetically unrelated languages, which nonetheless in many cases share strikingly similar features. The probability that many of these linguistic features have diffused across language boundaries suggests that the Amazonian lowlands encompass one or more broad linguistic areas or Sprachbunds, which may be further divisible into a number of sub-areas (cf. David Payne 1990, Doris Payne 1990, 2001, Aikhenvald and Dixon 1998). One of these, the Vaupés region of the Brazilian and Colombian border (see Map 1), has been discussed frequently in the literature as a place where linguistic exogamy, or obligatory marriage between language groups, has fostered the diffusion of many structural features

through intense language contact (Sorensen 1967; Gómez-Imbert 1996; Aikhenvald 2001, 2002).

The Vaupés region is here considered to be a “strong” linguistic area in the sense of Campbell et al. (1986: 536), in that its languages display a number of “shared traits [that] can be shown historically to be diffused — and cannot be ascribed to a common ancestor, to chance, or to universals” (cf. Aikhenvald and Dixon 1998: 244). A number of such traits — including tone/pitch-accent, nasalization as morpheme-level prosody, and verb serialization or compounding — have been identified as occurring consistently among Vaupés languages, but are crucially absent from some related languages outside the region, thus providing evidence for diffusion (primarily involving unilateral effects of Tukanooan languages on the Arawak language Tariana; see Aikhenvald and Dixon 1998; Aikhenvald 1999, 2002, etc.).

One feature that has been considered as particularly “diagnostic” of a linguistic area is evidentiality (Aikhenvald and Dixon 1998), here defined as a grammaticalized system for indicating the source of the information presented in a clause (cf. Chafe and Nichols 1986; Aikhenvald 2003b). Grammaticalized evidential systems are both relatively uncommon typologically — reported to occur in only about one quarter of the world’s languages [Aikhenvald 2004]¹ — and they tend to be clustered in relatively localized geographical regions (such as the Balkans, the Caucasus, the Himalayan region, and parts of North and South America), suggesting that evidentiality is extremely prone to diffusion (e.g., Sherzer 1976, Aikhenvald and Dixon 2003b, de Haan 2005). Thus the chances of evidentiality’s independent innovation are probably considerably lower than twenty-five percent; accordingly, Aikhenvald and Dixon (1998: 245) observe that the presence of evidentiality in adjacent languages is more likely to be the product of diffusion than of chance or universal linguistic tendencies. The fact that evidentiality is fairly wide-spread in Amazonia and neighboring regions — occurring in various language families including Arawak, Tukanooan, Yanomamian, Nambiquara, Panoan (cf. Aikhenvald and Dixon 1998), Cariban (e.g., Meira 2000, Mattei-Muller 2004), and in Quechua and Aymara in the Andes (e.g., Weber 1986, Hardman 1986) — therefore supports the characterization of Amazonia as a linguistic area². If we zero in on the Vaupés region, moreover, we find that the parallel complexity of the evidential systems found in these languages, in contrast to the minimal evidential specifications (or lack thereof) in their sister languages outside the Vaupés region, is a key indication of the region’s status as a strong linguistic area.



Map 1. Distribution of the Vaupés-Japurá (Makú) languages

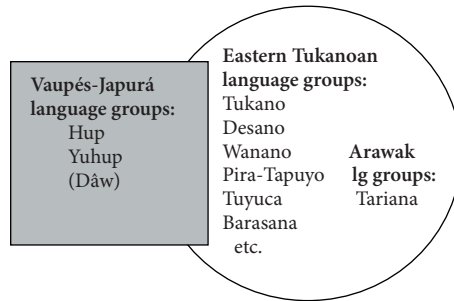
The Vaupés region is home to languages of three distinct and apparently unrelated language families. The discussion of the Vaupés linguistic area has focused on contact between two of these, Arawak and Tukanoan, which are both involved in the linguistic exogamy system. In particular, Alexandra Aikhenvald (2002, etc.) has examined in depth the contact between the Arawak language Tariana, and Tukano, the most widespread of the Eastern Tukanoan languages. However, the third language family has been essentially left out of the discussion of language contact in the region, in spite of being an integral part of the Vaupés system. This is the so-called Makú family, which I will refer to as Vaupés-Japurá (henceforth VJ);³ (see Map 1; see also Section 7 for a family tree). There are two main reasons for this family's neglect. First, and most importantly, there has been very little descriptive data available concerning these languages. Second, the speakers of the VJ languages do not take part in the linguistic exogamy system, and so they have been assumed to be only marginally involved in the language contact situation.

This paper has three main goals. First, it will contribute to our understanding of the dynamics of language contact in the Vaupés region by bringing a missing piece to the puzzle: information about the third language family involved in the contact situation. I will present data from my own research on the Hup language,⁴ one of the two VJ languages that are spoken in the Vaupés region proper. These data show that the system of evidentiality in Hup parallels the systems found in Tukano and Tariana, and has probably developed

relatively recently through contact with Tukanoan languages.⁵ The presence of the diagnostic feature of evidentiality is one important indication that even though Hup speakers do not participate in linguistic exogamy, they are nevertheless integrated into the Vaupés linguistic area. The second main goal of this paper is to provide a case study of how an evidentiality system can evolve within a given language. The discussion will address the language-internal processes of grammaticalization and semantic extension by which Hup developed its evidential markers to fit the categories it adopted via contact, and will show that Hup's use of its own linguistic resources has resulted in certain idiosyncratic extensions of the system that are interesting for the typology of evidential systems generally. Finally, while a comparison of the data from the VJ family provides strong support for the contact scenario, it also indicates that diffusion within the Vaupés linguistic area cannot account completely for the existence of grammaticalized evidentiality in Hup, because a basic evidential distinction can be reconstructed for the VJ language family. This observation leads us to reconsider Aikhenvald and Dixon's (1998: 252) proposal that evidentiality had two points of independent innovation in the northern Amazonian lowlands — in the Yanomamian and Tukanoan language families — and developed in other languages via diffusion.

2. The Hup language and the contact situation

The Hup language is spoken by around 1500 people in villages scattered between Brazil and Colombia. Living patterns have changed somewhat in the last generation, but the Hupd'əh have traditionally been forest-dwellers and semi-nomadic hunters who practice clan exogamy (cf. Reid 1979, Pozzobon 1991). The Tukanos, Tarianas, and other groups, on the other hand, live for the most part on the banks of major rivers, rely primarily on fishing and agriculture for subsistence, and practice exogamy based on linguistic/ethnic group affiliation⁶ (cf. Jackson 1983). Although — and in part because — the Hupd'əh and the river peoples exploit different ecological niches, the Hupd'əh have an extremely close socioeconomic relationship with their River Indian neighbors, especially the Tukanos; the two groups interact according to a “patron–client” system (cf. Ramos 1980: 14), which has much in common with the relationship between the Pygmies and Bantus in Africa (Turnbull 1961). Because the River Indians perceive the Hupd'əh as socially inferior, the burden is on the Hupd'əh to speak Tukano, and this has fostered a general one-sided bilingualism that appears



- Groups that interact through linguistic exogamy (obligatory marriage between language groups).
“Patrons” in a patron–client socio-economic relationship with the VJ groups.
Primary language of interaction with other language groups: Father’s language/Tukano.
- Groups that do not participate in the linguistic exogamy system.
“Clients” in a patron–client socioeconomic relationship with the Tukanoan/Arawak groups.
Primary language of interaction with other language groups: Tukano.

Figure 1. Interaction between language groups in the Vaupés region

to have been in place for many generations.⁷ In my experience, the comprehension level of Tukano among Hup adults (both male and female) is close to one hundred percent, and most are also fluent speakers. This intense language contact situation has brought about Hup’s integration into the Vaupés linguistic area, as this discussion will illustrate.

The linguistic and social interaction between Vaupés groups is summarized in Figure 1.

3. Evidentiality in Vaupés languages

Vaupés languages tend to have four evidential specifications in affirmative clauses. Table 1 (from Aikhenvald 2003a: 14) illustrates these for Tukano and Tariana:⁸ the specifications are ‘visual’, ‘nonvisual’ (sensory), ‘inferred’, and ‘reported’ (secondhand). The pattern for Tukano is like that of most of the other Eastern Tukanoan languages, although Tuyuca is noteworthy in having a further distinction between ‘apparent’ (based on direct evidence) and ‘assumed’ information (based on prior or general knowledge), where the other languages have only the inferred specification (Barnes 1990). Note that all the Tukano forms are suffixes, fusing evidentiality and tense, as well as person, number, and gender, and the Tariana specifications are enclitics that fuse evidentiality

Table 1. Evidentials, tense and person in Tukano and Tariana (from Aikhenvald 2002: 120)

	Present		Recent past		Remote past	
	Tucano	Tariana	Tucano	Tariana	Tucano	Tariana
<i>Visual</i>						
3sg.nf	-mi		-a-mi		-wī	
3sg.f	-mo	-naka	-a-mo	-ka	-wō	-na
3pl	-ma		-a-ma		-wā	
n3.p	-'		-a-pi		-wi	
<i>Nonvisual</i>						
3sg.nf	-sa-mi		-a-sī		-kātī	
3sg.f	-sa-mo	-mha	-a-sō	-mahka	-kātio	-mha-na
3pl	-sa-ma		-a-sā	(from -mha-ka)	-kātīā	
n3.p	-sa'		-a-si		-kātī	
<i>Inferred</i>						
3sg.nf			-a-pī		-pī	
3sg.f	same as nonvisual		-a-pō	-si-ka	-pō	-si-na
3pl			-a-pā		-pā	
n3.p			-a-pā		-pā	
<i>Reported</i>						
3sg.nf			-a-pi'		-pi'	
3sg.f	—	-pida	-a-po'	-pida-ka	-po'	-pida-na
3pl			-a-pa'ra		-pa'ra	
n3.p			-a-pa'ro		-pa'ro	

and tense. Other similarities between the two systems include a reduced choice of evidential specifications in non-affirmative clause types.

