EVIDENTIALS AND AREAL TYPOLOGY: A CASE STUDY FROM AMAZONIA

ALEXANDRA Y. AIKHENVALD and R. M. W. DIXON

1. Amazonian languages and areal linguistics

It is noteworthy that, in spite of the great genetic diversity of Amazonian languages, they all share certain structural features. There are many indications that Amazonian languages constitute a linguistic area. Languages of the same family spoken in different geographical regions typically show more structural similarities to their neighbours than to their genetic relatives.

Such is the case with the various Arawak subgroups, from the geographically most extended family in South America (Aikhenvald forthcoming-b). Unusual grammatical or phonological phenomena found in one subgroup but not in another are almost always explainable as the result of diffusion from neighbouring and not necessarily related languages.

A useful definition of ‘linguistic area’ (which is especially relevant for languages of the Americas) comes from Sherzer (1973: 760):

A linguistic area is defined here as an area in which several linguistic traits are shared by the languages of the area and furthermore, there is evidence (linguistic and non-linguistic) that contact between speakers of the languages contributed to the spread and/or retention of these traits and thereby to a certain degree of linguistic uniformity within the area. It is important to remember that languages which are unrelated or distantly related may very well and probably do disagree with regard to many traits and yet still be in the same linguistic area according to the above definition, since they share several traits (which one might want to call diagnostic traits).

It is often difficult to establish whether similarities between contiguous languages which belong to distinct subgroups of the same family are due to common linguistic origin or to diffusion. It is also difficult to make decisions concerning areal features among languages which display typological similarities. A very careful analysis of the grammars of individual languages and subgroups is needed, to proceed successfully with the analysis of areal diffusion patterns.

Properties shared by all, or almost all, Amazonian languages include:

(a) The majority of languages are polysynthetic and head marking; agglutinating with little fusion.

(b) There is typically one liquid phoneme, which is frequently a flap. There are usually more affricates than fricatives. The high unrounded central vowel \( i \) is frequent.

Correspondence relating to this paper should be addressed to A. Y. Aikhenvald, Research Centre for Linguistic Typology, Australian National University, Canberra, ACT 0200, Australia.
A typical Amazonian vowel system has five members: \( i, e, a, i, u/o \). There is typically contrastive nasalisation of vowels.

(c) Many languages have extensive classifier and/or gender systems. Gender assignment is often semantically transparent, and is not overtly marked on the head noun.

(d) There are very few oblique cases—often just a locative and an instrumental/comitative.

(e) Possession (either alienable or inalienable) is typically marked on the possessed noun, not on the possessor; the most widespread word order is possessor-possessed (e.g. ‘John his-canoe’).

(f) Often, just one core argument is cross-referenced on the verb. There may be different bound pronominal paradigms depending on which core argument is being cross-referenced in each particular instance.

(g) The rules for which core argument is cross-referenced can be complex (relating to the meaning of the verb, clause type, etc.) often giving rise to a ‘split-ergative’ system. Fully accusative systems of marking for predicate arguments are rarely encountered.

(h) The bound pronominal forms marking a possessor within an NP are typically the same as one of the bound pronominal paradigms for marking core arguments of a clause (sometimes the same as the A or \( A/S_a \) series, othertimes the same as the O or \( O/S_o \) series).

(i) Most (although not all) languages have prefixes; there are typically fewer prefix than suffix positions.

(j) If there are several prefix positions, the bound pronominal prefix(es) will typically appear further from the root than prefixes that mark valency-changing derivations (e.g. causative, applicative).

(k) Most verbal categories (e.g. tense, aspect, modality, direction) are expressed through optional suffixes.

(l) Subordinate clauses typically involve nominalised verbs, with the type of subordination being marked on the verb.

(m) If there is noun incorporation, typically only those nouns which are obligatorily possessed can be incorporated, and they typically precede the verb root.

(n) In many languages adverbs and adpositions may be incorporated into the verb, typically following the verb root.

(o) There is generally only a small class of lexical numbers.