Work by Alexandra Aikhenvald (2002, 2003b, etc.) has shown that Tariana owes most of its evidential system to diffusion from East Tukanoan languages. In addition to the parallels between the Tariana and Tukanoan evidential systems, the argument for diffusion is supported by the lack of such a system in Tariana's close Arawak relative Baniwa, which has only an optional reportive clitic, *-pida*. More evidence that these developments are relatively recent in Tariana includes the status of Tariana's evidential markers as partially analyzable enclitics, which can attach to any focused constituent in the clause (Aikhenvald 2002: 127), unlike the verbal suffixes found in Tukano (Ramirez 1997: 119–120). In some cases these enclitics appear to have grammaticalized from existing verb stems. Finally, the Tariana evidential markers are also less fusional than the Tukano forms, and combine evidentiality with tense only, while person, number, and gender are marked by prefixes on the verb.

4. Evidentiality in Hup

Discussions of the Vaupés linguistic area have characterized areal diffusion into the VJ languages as “superficial”, in contrast to the diffusion between the Tukanoan and Tariana systems (Aikhenvald and Dixon 1998: 250; Aikhenvald 1999: 394). However, this characterization was based mainly on data from Hup’s sister-language Dâw, which is spoken in a region peripheral to the Vaupés area proper. In the case of evidentiality, the correspondence between the evidential specifications made in Hup to those in Tukano (and Tariana) is much less superficial. Unlike Dâw, which has only one optional reportive clitic, Hup distinguishes the same four basic evidentiality choices as do its Vaupés neighbors, although it does not integrate evidentiality quite as tightly into its grammar. Other clues, such as the forms of the evidential particles, support the argument that the Hup evidential system developed its complexity relatively recently, and that areal contact with Tukanoan languages (primarily Tukano) was the catalyst.

4.1 Default evidential

The most basic evidentiality specification in Hup is the default category, which primarily involves information that has been perceived visually by a speaker. This category is formally unmarked in Hup, as we see in examples (1–2).

- (1) *ʔok-nih key-ham-kèt-yiʔ-ay tih=ʔāy-ah!*
 move-NEG see-go-stand-TEL-INCH=VIS:Ø 3SG=BN:FEM-DECL
 ‘She was just standing there looking, without moving!’ (speaker witnessed event) (cv.txt)

- (2) *manga hid-an taw-nih=kah*
 Margarita 3PL-OBJ yell.at-NEG=VIS:Ø=DISJ
 ‘Margarita didn’t yell at them, actually.’ (speaker was there) (cv.txt)

The encoding of visually acquired information with a minimally marked category is common to Tukano and Tariana as well. Although these languages require formal markers, these encode only tense in the case of Tariana, while the evidential specification itself is unmarked. In Tukano, similarly, the visual evidential category is the least-marked in the paradigm. Such minimal marking of visual evidence is typical cross-linguistically, and follows from the hierarchy of evidential specifications that is common to most languages, by which visual perception takes precedence over other kinds of perception or inference

when selecting an evidential specification (typically followed, as in Hup, by other kinds of first-hand perception) (see Aikhenvald 2003a: 22).

The default evidential specification in Hup is not limited to visually acquired information, but also occurs with clauses that refer to generally known facts — in particular those that currently hold true, as in example (3), as opposed to having taken place at some point in the past.⁹ This includes descriptive discourse involving how some activity is typically carried out (example 4). Such statements are also part of the visual evidential specification in Tukano and Tariana (Ramirez 1997: 127, Aikhenvald 2002: 119).

- (3) *tīhīy sōh=deh tih ham-kami=b'ay,*
 snake flood=rain 3SG go-time.of=again=VIS:Ø
 'When the Snake-Rain (and its constellation) comes around,

nup s'ah=so?=b'ay tīhīy=d'əh ni-iy=b'ay
 this earth=LOC=again snake=PL exist-IMPF=again=VIS:Ø
 here on earth there are (many) snakes.' (txt)

- (4) *hī hid hæv-æh; nut hæv-yo? ... hid san'*
 only 3PL scrape-DECL=VIS:Ø here scrape-SEQ 3PL leaf-cone
bī?-d'o?-oh.
 make-take-DECL=VIS:Ø
 'They just scrape it; having scraped this much, they make a leaf-cone.' (txt)

Personal experience narratives also tend to use the unmarked specification — in some cases, regardless of whether the actual event was visually available to the speaker, as in example (5). This specification is also commonly used in personal experience narratives in Tukano and Tariana (Aikhenvald 2003c: 137).

- (5) *ʔāh=tæh-ʔip ʔan tih d'o?-ʔūh-ūh, yēw.*
 1SG=child-father 1SG.OBJ 3SG take-APPL-DECL=VIS:Ø armadillo
 'My husband used to catch armadillos for me.' (txt)

Because this evidential specification is unmarked, and because visually acquired and personally experienced information is extremely common in discourse, some genres of Hup speech are almost totally lacking in evidentiality marking (e.g., explanatory or descriptive discourse; spells). Moreover, speakers can sometimes leave off an evidential marker where it would otherwise be expected. This tends to occur with reportive markers in narratives, where an evidential need not always occur with every clause in a chain of closely associated clauses; this is also the case in Tariana (Aikhenvald 2003d: 310). However, while an evidential specification is optional in forming a grammatical

clause, it is nevertheless pragmatically required (whether formally marked or unmarked) in most discourse situations.

An important question regarding this so-called “default” evidential specification in Hup is whether it is best represented as a zero-marked category in its own right, relating primarily to visual evidence (as it has been treated in Tukano and Tariana), or whether it is better understood as the truly default “elsewhere” grouping, without any underlying reality as an evidential specification. The fact that its realization is not limited to visual evidence alone is not enough to give it “elsewhere” status, given the semantic extensions we find in the other (marked) evidential categories in Hup (and in those of its Vaupés neighbors), as discussed below.

This question is essentially a matter of perspective. As this discussion will demonstrate, Hup developed its evidential categories to conform to a Tukano model, and the markers of these categories emerged through a multi-stage process of grammaticalization. These markers (as their putative sources suggest) underwent semantic extensions during their development into evidentials, until they had expanded to cover large domains of meaning (e.g., from ‘heard’ information to all nonvisually acquired information, see below). Accordingly, the formally unmarked domain was shrinking as the marked domains expanded. From a language-internal point of view, the “default” evidential specification is therefore what is left of the original domain of Hup grammar that had no specification for evidentiality at all. On the other hand, this unspecified domain gradually shrank down to fit a model category of its own, the so-called “visual” category of Tukanoan. Thus from the language-external or areal point of view, this category has an underlying reality of its own. Both of these perspectives were available to Hup speakers, because their general bilingualism in Tukano ensured the co-existing everyday reality of both systems within the Hup speech community.

4.2 Nonvisual evidential

Hup marks information that is acquired firsthand (i.e., sensorily) but nonvisually with the enclitic =*hɔ̃*, as in examples (6–9). The most common source for the information is hearing, but can also be smell, taste, or touch — just as we find for the nonvisual evidentials in Tukano and Tariana (Ramirez 1997: 130–1; Aikhenvald 2003d: 296).

- (6) *nasia pæ-sīw-īy=h̄ɔ̃*
 boat go.upriver-COMPL-IMPf=NONVIS
 ‘The boat already went upriver.’ (heard but did not see it) (OS)
- (7) *pæj=h̄ɔ̃*
 umari=NONVIS
 ‘It’s umari fruit.’ (smelling mess on baby’s foot) (OS)
- (8) *k’əh naw=h̄ɔ̃!*
 sweet good=NONVIS
 ‘It’s nice and sweet!’ (tasting something) (OS)
- (9) *hu²-d’əh ni-isap=h̄ɔ̃*
 pium-PL exist-AUG=NONVIS
 ‘There are a lot of piums (small biting insects)!’ (I can feel their bites) (OS)

The nonvisual evidential is used to express one’s own personal state, such as feeling sickness (example 10), emotion, desire or other mental state (examples 11–12), a heavy burden, cold, etc. It serves a similar function in Tukano (Ramirez 1997: 134) and Tariana (Aikhenvald 2003d: 298).