It is interesting to compare typological characteristics of the Amazonian linguistic area, in lowlands South America, with those of the Andean linguistic area in the adjacent mountains, which comprises the Quechua and Aymara families (see the map, Fig. 1). The Andean area is clearly different in almost all of the characteristics just listed. (b) There are two or three liquids; fricatives rather than affricates; and a three vowel system \( i, a \) and \( u \) with no contrastive nasalisation. (c–e) There are no classifier or gender systems; there is an extensive set of case markers; possession is marked both on possessor and on possessed. (f–h) Two core arguments are marked on the verb, in an entirely accusative system; bound pronominal markers of possession show some similarity to, but are not identical with, the forms marking core arguments on the verb. (i–k) There are no prefixes; and there is an obligatory suffixal system for tense and aspect. (l) Subordination does not involve nominalisation. (m–n) There is no incorporation of nouns, adverbs or prepositions (Cole, 1982: 161 mentions incorporation
Fig. 1. Rough locations of languages and language families with evidentials.
within nominalisation, but this is just a type of compounding, as in English). (o) There is a full set of lexical numbers. As regards characteristic (a), Andean languages are synthetic, and combine head and dependent marking; they are basically agglutinating with some fusion (subject, object and tense suffixes to the verb may be fused).

There are a number of grammatical properties which are not shared by all Amazonian languages but are found in the languages in certain regions and help to define these as linguistic subareas within a wider linguistic area. These properties include tone, serial verbs, switch-reference, and evidentiality. Their distribution can be important for hypothesising whether certain features were more widespread at some time in the past, and are being progressively lost, or whether they are recent innovations that are currently diffusing more and more widely.

In a given linguistic area not all shared features have the same significance as indicators of areal diffusion. Unusual features (which occur in only a small proportion of languages of the world) count for more than mundane features which have a wide general distribution. An unusual feature such as evidentiality is more likely to have diffused—and thus to be clearly criteria of a linguistic area—than anything as common as, say, a system of two tenses, past and non-past, which neighbouring languages may well develop independently. See Campbell et al. (1986: 535-6) who also comment that since 'meaningful linguistic areas are the historical product of linguistic diffusion, the stronger linguistic areas are those whose shared traits can be shown to be diffused—and cannot be ascribed to a common ancestor, to chance, or to universals'.

The distribution of each 'bundle' of areal features—together with data on language history and reconstruction—is crucial for determining the boundaries of linguistic subareas, and how these fit into the overall picture of the whole of Amazonia as one large linguistic area.

2. Evidentiality as a diagnostic feature

Evidentiality refers to a grammatical system whereby for every statement of a certain type made in the language, the evidence on which it is based must be stated. One of the classic systems is that in Tuyuca from the Tucanoan family (in north-west Amazonia) described by Barnes (1984). If a Tuyuca speaker wishes to say 'he played soccer' they must add an appropriate evidential suffix to the verb, according to whether they saw him play (visual or eyewitness evidential); or just heard the game (non-visual evidential); or saw some evidence that he had played, e.g. his footprint on the sports field (apparent evidential); or were told by someone else that he had played (secondhand or reported evidential); or assumed that he had played since he is known to play in the soccer match each week (assumed evidential).

Many evidentiality systems have a smaller number of choices than Tuyuca. The minimum, of course, is two, typically eyewitness and non-eyewitness. (See the collection of papers edited by Chafe and Nichols, 1986.) Every language does of course have the capability of providing information about the evidence on which a statement is based (e.g. They say... or I guess... or I saw... in English—see Chafe, 1986). But the existence of an obligatory morphological system of evidentiality choices—in certain tense(s) and/or certain clause types, etc.—is uncommon, and where it applies to all or most of the languages in a certain geographical area it will be a significant diagnostic for those languages constituting a linguistic area.
The greatest concentration of languages with evidentiality systems is in the Americas. In North America, evidentiality systems are found in languages from at least the Wakashan, Wintuan, Iroquoian, Yuman and Pomoan families (see the papers in Chafe and Nichols, 1986). In the Andes region they occur in languages from the Quechua and Aymara families (see Section 5 below). Their occurrence in Amazonian languages, from lowland South America, will be detailed in Section 3 below.