- (10) *ʔāh pe²-ey=h̄ɔ̃*
 1SG sick-IMPf=NONVIS
 ‘I’m sick.’ (OS)
- (11) *sadaka² yam ʔāh wi²-tu-y=h̄ɔ̃*
 chicken song 1SG hear-want-IMPf=NONVIS
 ‘I’d like to hear the “Chicken Song”.’ (OS)
- (12) *ʔāh hipāh-nih=h̄ɔ̃*
 1SG know-NEG=NONVIS
 ‘I don’t know.’ (OS)

These uses are largely limited to the first person, although they can also occur with second person interrogatives, as in example (13). Non-first-person expressions of personal states usually require a reportive marker or epistemic modality marker (with which the inferred evidential can co-occur, see discussion below). A similar pattern occurs in Tukano and Tariana, in which the visual or inferred specifications are preferred in these cases (Aikhenvald 2002: 121; 2003d). Such skewed occurrence of evidential specifications with first-person referents, termed conjunct-disjunct marking, has been identified in many evidential systems cross-linguistically (Curnow 2002, 2003).

- (13) *yí?íyí? níŋ hipāh-nih-h̄h̄-ʔʔ?*
 this 2PL KNOW-NEG-NONVIS-INT
 ‘This, don’t you all know it?’ (txt)

Finally, the nonvisual evidential in Hup can also be extended creatively (i.e., optionally) to emphasize a personal opinion or thought, as in example (14), and can even serve to moderate a statement to make it more polite, as in example (15). This would seem to reflect an understanding or folk belief that visually acquired information is the most definite or objective type; nonvisual is therefore less objective and more moderate. Ramirez (1997: 135) reports the use of non-visual for “calculated deduction”, expressing strong epistemic sense like “must”, and Aikhenvald (p.c.) observes that both Tukano and Tariana use a nonvisual for present-tense inferences, since the inferred evidential in these languages occurs only with past tense. This use of the Hup nonvisual may therefore be a response to the Tukano model, even though Hup allows the inferred evidential to occur in the present tense.

- (14) *s’ub tog-d’əh=n’uh wid-b’ay-ay=h̄h̄ hid=b’ay*
 S’ub daughter-PL=REF.FOC arrive-return-IMPV=NONVIS 3PL=FOC
 ‘I’m thinking of S’ub’s daughters who left and never came back.’ (cv.txt)
- (15) *sim’əh=h̄h̄*
 little-NONVIS
 ‘It seems very little.’
 (when being given some cooking oil) (OS)

Hup differs from its Vaupés neighbors in other aspects of the nonvisual extensions as well. For example, Hup nonvisual =*h̄h̄* is not used to refer specifically to an act of *not* seeing something, as it is in Tukano and Tariana; instead such an act is classed in Hup as general personal experience and falls into the unmarked default category. The Hup nonvisual is also not used to describe accidental, uncontrollable actions, which are sometimes expressed with the inferred evidential, whereas Tariana and Tukano prefer the nonvisual (Aikhenvald 2003c: 138, Ramirez 1997: 133).

As the examples in this section illustrate, the nonvisual marker =*h̄h̄* — like the inferred evidential described below — tends to occur in predicate-final position. It usually occurs as enclitic to a verb, outside the core verbal aspect and tense suffixes. When no verb is present, =*h̄h̄* can appear with nouns or adjectives in predicate nominal or predicate adjective clauses. Its status as a clitic, and the fact that it is not fully obligatory, suggest its relatively recent development.

The claim that the nonvisual evidential category was borrowed into Hup from its Vaupés neighbors is supported not only by the close parallels in its uses, but also by the fact that it is lacking from the other VJ languages, with the exception of Hup's closest Vaupés relative Yuhup. However, the source of the nonvisual marker was system-internal to Hup, as this discussion will demonstrate.

The most likely source candidate is the verb stem *hɔh*- 'produce sound'. The grammaticalization of verb stem to enclitic probably came about via verb compounding, in which the final verb of the compound lost its final tense–aspect–mode suffix morphology and took on clitic status. Aikhenvald (2002:127) proposes a similar auditory source (from a compounded verb root 'hear, perceive') for the nonvisual marker in Tariana, and observes that the grammaticalization of a compounded verb is a typical process among East Tukanoan languages. This suggests that Hup's nonvisual evidential fits the Vaupés areal profile not only in its patterns of use, but also in its process of development.

Such a stage-by-stage grammaticalization scenario has precedent elsewhere in Hup, and comparison with other transitional forms can give us insight into exactly how this development may have occurred. In the first stage, the verb 'produce sound' probably was a productive compound-final form, with the meaning 'do (verb) and produce noise'. It would have been a short step from this to a more manner-related meaning, 'produce noise in doing (verb)'. Verb compounding is extremely productive in Hup, and compounded stems frequently take on modal or Aktionsart functions with varying degrees of abstraction, as we see in compounds like *?id hipāh*- (speak-know-) 'know how to speak', and *bi?-key*- (work-see-) 'try to do (something)'. Such integrated compounds involving the verb 'produce sound' are currently attested in Hup, as in example (16).

- (16) *yam-hɔh-nih-yi?* *nih!*
 sing-make.sound-NEG-ADVR be.IMP
 'Don't make (so much) noise singing!'

Through more and more frequent use (presumably motivated by the Tukano evidential model), the compound-final verb 'make noise' would have taken on an increasingly secondary status to the preceding stem until it had become an auxiliary. This process is illustrated by co-existing variants of other Hup verbs, which in some cases also undergo phonological reduction in their transition to more grammatical functions — typically involving the loss of final consonants, just as the developing evidential marker presumably lost its final *-h* (*hɔh* → *hɔ*).

For example, the stem *tuk-*, when used on its own, is a normal transitive verb meaning ‘want’. In compounds, however, it has a modal use, and its two phonological variants encode different degrees of forcefulness, as illustrated in examples (17–18); it is the full, unreduced version *-tuk-* that is used for an insisting request, while the reduced version *-tu-* is neutral. Finally, the more grammaticalized (and phonologically reduced) variant *tu-* can also indicate immediate future, as in example (19).

- (17) *sug ʔāh wiʔ-tuk-uy=h̄*
 fiddle 1SG hear-want-IMPF=NONVIS
 ‘I want to hear the fiddle!’ (emphatic) (OS)
- (18) *sug ʔāh wiʔ-tu-y=h̄*
 fiddle 1SG hear-want-IMPF=NONVIS
 ‘I’d like to hear the fiddle.’ (non-emphatic) (OS)
- (19) *deh doj-tu-y*
 water rain-‘want’-IMPF
 ‘It’s about to rain.’ (OS)¹⁰

At this stage in its existence, the verb stem ‘produce sound’ would have had two distinct realizations — one primarily lexical (as an independent verb root), and the other primarily grammatical (as a manner-related evidential). However, these would have occurred in formally identical constructions. Possibly in response to a need to differentiate these, the next stage would have involved the more grammaticalized form of the verb detaching itself from the core of the verbal construction (defined by stress and placement of the suffix), and moving to the periphery as an unstressed enclitic. At this point, the verb stem and the evidential particle would have become formally distinct.

This transition can be illustrated in the case of other Hup forms, which display flexibility in appearing as either suffixes or enclitics. One such example is the form *yāh*, which can occur either as a verb stem meaning ‘request, command’ (see example 20), or as a frustrative marker ‘in vain’ (which probably derives from the verb root). This frustrative form can itself appear either next to the verb stem as a stress-bearing suffix (example 21) or in a clitic position (example 22). When it occurs in the suffix position, the frustrative marker is formally indistinguishable from the verb. (In the examples below, underlining denotes stressed syllables.)

- (20) *deh sāy-an tih hop-yāh-āh*
 water beetle-OBJ 3SG immerse-command-DECL
 ‘He sent the water-beetle down into the water.’ (txt)
- (21) *nuw-an ?āh tuk-yāh-āh*
 this-OBJ 1SG want-FRUST-DECL
 ‘I’d like this one (but I don’t expect to get it).’ (OS)
- (22) *nuw-an ?āh tuk-uy=yāh*
 this-OBJ 1SG want-IMP=FRUST
 ‘I’d like this one (but I don’t expect to get it).’ (EL)

Less canonical uses of the evidential marker =*h̄s* appear occasionally in Hup discourse, and provide additional evidence for its grammaticalization from a compounded verb. In (23), for example, *h̄s* takes the same form as a normal compound-final verb stem (i.e., it occurs between the preceding stem and the suffix and receives stress), but it appears without the final *-h* of the verb stem ‘make noise’, and has an evidential reading.¹¹ In example (24), the form *h̄s* also has the stress and placement of a compounded verb stem, but also lacks the final *-h* and has an ambiguous reading.