Evidentials are only found in a few, scattered language outside the Americas. These include some Tibeto-Burman languages—including Sherpa (Woodbury, 1986), Tibetan (DeLancey, 1986), and Akha (Thurgood, 1986). Evidentials were the only verb inflection in a Chinese–Russian pidgin (Nichols, 1986). There is an incipient system in Balkan languages (Bulgarian, Macedonian and Albanian) and also in Turkish. Among Uralic languages, Estonian has a special verbal form used to refer to a reported action (Laanest, 1975: 155–6). Languages of the Permian subgroup of Uralic have a system of eyewitness/non-eyewitness evidentials in past tense (Tepljashina and Lytkin, 1976: 177). Classical Japanese had an evidentiality system (M. Shibatani, personal communication) but only a few traces remain in the modern language (Aoki, 1986). Evidentials are not found elsewhere in Indo-European, Uralic, Afroasiatic, Caucasian or Altaic languages.

The next section looks at the structure and distribution of systems of evidentials in Amazonian languages. The interaction of areal and genetic properties of evidentials is considered in Section 4, and some conclusions are presented in Section 5.

3. Types of evidentials in Amazonian languages

In lowland South America there are two regions\(^5\) characterised by the occurrence of evidentials. Section 3.1 describes what we can call a region of ‘weak evidentiality’ in southern Amazonia, where there are small systems with two of three evidential values in some (but not all) of the languages; there is just one language with a four-term system. Then Section 3.2 deals with a region in northern Amazonia characterised by what we can call ‘strong evidentiality’; here almost every language shows an evidentiality system, and these often have four or five terms.

3.1. Evidentials in southern Amazonia

Evidentiality systems found in the languages of southern Amazonia are relatively simple, being mainly based on the binary oppositions ±eyewitness and ±reported.

Three kinds of systems are attested: (1) ±reported; (2) ±eyewitness with an optional marking for reported; and (3) a four-term system of eyewitness, inferred, reported and gut feeling. But note that evidentiality systems appear only to occur in some of the languages from the southern region shown on the map. (However, for many of the languages in the region covered by the map there is as yet no adequate grammatical description and we simply do not know whether or not they have evidentials.)

(i) A REPORTED/NON-REPORTED system is found in a number of South Arawak languages such as Terêna, Ignaciano, Waurá, Pareci and Piro. It is also present in languages from the Panoan family (Loos forthcoming). Traces of the system are found in some Tupi-Guaraní languages—Kamaiurá, spoken on the Upper Xingu River (Seki forthcoming), distinguishes reported and non-reported in the past tense. Evidentiality
marking in these languages appears to be of relatively recent origin, for the following reasons:

(a) Evidentiality is marked by particles or clitics rather than affixes; these are not fused with any other morpheme.

(b) Different markers are used even in languages from the same genetic group (e.g. the South Arawak subgroup).

(c) Evidentials often do not interact with tense-aspect distinctions (in the way that they do in Tuyuca, described above).

(ii) An EYEWITNESS/NON-EYEWITNESS system is found in just one language, Madi, from the Arawá family (Dixon and Vogel, forthcoming). The Jarawara and Banawa dialects of this language have an obligatory eyewitness choice in the three past tenses—immediate past (from a few seconds to a few weeks ago), recent past (a few weeks to a few years ago), and far past (anything more than a few years ago). Note that tense and evidentiality are fused into one suffix.6 However, the Jamamadi dialect of Madi only makes an evidentiality distinction in recent past and far past. The immediate past eyewitness form in Jarawara and Banawa is morphologically anomalous, in a way which suggests that it was recently introduced, and formed partly by analogy. This, and the fact that the evidentiality system is not known to operate in any other language from the family,7 suggests that it is a recent development.

The Madi dialects also have an optional verbal suffix indicating that the statement of the clause is reported. This can be used with non-eyewitness past tenses (traditional narratives normally have non-eyewitness far past tense marking plus the reported suffix on each main clause verb); it cannot, of course, be combined with eyewitness past tenses.

(iii) The most complex evidentiality system in southern Amazonia is that in the Nambiquara family (Lowe, 1972: 368–9; and forthcoming). There is obligatory marking on the verb for whether a statement is EYEWITNESS, INFERRED or REPORTED or whether the speaker has a GUT FEELING that something has happened. (No information has been provided on whether there is any interaction between tense and evidentiality. More data is needed.)

3.2. Evidentials in northern Amazonia

Languages in the central part of the northern region (around the Vaupes River basin on the Brazil/Colombia border) have the most complex evidentiality systems, with four or five members. In these languages evidentiality may interact with mood, so that some specification of evidence is included with interrogatives. We find smaller systems, with just two or three evidentiality choices, in languages on fringes of the region—north towards Venezuela, west towards Equador, and south-west into Peru.

(i) A two-term system of REPORTED/NON-REPORTED is found in a number of North Arawak languages—Resigaro (on the border between Peru and Colombia), Piapoco (in Colombia), Baniwa of Içana (on the border between Brazil, Colombia and Venezuela—see Section 4.2) and Achagua (on the border between Colombia and Venezuela).

These simple systems appear to be of recent origin—they are not obligatory and employ different clitic forms in these rather closely related languages. Evidentiality clitics generally attach to the verb, but in Achagua the clitic -minaa, used to mark
reported or assumed action, may instead go on a topicalised constituent (Wilson, 1992: 102–3).

A system of REPORTED/NON-REPORTED is also found in Arabela, from the Zaparoan family in north-east Peru (Wise, forthcoming).

(ii) Another two-term system, of EYEWITNESS/NON-EYEWITNESS, is found in Koreguaie, Secoya and probably also in Siona, three languages of the West Tupanoan subgroup, spoken in Colombia and Equador (Cook and Criswell, 1993: 86–87, Wheeler, 1987). We discuss the question of their origin in Section 4.1.

(iii) In Bora, of the Witotoan family, in north-east Peru, there is specification of VISUAL (with zero marking) versus NON-VISUAL (marked by suffix -hja). Within non-visual there is a further contrast between REPORTED (marked by -va, which precedes -hja) and NON-REPORTED (with zero realisation) (Thiesen, 1996: 97).

(iv) A three-term system of REPORTED, INFERRED and UNMARKED was used in Bahwana, a North Arawak language spoken on the Middle Rio Negro in Brazil (Ramirez, 1992: 64–65). Materials on this language are, unfortunately, very limited (being based on work with the last speaker who is now dead) and it is impossible to say under what conditions an evidentiality specification was obligatory, and what its interrelations were with other grammatical systems.

(v) It appears that languages of the Yanomami family (which straddles the border between Brazil and Venezuela) have markers of evidentiality, some shown by affixes to the verb and others by particles. From the information in available grammars it appears that the number of evidentials varies, according to the distance of individual languages from the Vaupes River basin, which is the focus of the northern region of evidentiality. Xamatauteri is the language closest to the Vaupes; here Ramirez (1994: 316ff.) recognises four evidential markers—EYEWITNESS, NON-EYEWITNESS, INFERRED and REPORTED; note that the non-evidence marker only appears in recent past tense, and there is no evidentiality specification in future tense. The language to the north of Xamatauteri is Sanuma, for which Borgman (1990: 165–73) describes three evidential markers—EYEWITNESS (used in present and past tenses), VERIFICATION, by seeing evidence or by hearing from someone who has first-hand knowledge of the state or event (also in present and past) and SUPPOSITION (used in present, past or future). The language furthest from the Vaupes (north-east of Xamatauteri and east of Sanuma) is Yanam which Gomez (1990: 97) says shows only two evidentials, EYEWITNESS and NON-EYEWITNESS.

(vi) The most complex evidentiality systems from anywhere in the world are found in languages from the East Tucano subgroup, spoken in the Vaupes River region of Brazil/Colombia—in Bara, Carapana, Siriano, Tatuyo, Desano, Tuyuca, Tucano, Wanano-Piratapuyo, Barasano and Macuna (Barnes, forthcoming). And, as we shall discuss in Section 4.2, the system has diffused into Tariana, a North Arawak language spoken in the same region.