- (23) *isana ma-at ni-h̄s-ṣp=?ih*
 Içana river-OBL be-Noise-SUB=BN:MSC
 ‘The person that I believe/hear is living on the river Içana...’ (txt)
- (24) *himun=hɔb d’o?-d’əh-?ay ham, yup*
 paxiuba.tree=hollow take-send-LOC.SHFT go.IMP that
nɔh-kəd-hi-h̄s-an
 fall-pass-descend-Noise-OBJ
 ‘Go fetch a paxiuba-tree-hollow, over there where (I heard) that noise of something falling.’ (txt)

Finally, the extension of the developing evidential form from the purely auditory domain to include non-auditory functions such as smell, touch, and thought also has a synchronic precedent in Hup. There is a semantic distinction in the lexicon between visual and nonvisual sensation, in which the verb form *key-* ‘see, look’ normally refers to visual perception, while the form *wi?*- is used to express both ‘hear’ and ‘understand’, and together with the incorporated noun ‘smell’ forms the compound *sih-wi?*- (smell-hear) ‘smell’.¹²

4.3 Inferred evidential

Hup speakers use the enclitic =*sud* [ʃudⁿ] to designate an inference, usually based on some form of tangible proof, as in examples (25–27). This proof is often, but not necessarily, visual evidence. In (25), for example, the husband infers from a sore on his wife's head, as well as from her illness and her story of what had happened during the day, that the evil being Curupira has sucked out her brain.

The inferential marker patterns like the nonvisual marker in that it generally cliticizes to predicates — usually verbs, but also nouns or adjectives where no verb is present. In predicate nominal and adjective clauses without a copula verb, the evidential can also occur with a focused non-predicate constituent, as in example (31) below.

- (25) *ʔam-an dohʔāy ʔunʹ-ni-iy=sud*
 2SG-OBJ Curupira suck-exist-IMP=INFR
 'Curupira has sucked you (your brain), apparently.' (txt)

- (26) *bʔoy yo-hipāh-nih=sud ʔam-ah*
 traira.fish dangle-know-NEG=INFR 2SG-DECL
 'It looks like you don't know how to carry traira fish.' (watching his
 stumbling) (txt)

- (27) *tih=doʔ sʔom-sĩwĩy=sud*
 3SG=child bathe-already=INFR
 'The child already took a bath.' (we see he's wet) (txt)

The evidence for the inference need not always be tangible, as in examples (28–29).

- (28) *ʔāh himihin-yiʔ-iy=sud*
 1SG forget-TEL-IMP=INFR
 'I forgot it, apparently.' (e.g., looking for something) (OS)

- (29) *sʔam ʔāh sʔh-ni-iy=sud=yāh, hipāh-nih=hʔ*
 yesterday 1SG dream-exist-IMP=INFR=FRUST know-NEG=NONVIS
 'I must have dreamed last night, I don't know (~I can't remember the
 dreams).' (EL)

The inferred evidential is also used to comment on an observable state, as in example (30), and can even express a creative visual comparison (31–32). A similar use is reported for Tariana (Aikhenvald, p.c.).

- (30) *nup pay=sud!*
 this bad=INFR
 ‘This looks bad/wrong.’ (child referring to a folded-over notebook page)
 (OS)
- (31) *nup=sud pati tãh-ʔip*
 this=INFR Pattie child-father
 ‘This one looks like Pattie’s husband!’ OR ‘This one, apparently, is Pattie’s husband!’ (kid pointing at picture of ugly person in book, teasing me) (OS)
- (32) *hat kæg=sud*
 alligator bone=INFR
 ‘It looks like an alligator bone (but I know that it isn’t).’
 (used in reference to a plastic hairbrush, which the speaker knows is not an alligator bone) (OS)
 But could also mean: ‘It must be an alligator bone (I think that it is).’

The inferred evidential marker can co-occur with the epistemic modality marker =ʔũh, in the partially (phonologically) fused form =suʔn’uh, as in (33). This use is preferred for expressions of inference or speculation when no evidence is on hand, or when the evidence is too vague to be very conclusive. Tukano and Tariana, in contrast, prefer the nonvisual evidential for expressing doubt or speculation, a role that it cannot play in Hup.

- (33) *hup kəwæg pog=suʔn’uh!*
 person eye big=INFR.EPIST
 ‘It must have been that big-eyed one (who ate my fruit)!’ (txt)

Aside from its occurrence in these expressions of speculation, most uses of the inferred evidential in Hup closely parallel those found in Tukano and Tariana. However, the inferred marker has an additional, functionally bizarre extension in Hup: it is used to mark a referent as deceased. Such an extension of an evidential marker is not reported in the neighboring languages and appears to be typologically rare.¹³

Use of this “deceased marker” occurs most commonly with kin terms, but is acceptable for human referents in general, and even domestic animals, as in examples (34–36). These constructions involving =*sud* differ formally as well as functionally from the usual inferred evidential constructions, in that the marker cliticizes directly to a nominal constituent in a larger clause, instead of a predicate, and it has scope over this constituent rather than over the entire predication. The possibility that these two uses of the form =*sud* (as evidential

and as deceased marker) reflect homonymy rather than polysemy cannot be ruled out, but is unlikely. In addition to the formal resemblance, the two uses have a semantic link in that they both involve situations or entities that are considered to have or have had an actual existence, which is currently relevant although not currently accessible. I will return to this point below.

- (34) *ʔin pāyh=sud=wəd* *peʔ-ni-ih*
 1PL father's.bro=DCSD=old/respected sick-be-DECL
 'Our late uncle was sick.' (txt)

- (35) *nata, naʔ-yiʔ-ḭp=ʔāy=sud*
 Natasia, die-TEL-SUB=BN:FEM=DCSD
 'Natasia, the one who died.' (EL)

- (36) *ni yaʔamboʔ=sud naw-ʔeʔ*
 1SG.POSS dog=DCSD good-PERF.NR
 'My (dead) dog was a good one.' (EL)

As in the case of the nonvisual evidential, the parallels between the inferred evidential categories in Hup and its Vaupés neighbors, coupled with the absence of this inferential specification and marker in the other VJ languages, is strong evidence for the category's diffusion into Hup. On the other hand, the form and additional functions of the Hup inferential marker show that Hup also brought its own resources to bear in developing and expanding this category.

The best source candidate for the inferred evidential is the verb stem *sud-*, 'be located inside something else'. This verb is used for animals in underground burrows or hollow trees, people in their clothes, objects inside boxes, bags, or folders, items wrapped up in something else, and so on, and generally implies that the object in question is not available for direct inspection. The grammaticalization of this form from verb stem to enclitic probably followed a similar path to that proposed above for the nonvisual marker.

The conceptual link between the three manifestations of the form *sud* — the verb 'be inside', the inferred evidential, and the 'deceased' marker — is not nearly as obvious as that between the verb 'produce sound' and the nonvisual evidential. On what basis can we claim that they are actually historically related, and not simply cases of chance homonymy? In fact, all three realizations of the form *sud*, in spite of their different functions, share a core semantic and pragmatic feature. They are all concerned with a referent that is believed to have an actual existence in some alternative 'location' — physical, temporal, or epistemological — but that is not currently accessible to direct experience. As a verb stem, *sud* expresses physical presence which is nevertheless often intangible,

and thus not completely certain — this would be especially frequent, for example, in commentary about fleeing game animals (a common topic in Hup life). It would be a relatively short conceptual step to the inferential marker, which denotes an alternative epistemological world, a possible state or event. Moreover, a locational source for an inferred evidential has precedent in a few other languages of the world, such as Wasco-Wishram (Silverstein 1978). From this point, it is not in fact a huge leap to the ‘deceased’ marker, which places its referent in an alternative temporal and metaphysical world (that of memory). These conceptual jumps are in keeping with the cross-linguistic tendency “to use vocabulary from the external (sociophysical) domain in speaking of the internal (emotional and psychological) domain,” including expressions of modality (Sweetser 1997: 49). Moreover, there is precedent elsewhere in Hup for such equations of physical and metaphysical concepts. For example, the form *mi?* can be used spatially to mean ‘under’, temporally as ‘at the same time as’, and modally as ‘in spite of’, and the form *big* can be used as the adjective ‘old’ and as a verbal marker of habitual aspect.

4.4 Reportive evidential

Hup marks secondhand, reported information with the enclitic =*mah*, as in examples (37–38). This evidential form can also be used when inquiring about or quoting someone else’s speech, as in (39). These are both functions of the reportive in Tukano and Tariana (Ramirez 1997: 141–2; Aikhenvald 2003d: 302–3).

- (37) *tih ham-teg=mah*
 3SG go-FUT=REP
 ‘He’ll go (he or another said so).’ (OS)
- (38) *‘titi? yuw-uh!’ nɔ-ɔy=mah.*
 dirty that-DECL say-IMP=REP
 “‘That one is dirty!’ he said, they say.’ (txt)
- (39) *hi?nih=mah?*
 what=REP
 ‘What did he say?’ (OS)

A reportive construction is commonly used when giving a personal name, including one’s own, as in example (40). This use is not reported for Tukano, and in Tariana only occurs in non-first-person quotative contexts (i.e., ‘he says his name is...’; Aikhenvald, p.c.).

- (40) *huy=mah* *ʔāh-āh*.
 (name)=REP 1SG-DECL
 ‘I am (called) Huy.’ (int.txt)

The reported marker in Hup is the only evidential that can occur in imperative clauses, as in example (41). This is also the case in Tukano and Tariana, where (as in Hup) it acts as a quotative and involves the repetition of commands (see Aikhenvald 2003c: 146).