In Section 2 we described the five terms in the Tuyuca system—VISUAL, NON-VISUAL, APPARENT, SECONDHAND and ASSUMED. Other languages have the same system, with cognate markers. In a number of East Tupanoan languages (e.g. Tucano and Desano) there are just four evidentials, a single suffix covering both assumed and apparent.8

In all of these languages, evidentiality interacts with tense. It is not marked in future, and there is typically no secondhand specification in present tense. Table 1 illustrates the evidential suffixes in past and present tense for Tuyuca; note that there
Table 1. Evidentials in Tuyuca—affirmative mood (from Barnes, 1984:258)

<table>
<thead>
<tr>
<th></th>
<th>Visual</th>
<th>Nonvisual</th>
<th>Apparent</th>
<th>Secondhand</th>
<th>Assumed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Past tense</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3m.sg</td>
<td>-w-i</td>
<td>-i-i</td>
<td>-y-i</td>
<td>-yi-gi</td>
<td>-lly-i</td>
</tr>
<tr>
<td>3f.sg</td>
<td>-w-o</td>
<td>-i-o</td>
<td>-y-o</td>
<td>-yi-go</td>
<td>-lly-o</td>
</tr>
<tr>
<td>3pl</td>
<td>-w-a</td>
<td>-i-a</td>
<td>-y-a</td>
<td>-yi-ra</td>
<td>-lly-a</td>
</tr>
<tr>
<td>other</td>
<td>-w-i</td>
<td>-i-i</td>
<td>-y-u</td>
<td>-yi-ro</td>
<td>-lly-u</td>
</tr>
<tr>
<td><strong>Present tense</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3m.sg</td>
<td>-i</td>
<td>-y-i</td>
<td>-lly-i</td>
<td>-k-i</td>
<td></td>
</tr>
<tr>
<td>3f.sg</td>
<td>-y-o</td>
<td>-g-o</td>
<td>-lly-o</td>
<td>-k-o</td>
<td></td>
</tr>
<tr>
<td>3pl</td>
<td>-y-a</td>
<td>-g-a</td>
<td>-lly-ra</td>
<td>-k-a</td>
<td></td>
</tr>
<tr>
<td>other</td>
<td>-u</td>
<td>-g-a</td>
<td>-k-un</td>
<td>-k-u</td>
<td></td>
</tr>
</tbody>
</table>

‘Other’ refers to 1st and 2nd person singular and plural, and to 3rd person inanimate. There is a gap at ‘other, apparent’ in present tense. Apparent cannot be used with 1st person subject in present tense; for a second person subject one of the third person forms is used.

There are different forms depending on the person, number and gender of the subject of the clause.

There are also interesting interactions between person and evidentiality. Non-visual evidentials are preferred when speakers refer to their own cognitive states (e.g. knowing or understanding). Assumed evidentials are preferred when speaking about a third person—there is no way of perceiving that such a state exists inside another person, one can really only assume it (Malone, 1988: 131).

In Tucanoan languages with a four or five-term evidential system in affirmative mood there is a reduced system in interrogative mood. In the remote past tense there is an opposition between visual, non-visual and an evidentiality choice covering apparent, assumed and secondhand; in present tense this reduces still further to a contrast between visual and all other possibilities. The Tuyuca paradigm is given in Table 2.

4. Development of evidentials in northern Amazonia

4.1. Origins of evidentials in the East and Central Tucanoan languages

Malone (1988, p.c.) provided a fine internal reconstruction of the forms in Tables 1 and 2, and undertook a comparison of evidential systems across Tucanoan languages. She concluded:

(a) There was at first a contrast ±FIRSTHAND, and there then developed a secondary distinction within -firsthand of ±REPORTED (giving secondhand and assumed evidentials).

(b) There was then development of a contrast ±VISUAL within +firsthand (giving visual and non-visual evidentials) and of ±DIRECT within -firsthand, +direct giving the apparent evidential and -direct relating to the secondhand and assumed specifications. She suggested that these developments ‘may have been forced by a shift from a seman-

Table 2. Evidentials in Tuyuca—interrogative mood (from Barnes and Malone, 1992)