- (41) *næn=mah!*
 come=REP
 ‘Come here, (they) said!’ (OS)

The reportive particle differs from the nonvisual and inferential evidential markers in Hup in its positioning and distribution. It does not occur inside subordinate clauses, unlike the other two forms. Furthermore, it can cliticize to any focused constituent of a clause, whereas the other two evidential particles cliticize to predicates. In narrative, the reportive marker is much more likely to occur in second position in the clause than on the verb, as we see in example (42). This difference in positioning sets the reportive evidential off from the others as a distinct subsystem within the larger evidential system.

- (42) *nuh-kəbək-nʌn=mah* *hid pəʔ-əh,* *dəb!*
 head-break-PL.OBJ=REP 3PL ritually.present-DECL many
 ‘They made a ritual presentation of sauva (lit. ‘head-breaker’) ants, they say, lots of them!’ (txt)

There is a diachronic explanation for this synchronic idiosyncrasy: the reportive evidential has a different history from the nonvisual and the inferred specifications. Unlike the other evidential particles in Hup, no likely source can be identified for the reportive marker — but cognate forms can be identified in at least three other Vaupés-Japura languages (a point I return to in Section 7).¹⁴ It is therefore probable that the reportive specification is the earliest evidential form in Hup, and cannot be attributed to diffusion from Tukano. Rather, the prior existence of a basic evidentiality contrast in the language may have facilitated the absorption of the other evidential categories by providing a model.

Nevertheless, its distinct history did not keep the reportive evidential immune to outside influences. Its patterns of use have undoubtedly been shaped by contact, as the parallels between the Hup reportive and those of Tukano and Tariana illustrate. Not even its form has been completely untouched. As we have already seen, Tukano has portmanteau morphemes combining evidentiality,

recent and distant past tense, and person, number, and gender. But since Hup does not mark person on the verb anyway, and formally-marked tense distinctions are minimal,¹⁵ it is no surprise that such portmanteau forms are absent in Hup. On the other hand, the reportive marker =*mah* often co-occurs with one of two optional contrast particles, one specified for recent past (=p*ah*; example 43), the other for distant past (=s'*am*, reduced to =s'*ā*[*h*] in some dialects; example 44), especially in narrative.¹⁶ In the upriver (Umari Norte) dialect, these two enclitics have become phonologically merged to create the form =*maám*, as we see in (45). This appears to reflect an incipient tense-evidential fusion in Hup, almost certainly inspired by the Tukano model.

- (43) *ni=mah=pah* *yuw-uh!*
 1SG.POSS=REP=CONTR.REC that-DECL
 'It was mine, (someone just said)!' (txt)
- (44) *yiniy=mah=s'am* *tih bi²-ih,* *hup-n'an* *tih bi²-ih*
 SO=REP=CONTR.DIST 3SG make-DECL person-PL.OBJ 3SG make-DECL
 'Thus (long ago, they say) he made (them), he made people.' (txt)
- (45) Upriver (Umari Norte) dialect:
s'ug-ut=maám *tih wɔn-kot=mah-ah*
 forest-OBL=REP. CONTR.DIST 3SG follow-go.in.circles=REP-DECL
 'In the forest, (long ago, they say), he wandered following (the tapir).' (txt)

5. The distribution of evidentials across clause types and tense-aspect-mode distinctions

In general, Hup's evidential paradigm reflects a leveling or "regularizing" effect that suggests its relative newness compared to the Tukanoan system (cf. Aikhenvald 2002: 121 for Tariana). While Hup, Tukano, and Tariana treat evidentiality similarly in declarative, imperative, and negated clauses, Hup is more permissive than its neighbors in its distribution of evidential markers across tense-aspect-mode distinctions and other types of clauses (corresponding to speech acts). For example, Hup interrogative clauses can accommodate all evidential specifications,¹⁷ whereas Tukano and Tariana reduce their system to three (minus reported). Evidentials are also reported as absent from exclamatory clauses in Tukano and Tariana (e.g., Aikhenvald 2002: 126), but are possible in Hup, as in examples (46–47).

(46) *kʰ-isap=hʒ!*
 hot-AUG=NONVIS
 ‘It’s really hot!’ (OS)

(47) *peʔ=mah! peʔ, sɔ-ɔw-ɔh!*
 power=REP power rainbow-EMPH-DECL
 ‘(He has) evil power, they say! Evil power, that rainbow (spirit)!’ (txt)

Hup is also more flexible regarding tense distinctions and evidentiality. All evidential specifications can co-occur with the future suffix in Hup, as well as with the past-contrast clitics and in clauses lacking overt tense marking (the most common case).¹⁸ Tukano and Tariana, on the other hand, are reported as not distinguishing evidentiality in the future tense at all (Aikhenvald 2002: 126; 2003: 122, etc.) — although the Tariana nonvisual marker does co-occur with the future marker in some constructions (Aikhenvald 2002: 126), and Tukano and Tariana use [evidential + present or past tense] suffixes in certain expressions of future (Ramirez 1997: 136, 166; Aikhenvald 2002: 123). Also unlike Tukano and Tariana, Hup makes an inferred distinction in the present tense. Finally, both Hup and Tariana, but not Tukano, make a reportive specification in present-tense statements.

Table 2 summarizes the distribution of evidential specifications among the representatives of the three language families in the Vaupés.

Table 2. Distribution of evidential marking across Tukano, Tariana, and Hup clause types and tense distinctions (V visual, N nonvisual, I inferred, and R reported).

	Tukano	Tariana	Hup
Declarative	V, N, I, R	V, N, I, R	N, I, R
Imperative	R; N (if action is to be performed at a distance)	R	R
Interrogative	V, N, I	V, N, I	N, I, R
Exclamatory	No evidential marking	No evidential marking	N, I, R
Negative	V, N, I, R	V, N, I, R	N, I, R
Evid. + past tense markers	V, N, I, R (portmanteau forms)	V, N, I, R (portmanteau forms)	N, I, R (tense marking optional)
Evid. + present tense markers	V, N (portmanteau forms)	V, N, R (portmanteau forms)	N, I, R (tense unmarked)
Evid + future tense markers	No evidential marking	N	N, I, R

Information on Tukano and Tariana from Aikhenvald 2002, 2003b, 2003d; Ramirez 1997.

6. Development of an additional inferred evidential

Hup has developed another inferred evidential, in addition to =*sud*: the form *-ni-* (which is obligatorily followed by additional inflectional material). Like =*sud*, *-ni-* indicates an inference about an event, where the result of the event is accessible but the speaker did not actually see it taking place, as in examples (48–50). In many cases the two forms are judged by speakers to be interchangeable, and indeed can co-occur within a single sentence, as in example (48). However, use of *-ni-* tends to place less emphasis on the actual act of inferring, in comparison with =*sud*. Also, *-ni-* can occur in interrogative clauses, but only clause-finally (unlike the other evidentials), and it is not found in imperatives. Finally, it only occurs in reference to a past event — as do all inferential distinctions in Tukano and Tariana. Hup strongly favors the use of *-ni-* over =*sud* in narrative, as in example (50).

- (48) *ʔəg-hūʔ-yiʔ-iy=sud, diʔ pā tih ʔəg-yiʔ-ni-h*
 drink-finish-TEL-IMPF=INFR remain NEG.EX 3SG drink-TEL-INFR2-DECL
 ‘He drank it all up; he drank it up and left none.’ (we see from empty pot)
 (txt)
- (49) *yup hɔtdʔah=mah hid ye-ni-ip=bʔay-ah*
 that other.side=REP 3PL enter-INFR2-SUB=again-DECL
 ‘There on the other side of it (they say), they got in again.’
 (assumption based on fact that fish has disappeared from house) (txt)
- (50) *poh, deh teg kʔet-ʔeʔ-ni-h*
 high water tree stand-PERF.NR-INFR2-DECL
 ‘Really high, the water-tree stood.’ (but is said to have later been felled to
 create the Amazonian river system) (txt)

With a first person, *-ni-* can only be used in reference to actions that the speaker has no memory of performing, usually because he or she was too young to remember, drunk, or asleep.

- (51) *nʔiʔkan ʔāh masa-ni-h, ped.*
 over.there 1SG be.born-INFR2-DECL (name)
 ‘I was born over there, Ped.’ (txt)

While the two inferential forms have a functional resemblance, *-ni-* is formally quite different from the other evidentials in Hup. It occurs only as a verbal suffix, combines directly with verb roots, and always receives the primary stress in the verb word. These features give it a much more verb-like character than the

other evidential markers, which can cliticize to nominal constituents, usually take tense–aspect suffixes between them and verb roots, and are unstressed. There is little doubt that the *-ni-* evidential derives from the verb stem *ni-* ‘be, exist’ (see below), which can itself occur as a verbal auxiliary (i.e., as the final — fully verbal — constituent in verb compounds). The restricted distribution and different patterning of *-ni-* relative to the other evidentials suggest that it has developed its use as an evidential fairly recently.

As an areal feature, the *-ni-* evidential is truly remarkable. Not only a similar evidential specification, but an almost identical form, exists in many Vaupés languages, including Tukano, Tariana, Desano (Miller 1999: 64), and Wanano (Malone 1988: 135); (see discussion in Aikhenvald 2002: 123). It also exists in Hup’s closest relative Yuhup (Ospina 2002: 181).

In Tukanoan languages, this inferred evidential reading is produced by a construction involving the auxiliary verb *nii*, ‘be’. This verb is virtually identical in form and meaning to Hup *ni-* ‘be, have’, which also occurs in the VJ languages Yuhup (my fieldnotes) and Dâw (V. Martins 1994: 154). Tariana, on the other hand, has reanalyzed the combination of the anterior aspect marker *-nhi* and past visual evidentials to create an inferred evidential (*-nhina*, *-nihka*) that closely resembles (both formally and functionally) the one found in the Tukanoan languages (Aikhenvald 2002: 123).

It is likely that the Hup inferred evidential construction is the calqued equivalent of the Tukano construction, which is built from [verb stem + nominalizer + ‘be’ + visual evidential-tense-person/number/gender] (Ramirez 1997: 140), as we see in example (52). Nominalized forms of verbs in Hup can be derived simply by stripping the otherwise obligatory suffix from the verb stem, and as we have seen the visual evidential specification is also unmarked in Hup. Thus, just as the Hup form is the semantic parallel of its Tukano counterpart, it is also essentially its formal equivalent: [verb stem + Ø + ‘be’ + Ø], with a reduced form of the verb ‘be’.

(52) Tukano:

yaa wesé ma’a wið-’karā nii-áma
 POSS field path obstruct-NOM.PL.PERF be-REC.PAST/VIS/3PL
 ‘They’ve blocked the path to my manioc field’ (proof: logs across the
 path) (Ramirez 1997:140)

Notwithstanding this partial isomorphism between Hup and Tukano, Hup already had the resources it needed for developing this construction, just as it had in the cases of the other evidential markers in its repertoire. This is indicated by the presence of the verb *ni-* ‘be, have’ in Hup and its nearby sister lan-

guages, which was probably borrowed from the Eastern Tukanoan languages at some earlier date (cf. its existence in Dâw for which no inferential evidential is reported).

7. Evidentiality in other VJ languages and its implications

The case that Hup owes its complex evidential system to diffusion depends on comparative evidence from other VJ languages. These make up a small, fairly localized family¹⁹ (see Figure 2; also Map 1 in Section 1), but they differ widely in their expression of the grammatical category of evidentiality, as can be seen in Table 3 below. Within the Vaupés linguistic area, Hup's sister language Yuhup displays a set of evidential markers that closely match those found in Hup (Ospina 2002: 181–3), and the Hup and Yuhup systems parallel those found in Tariana and the Eastern Tukanoan languages (of which Tuyuca, included below for comparison, is the most complex). Crucially, however, the VJ languages outside the Vaupés linguistic area lack most of these evidentials. Dâw, located on the periphery of the Vaupés, has only the optional reportive clitic =*mah* (S. Martins 1994: 106), and the geographically more distant language Nadëb similarly has only a reportive particle or clitic *mih* (Weir 1984: 254).²⁰

While the complexity of the Hup and Yuhup systems constitutes strong support for the robustness of evidentiality as an areal feature in the Vaupés region, the existence of an apparently cognate reportive particle across the Vaupés-Japurá language family is equally strong support for the existence of evidentiality in these languages prior to contact with Tukano. Because Nadëb, in particular, is spoken far away from the Vaupés region and its speakers have essentially no contact with speakers of other Vaupés-Japurá languages, the most likely historical scenario is that the reportive particle in these languages was present in their common ancestor. Moreover, since the speakers of the Eastern Tukanoan languages are thought to have entered the Vaupés region from the

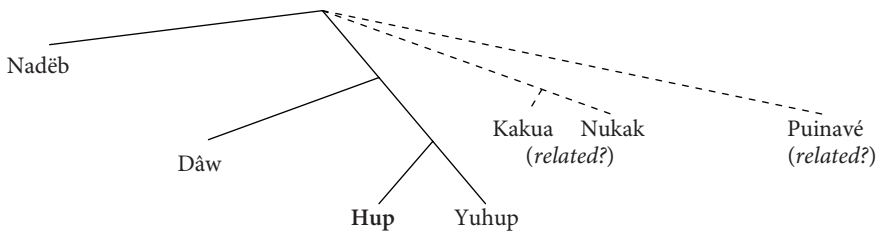


Figure 2. The Vaupés-Japurá (Makú) language family.

west within the last 500–1000 years (cf. Nimuendajú 1982; Aikhenvald 1996: 79–80),²¹ and Nadëb is too different from the other VJ languages for such a recent split date to be likely, it is improbable that the presence of this particular evidentiality specification in the VJ family can be attributed to direct diffusion from Tukanoan languages.

These observations may have crucial implications for our understanding of the development of evidentiality in northern Amazonia²² in general. Aikhenvald and Dixon (1998: 252) propose that evidentiality may have been separately innovated in the northern Amazon basin (i.e., excluding the Andean region) in at least two language families, Yanomamian and Tukanoan, each of which has a basic reconstructible evidentiality distinction.²³ The other languages in the region with evidentiality — North Arawak, Bora (Witotoan), and Arabela (Zaparoan) — are considered to have acquired it via diffusion from Tukanoan languages. However, Aikhenvald and Dixon (1998: 252) also claim that “there is no doubt” that the VJ language Dâw “developed an evidential marker [reportive =*mah*] under areal influence from Tukanoan.”

In fact, as the cognate forms for the reportive particle show, a basic evidentiality distinction (presumably between reported information and everything else) is as likely to have been independently innovated in Dâw and its sister languages as it was in the Yanomamian and Tukanoan languages. This leaves us with two possible conclusions. Either evidentiality had at least *three* points of independent innovation in families located in the same general region — and possibly more if we consider Cariban and the languages of the Andean region — or the existence of evidentiality in the various proto-languages can itself be attributed in part to diffusion.²⁴ Given what we know about grammaticalized evidentiality — particularly its strong areal tendencies (e.g., Aikhenvald 2003b, de Haan, to appear) — the possibility of so many points of independent innovation seems relatively unlikely. Rather, I propose that some of the several reconstructible evidentiality systems in the region may actually reflect a much earlier language contact situation, one that existed long before the current one.

An ancient language area that included two or more of these families' parent languages is not improbable. The high mobility of Amazonian groups generally is reflected in the widely dispersed geographical distribution of many language families throughout the Amazonian region, such as Arawak and Tupi. Widespread trading networks are also known to have existed in northwest Amazonia and adjacent regions prior to the Spanish and Portuguese invasions (e.g., Arvelo-Jiménez and Bjord 1994; Porro 1994).

Table 3. Evidentiality in languages of the Vaupés region and the VJ family

Languages of the Vaupés region			VJ languages outside the Vaupés region proper			
	Tukano Tuyuca (among other E. Tuk. lgs) (<i>East Tukanoan</i>)	Tariana (<i>Arawak</i>)	Hup Yuhup (<i>Vaupés-Japurá</i>)	Dâw	Nadëb	
Visual	paradigm (evid.-person-tense-number)	paradigm (evid.-person-tense-number)	paradigm (evid.-tense)	(-ø)		
Nonvisual	paradigm (evid.-person-tense-number)	paradigm (evid.-person-tense-number)	paradigm (evid.-tense)	=hɔ̃	=hɔ̃	
Inference	paradigm (evid.-person-tense-number)	paradigm (evid.-person-tense-number)	paradigm (evid.-tense)	=sud		
Inference2	-nii construction	paradigm (evid.-person-tense-number)	-nhina, -nihka	-ni- constr.	-ni constr.	
Reported	paradigm (evid.-person-tense-number)	paradigm (evid.-person-tense-number)	paradigm (evid.-tense)	=mah	=mah	=mah mih

Sources: Tukano: Ramirez 1997; Tuyuca: Barnes 1990; Tariana: Aikhenvald 2003a, etc.; Yuhup: Ospina 2002:181; Dâw: S. Martins 1994:106; Nadëb: Weir 1984:254.

Since features that are easily spread via diffusion may also be easily lost, identifying an ancient linguistic area may be very difficult indeed. However, if more typologically uncommon features (“quirks”) like evidentiality — especially those that may be truly diagnostic of language contact — can be reconstructed for the same set of otherwise distinct proto-languages, we will have more evidence on which to base such a claim. Conversely, further cross-linguistic studies of evidentiality systems may give us more insight into the probability of their independent innovation, as well as their likelihood of diffusing relative to other diagnostic features.

8. Conclusions

The grammatical category of evidentiality straddles the realms of grammar and of discourse. It is a structural linguistic feature that relates directly to cultural expectations regarding politeness and responsibility for information (cf. Beier et al. 2002). It is undoubtedly its high salience in discourse that makes evidentiality particularly prone to diffusion among languages, as we have seen illustrated in the near mirror-images of evidential distinctions in the Vaupés languages.

The dynamics of areal diffusion of features like evidentiality may have much to teach us. The standard model of language contact presents a linguistic feature as fully formed in one language before it is borrowed into another. However, places like the multilingual Vaupés challenge our conceptions about language boundaries. In such situations, where speech communities overlap and intersect with language groups and even language families, a linguistic innovation — like the development of the *ni* inferred evidential construction — could be expected to spread not simply from language to language, but rather from one multilingual speech community to another.