<table>
<thead>
<tr>
<th></th>
<th>Visual</th>
<th>Nonvisual</th>
<th>Apparent/secondhand/assumed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Present tense</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-i (&lt; *-RI)</td>
<td>-ga-ri</td>
<td>-ga-ri</td>
</tr>
<tr>
<td>Remote past tense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ri</td>
<td>-ta-ri</td>
<td>-yi-ri</td>
</tr>
</tbody>
</table>
tic emphasis on speaker distance in time and space, or cognitive distance from an action or state, to an emphasis on the kind of evidence that the speaker sees. It is hard to tell whether the four/five-term evidentiality system developed in proto-East-Tucano and proto-Central-Tucano, and was then inherited by the modern languages, or whether it developed in one language and spread by diffusion. This is an area of rampant multilingualism, due to obligatory exogamy, which provides a perfect pasture for the diffusion of grammatical categories and forms (see Sorensen, 1967; Aikhenvald, 1996a).

4.2. Evidentials in Tariana

Tariana is the only language of the North Arawak subgroup spoken among the East and Central Tucanoan languages in the Vaupes region. In the prevailing situation of obligatory multilingualism it has experienced a heavy phonological, morphological, syntactic and semantic impact from its Tucanoan neighbours. Although the lexicons are kept strictly apart, Tariana shows a remarkable structural similarity to the East Tucano languages of the Vaupes, especially to Tucano.

One of the points of similarity is the system of evidentials. Tariana makes four specifications, the same as Tucano—VISUAL, NON-VISUAL, INFERRED (corresponding to apparent and assumed in Tuyuca) and SECON DHAND. For instance, in describing an event such as ‘The dog ate the fish’ use of the visual evidential would imply that the speaker saw the dog eat the fish. The non-visual marker might be used if the speaker heard the noise of the dog eating, in the next room. The secondhand evidential would be employed if someone told the speaker that the dog had eaten the fish. And the inferred evidential could be used if there were fishbones on the floor near the dog, which looked satisfied as if after a good meal (this would correspond to the apparent marker in Tuyuca) or if the speaker assumed the dog must have been the culprit, since the fish was raw and only a dog would eat raw fish (this corresponds to the assumed evidential in Tuyuca).

Like the East Tucano languages, Tariana makes an evidentiality choice just in present and past tenses, the two specifications occurring in one portmanteau suffix. The paradigm is given in Table 3.

Like Tuyuca and Tucano, Tariana also includes a three-way evidential specification in its mechanism of interrogative marking. The paradigm is given in Table 4. Note that in Tariana the tense distinction is neutralised in non-visual evidential, whereas for Tuyuca (Table 2) an evidentiality distinction is neutralised in present tense.

There is both ethno-historical and linguistic evidence that the Tarianas are comparative newcomers to the Vaupes region, having arrived after the Tucanoans (Aikhenvald, 1996b). They came down from tributaries of the Íçana river, where they had lived in contact with Baniwa of Íçana, their closest relative within the North Arawak sub-

<table>
<thead>
<tr>
<th>Present tense</th>
<th>Visual</th>
<th>Non-visual</th>
<th>Inferred</th>
<th>Secondhand</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-naka</td>
<td>-mka</td>
<td>-sika</td>
<td></td>
</tr>
<tr>
<td>Recent past tense</td>
<td>-ka</td>
<td>-makhia</td>
<td>-nikka</td>
<td>-pidaka</td>
</tr>
<tr>
<td>Remote past tense</td>
<td>-na</td>
<td>-shina</td>
<td>-pidana</td>
<td>-nkina</td>
</tr>
</tbody>
</table>
group. Baniwa and Tariana still have about 80% vocabulary in common. However, the grammatical differences are striking, mainly owing to Tariana having accommodated itself to the structural patterns of Tucanoan languages.

Evidentials are a point of major difference. As mentioned under (i) in Section 3.2, Baniwa of Içana shows rudimentary grammatical marking for evidence—a clitic -pida is optionally attached to the verb (or to the topic NP) to mark non-firsthand information. It is often used in traditional stories and other narratives which do not involve the speaker directly. This marker is the basis for two of the Tariana evidentials given in Table 3—secondhand recent past -pidaka and inferred remote past -pidana; the last of these is typically used in narratives which do not involve the personal experience of the speaker.