The evidential system in Hup provides us with a piece of the puzzle for understanding the Vaupés linguistic area, and illustrates the robustness of evidentiality as a diagnostic feature of such a linguistic area. In Hup, we can unravel the story of an evidential system's historical development, beginning with one optional evidential distinction and ending with a full-fledged, grammaticalized system of as many as five specifications. Areal influences were the catalyst for the transition and lent the blueprints for the new system, but language-specific processes of grammaticalization and semantic extension provided its nuts and bolts. Finally, the comparison of evidentiality in Hup and its sister languages suggests that the current Vaupés linguistic area is only one of many possible combinations of languages and language families that may have interacted in the past, and that evidentiality may be a diagnostic feature of such ancient linguistic areas as well as of contemporary ones.

Example source coding

(cv.txt) = conversation text

(EL)= elicitation contexts

(int.txt) = interview text

(OS)= overheard speech (doublechecked later with consultants)

(txt) = narrative text

Interlinear abbreviations

ADVR – Adverbializer	INFR2 – Inferential evidential 2
APPL – Applicative	INT – Interrogative
AUG – Augmentative	LOC – Locative
BN:FEM – Bound noun: feminine	LOC.SFT – Locative shift
BN:MSC – Bound noun: masculine	NEG – Negative
COMPL – Completive	NEG.EX – Negative existence
CONTR.DIST – Distant past contrast	NONVIS – Nonvisual evidential
CONTR.REC – Recent past contrast	POSS – Possessive
DECL – Declarative clause marker	PERF – Perfective
DISJ – Disjunctive	PERF.NR – Perfective, nominalized form
DSCD – “Deceased” marker	OBJ – Object
EMPH – Emphasis	OBL – Oblique
EPIST – Epistemic modality	PL – Plural
FOC – Focus	REF.FOC – Referential focus
FRUST – frustrative	REP – Reportive evidential
IMP – Imperative	SEQ – Sequential
IMPF – Imperfective	SG – Singular
INCH – Inchoative	SUB – Subordinator
INFR – Inferential evidential	TEL – Telicity

Notes

* This material is based upon work supported by a Fulbright-Hays Dissertation Research Grant, by the National Science Foundation under Grant no. 0111550, and by the Max Planck Institute for Evolutionary Anthropology. In Brazil, the Museu Paraense Émilio Goeldi and the Instituto Socioambiental provided considerable assistance with practical issues relating to fieldwork, which is gratefully acknowledged. I owe my thanks to the Hupd’əh for teaching me their language, and I would also like to thank Alexandra Aikhenvald, Bernard Comrie, Orin Gensler, Don Stilo, and two anonymous reviewers for their many helpful comments on this paper.

1. It is possible that the occurrence of evidentials is somewhat more common cross-linguistically, but that they have been overlooked in a significant number of grammatical descriptions. If this turns out to be the case, then the degree of evidentiality’s usefulness as a diagnostic feature of language contact may require reevaluation.

2. An association between the Andean languages and those of Amazonia is more tenuous. Little or no evidence of linguistic diffusion between these regions has been put forward to date, and Aikhenvald and Dixon (1998: 253) maintain that the Andean area shows a “significantly different typological profile from the Amazonian area”. However, there is archaeological and other evidence of interaction and economic exchange between the Andes and the

Amazonian lowlands (e.g., Lathrap 1973, Kurella 1998, Taylor 1999), which suggests that the question of linguistic interchange between these regions awaits further evaluation.

3. I prefer to use the name 'Vaupés-Japurá' (originally suggested by Henri Ramirez) for two reasons: (1) There is some confusion surrounding the name 'Maku', which occurs in the literature in reference to several unrelated language groups in Amazonia (for example, Campbell 1997: 183 lists the VJ languages under 'Puinavean (Makú stock)', and refers to an extinct 'Maku' language family); (2) The name 'Maku' (probably from Arawak *ma-aaku* [NEG-talk] 'do not talk') is widely recognized in the Vaupés region as an ethnic slur, directed against the members of this ethnic/linguistic group. Vaupés and Japurá are the names of two rivers that delineate the general area in which these languages are spoken.

4. Also known as Hupda or Jupda. Speakers call their language 'Hup' (= 'human'); Hupd'əh is the plural ethnonym, lit. 'people'. This research is based on about 13 months of fieldwork (of which 10 were basically consecutive), conducted in Hup villages along the Tiquié river.

5. The relationship between Tariana and Hup is thus a fascinating example of a case where two languages have come to resemble each other through contact with an intermediary. A similar situation has been described by Tosco (2000: 359) for Ethiopia, where several different non-Semitic languages are undergoing a shift toward Tigre. (Thanks to Alexandra Aikhenvald for bringing this to my attention.)

6. This relationship is subject to additional rules which further restrict marriage between these groups; for example, the Tukano cannot marry the Bará (see Jackson 1983, Aikhenvald 2002: 22). River Indian men do occasionally marry Hup women, but the reverse is unheard of — a fact which reflects the patrilineal River Indian social structure and the social inequality between these two groups.

7. A related situation is found among the Romani peoples, social "outcasts" in European society who have maintained a similar long-term, pervasive, and one-sided bilingualism that has brought about extensive structural shift in their language (Matras 2002).

8. The four-term Tariana system is typical among older, more traditional speakers (Aikhenvald, p.c.).

9. Reference to on-going events that are not accessible to direct experience, such as habits of peoples in other lands, usually involves the reported evidential.

10. A similar construction occurs in both Tukano and Tariana (Aikhenvald, p.c.).

11. However, this evidence is weakened by the fact that a few clitics in Hup did not derive originally from verbs, and also show this positional flexibility. For example, the pair of future-tense suffixes *-teg* and *-te* can be demonstrated to have grammaticalized from a noun. Nevertheless, these are exceptions to the general rule in Hup, since most clitics for which a path of grammaticalization can be proposed probably have verbal origins.

12. These extensions are likewise found in Tukano and Tariana (Aikhenvald, p.c.), and may represent yet another areal feature.

13. One of the lexical uses of the unproductive inferred evidential in the Bolivian isolate Mosestén (Sakel 2002: 268) is to refer to an unwitnessed death upon finding the body. However, this is a reflection of a true verbal evidential use (by which the speaker infers an event upon seeing the evidence), rather than a ‘deceased marker’ like the one found in Hup.
14. Aikhenvald (2003d; p.c.) notes that the reported evidential in Tariana behaves differently from the other evidential forms. It is also the only ‘archaic’ one in the system, and has cognates with reported evidential forms in other Arawak languages.
15. With the exception of the future specification (which is itself optional in certain contexts), tense information in Hup is usually communicated via aspectual distinctions.
16. The order of evidential + contrast/tense marker is fixed; apparently, the tense semantics can refer *either* to the time of the report, or to the time of the event.
17. These usually reflect the information source of the hearer, rather than the speaker.
18. Aspect, rather than tense, is more likely to be overtly marked on Hup verbs. In general, evidential markers combine freely with aspectual inflection.
19. The relationship of the languages Kakua (also known as Bará), Nukak, and Puinavé (all spoken in Colombia) to the rest of the family is still in doubt, and very little descriptive data from these languages is currently available. The exact nature of Nadëb’s relationship to other VJ languages also has yet to be fully defined, but the available data indicate a clear connection (based on numerous cognates exhibiting regular sound changes).
20. Note that the claim made by Martins and Martins (1999: 261), that evidentiality is lacking in all VJ languages except Dâw, is obviously incorrect.
21. But note that this historical scenario is somewhat speculative and awaits further clarification.
22. Northern Amazonia is here defined as the region encompassing the lowlands of the Amazon river watershed north of the Amazon/Solimões itself, covering parts of Brazil, Colombia, Ecuador, Venezuela, Suriname, and the Guianas.
23. Yanomamian languages near the Vaupés region (Xamatauteri and Sanuma), like the VJ languages, reportedly have several evidentiality specifications, but the most distant language (Yanam) has only a single distinction (eyewitness — non-eyewitness) which, according to Aikhenvald and Dixon (1998: 247), probably cannot be attributed to diffusion from Tukanoan languages. Aikhenvald and Dixon do not discuss evidentiality in Cariban languages, which also seems unlikely to be a result of diffusion from Tukanoan.
24. One alternative possibility is that the reconstructed form could have been only an *incipient* evidential at the time of the VJ family’s split, and might have gained full evidential status later in the respective VJ languages through contact with various neighboring Tukanoan and Arawak languages (thanks to an anonymous reviewer for suggesting this possibility).

References

- Aikhenvald, Alexandra. 1996. "Areal diffusion in northwest Amazonia: the case of Tariana". *Anthropological Linguistics* 38:1, 73–116.
- Aikhenvald, Alexandra. 1999. "Areal diffusion and language contact in the Içana-Vaupés basin, north-west Amazonia". In: Dixon and Aikhenvald (eds.), *The Amazonian languages*, 385–416. Cambridge: Cambridge University Press.
- Aikhenvald, Alexandra. 2001. "Areal diffusion, genetic inheritance, and problems of sub-grouping: a North Arawak case study". In: Dixon, R.M.W. and Alexandra Aikhenvald (eds.), *Genetic inheritance and areal diffusion: problems in comparative linguistics*, 167–194. Oxford: Oxford University Press.
- Aikhenvald, Alexandra. 2002. *Language contact in Amazonia*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra. 2003a. "Mechanisms of change in areal diffusion: new morphology and language contact". *Journal of Linguistics* 39:1–29.
- Aikhenvald, Alexandra. 2003b. "Evidentiality in typological perspective". In: Aikhenvald, Alexandra and R.M.W. Dixon (eds.), *Studies in evidentiality*, 1–31. Amsterdam/Philadelphia: John Benjamins.
- Aikhenvald, Alexandra. 2003c. "Evidentiality in Tariana". In: Aikhenvald, Alexandra and R.M.W. Dixon (eds.), *Studies in evidentiality*, 131–164. Amsterdam/Philadelphia: John Benjamins.
- Aikhenvald, Alexandra. 2003d. *A grammar of Tariana*. Cambridge: Cambridge University Press.
- Aikhenvald, Alexandra. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra and R.M.W. Dixon. 1998. "Evidentials and areal typology: a case study from Amazonia". *Language Sciences* 20:241–257.
- Arvelo-Jiménez, Nelly and Horacio Biord. 1994. "The impact of conquest on contemporary indigenous peoples of the Guiana shield: the system of Orinoco regional interdependence". In: Anna Roosevelt (ed.), *Amazonian Indians from prehistory to the present: anthropological perspectives*, 55–78. Tucson: University of Arizona Press.
- Barnes, Janet. 1990. "Classifiers in Tuyuca". In: Doris L. Payne (ed.), *Amazonian linguistics: studies in lowland South American Indian languages*, 273–292. Austin: University of Texas Press.
- Beier, Christine, Lev Michael, and Joel Sherzer. 2002. "Discourse forms and processes in indigenous lowland South America: an areal-typological perspective". *Annual Review of Anthropology* 31:121–45.
- Campbell, Lyle. 1997. *American Indian languages*. New York: Oxford University Press.
- Campbell, Lyle, Terence Kaufman, and Thomas C. Smith-Stark. 1986. "Meso-America as a linguistic area". *Language* 62:530–570.
- Chafe, W. and Joanna Nichols. 1986. *Evidentiality: the linguistic coding of epistemology*. Norwood, NJ: Ablex.
- Curnow, T. J. 2002. "Conjunct/disjunct marking in Awa Pit". *Linguistics* 40: 611–27.
- Curnow, T. J. 2003. "Types of interaction between evidentials and first person subjects". *Anthropological Linguistics* 44. 178–96.

- Gómez-Imbert, Elsa. 1996. "When animals become 'rounded' and 'feminine': conceptual categories and linguistic classification in a multilingual setting". In: Gumperz, John and Stephen Levinson (eds.), *Rethinking linguistic relativity*, 438–469. Cambridge: Cambridge University Press.
- Gumperz, John and Stephen Levinson. 1996. *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.
- Haan, Ferdinand de. 2005. "Evidentiality". In: M. Dryer, M. Haspelmath, D. Gil and B. Comrie (eds.), *World atlas of language structures*. Oxford: Oxford University Press.
- Hardman, M. J. 1986. "Data-source marking in the Jaqi languages". In: Chafe and Nichols (eds.), *Evidentiality: the linguistic coding of epistemology*, 113–36. Norwood, NJ: Ablex.
- Hill, J.D. and F. Santos-Granero (eds.). 2002. *Comparative Arawakan histories. rethinking language family and culture area in Amazonia*. Urbana/Chicago: University of Illinois Press.
- Jackson, Jean. 1983. *The fish people: linguistic exogamy and Tukanoan identity in northwest Amazonia*. Cambridge: Cambridge University Press.
- Kurella, Doris. 1998. "The Muisca: Chiefdoms in transition." In E.M. Redmond (ed.), *Chiefdoms and chieftaincy in the Americas*, p.189–216. Gainesville: University Press of Florida.
- Lathrap, Donald W. 1973. "The antiquity and importance of long-distance trade relationships in the moist tropics of Pre-Columbian South America". *World Archaeology* 5:2.
- Malone, Terrell. 1988. "The origin and development of Tuyuca evidentials". *International Journal of American Linguistics*, 54:119–40. Chicago: University of Chicago Press.
- Martins, Silvana A. 1994. *Análise da morfossintaxe da língua Dâw (Makú-Kamã) e sua classificação tipológica*. Master's thesis, UFSC, Santa Catarina.
- Martins, Silvana A. and V. Martins. 1999. "Makú". In: Dixon, R.M.W. and Alexandra Aikhenvald (eds.), *The Amazonian languages*, 251–267. Cambridge: Cambridge University Press.
- Martins, Valteir. 1994. *Análise prosódica da língua Dâw (Makú-Kamã) numa perspectiva não-linear*. Master's thesis, UFSC, Santa Catarina.
- Mattei-Muller, Marie-Claude. 2004. "Voir et savoir en Panaré". In: Landaburu, Jon and Zlatka Guentcheva (eds.), *L'Énonciation médiatisée et modalité épistémique*, Paris: CNRS (Centre National de la Recherche Scientifique).
- Matras, Yaron. 2002. *Romani: a linguistic introduction*. Cambridge: Cambridge University Press.
- Meira, Sergio. 2000. *A reconstruction of Proto-Taranoan: phonology and morphology*. Meunchen: Lincom Europa.
- Miller, Marion. 1999. *Desano grammar*. Arlington: SIL and the University of Texas at Arlington [*Studies in the Languages of Colombia* 6].
- Nimuendajú, Curt. 1982. *Textos indigenistas*, no. 8:1–39. São Paulo: Ed. Loyola.
- Ospina Bozzi, Ana Maria. 2002. *Les structures élémentaires du Yuhup Makú, langue de la Amazonie Colombienne: morphologie et syntaxe*. Ph.D dissertation, Université Paris 7 — Denis Diderot.
- Payne, David. 1990. "Some widespread forms in South American languages". In: Doris Payne (ed.), *Amazonian linguistics*, 75–87. Austin: University of Texas Press.

- Payne, Doris. 1990. "Morphological characteristics of lowland South American languages". In: Doris Payne (ed.), *Amazonian linguistics*, 75–87. Austin: University of Texas Press.
- Payne, Doris. 2001. "Review of *The Amazonian Languages*, ed. by Dixon and Aikhenvald". *Language* 77, 3:594–598.
- Porro, Antonio. 1994. "Social organization and political power in the Amazon floodplain: the ethnohistorical sources". In: Anna Roosevelt (ed.), *Amazonian Indians from prehistory to the present: anthropological perspectives*, 79–94. Tucson: University of Arizona Press.
- Pozzobon, Jorge. 1991. *Parenté et démographie chez les Indiens Makú*. Ph.D dissertation, Université de Paris VII.
- Ramirez, Henri. 1997. *A fala Tukano dos Ye'pa-Masa*. Inspeção Salesiana Missionária da Amazônia, CEDEM: Manaus.
- Ramos, Alcida Rita. 1980. "Introduction". In: Ramos, Alcida Rita, *Hierarquia e Simbiose: Relações Intertribais no Brasil*, 1–17. São Paulo: Editora Hucitec.
- Reid, Howard. 1979. *Some aspects of movement, growth and change among the Hupdu Maku Indians of Brazil*. Ph.D dissertation, University of Cambridge.
- Roosevelt, Anna. 1994. *Amazonian Indians from prehistory to the present: anthropological perspectives*. Tucson: University of Arizona Press.
- Sakel, Jeanette. 2002. *A grammar of Mose'tén*. Ph.D dissertation, University of Nijmegen, The Netherlands.
- Sherzer, Joel. 1976. *An area-typological study of American Indian languages north of Mexico*. New York: North Holland. [*North Holland Linguistic Series*, 20].
- Silverstein, Michael. 1978. "Deixis and deducibility in a Wasco-Wishram passive of evidence". *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 4:238–54.
- Sorensen, Arthur P. 1967. "Multilingualism in the northwest Amazon". *American Anthropologist* 69:670–84.
- Sweetser, Eve. 1997. *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Taylor, Anne Christine. 1999. "The western margins of Amazonia from the early sixteenth to the early nineteenth century." In F. Salomon & S.B. Schwartz (eds.), *The Cambridge history of the native peoples of the Americas*, Volume III: South America, Part 2, p.188–256. Cambridge: Cambridge University Press.
- Tosco, Mauro. 2000. "Is there an 'Ethiopian language area'?" *Anthropological Linguistics* 42: 329–65.
- Turnbull, Colin. 1961. *The forest people*. New York: Simon and Schuster.
- Weber, D. J. 1986. "Information perspective, profile, and patterns in Quechua". In: Chafe and Nichols (eds.), *Evidentiality: the linguistic coding of epistemology*, 137–55. Norwood, NJ: Ablex.
- Weir, E. M. Helen. 1984. *A negação e outros tópicos da gramática Nadëb*. Master's thesis, UNICAMP, Campinas.

Author's address:

Patience Epps
Department of Linguistics
Max Planck Institute for Evolutionary Anthropology
6 Deutscher Platz
D-04103 Leipzig, Germany
epps@eva.mpg.de