A marker (with different form) for non-firsthand evidence does occur in a few other North Arawak languages. Other members of this genetic subgroup lack evidentials, e.g. Bare (spoken to the south of the Vaupes region) and Warekena (spoken on other tributaries of Içana)—see Aikhenvald (1995, forthcoming-a). Tariana is the only North Arawak language to have acquired a complex and obligatory morphological system of evidentials, on the Tucano pattern.

4.3. Evidentials in Makú languages

There is one other type of language spoken in the Vaupes region—four languages from the small Makú family (these were probably the earliest languages in the region; Nimuendajú, 1982: 169). Nowadays the Makú (who are hunters and gatherers) are accorded inferior status by the Tucanos and Tarianas (who are agriculturalists) and are not integrated into the marriage system with its concomitant multilingualism. As a result there has been far less Tucanoan influence on the grammars of these Makú languages than there has been on Tariana. There has, however, been some influence.

The only one of the Vaupes Makú languages for which we have fair grammatical information is Dâw, and it does have an optional clitic marking non-firsthand evidence (Martins and Martins forthcoming). Makú languages spoken outside the Vaupes region (for which there are reliable descriptions) do not show any evidentiality markers.

4.4. Evidentials in Tucanoan languages outside the Vaupes

Tucanoan languages spoken outside the Vaupes area have a very reduced system of evidentials. Retuarã (also known as Tanimuca), from the Central Tucano subgroup, is spoken in Colombia next to Yucuna, a language from the North Arawak subgroup that lacks evidentials. Yucuna is the dominant language in this region, with speakers of Retuarã also being bilingual in it. As a result of Yucuna influence, Retuarã has lost
the high unrounded central vowel, ï, it has simplified the system of classifiers (Gomez-Imbert, 1996: 445, and personal communication); and it has reduced its system of evidentials. There are now just three evidential specifications—AUDITORY, ASSUMED and REPORTATIVE—which are optional and used only rarely (Strom, 1992: 90–1). Besides their optionality, evidentials in Retuarâ differ from those in Tucanoan languages from the Vaupes region in being marked with separate suffixes (rather than being portmanteau with tense) and with there being no interdependence between tense or mood and evidentials.

In the West Tucano subgroup, there is a simple specification ±EYEWITNESS for Koreguaje, for Secoya and probably also for Siona. The fourth member of this subgroup Orejón (spoken in Peru) has no evidentials at all (Malone, 1988: 124).

5. The multiple evolution of evidentials

There are two kinds of grammatical diffusion. The most common is diffusion of a category; less common is diffusion of grammatical forms. For example, if a language moves into a linguistic area characterised by each language having a system of noun classes (genders) then it is likely to innovate a noun class system. But it will do this by developing noun class markers from its own internal resources (e.g. by grammaticalising classifiers to be affixes marking noun class) rather than by borrowing the actual forms from a neighbouring language.

It is clear that in parts of Amazonia the category of evidentials has diffused from one language to some of its neighbours. Or, that a language with a small number of evidentials has augmented its system through contact with a language that already had a larger set. But it is also clear that all the occurrences of evidentials cannot be explained through this category developing just once, in a single language, and diffusing from there. The available information points to a grammatical category of evidentials having been independently innovated in at least six separate places.

Looking first at the occurrences of small systems of two or three evidentials in southern Amazonia (Section 3.1). The evidence here suggests at least four separate innovations:

(1) Kamaiurá, on the Upper Xingu River
(2) Nambiquara, on the headwaters of western tributaries of the Madeira River
(3) Madi, close to the Purús River

These show different types of evidentials, with different markers. The territories of groups (1–3) are geographically separate. Although we do not know the full historical picture, it is unlikely that any of these groups has been in contact with either of the others in recent times (and the evidentiality systems in Kamaiurá and Madi, at least, are clearly of recent origin).

(4) Languages of the Panoan family, and various languages of the South Arawak subgroup. The Panoan languages are closely related (and the family must have a shallow time depth); it is possible that an evidentiality system developed in proto-Panoan. As stated in Section 3.1, in South Arawak languages the evidential markers differ from language to language; the category may have diffused but the forms must have developed separately in each language. These South Arawak languages are spoken in
the same general region as Panoan. It is quite likely that a REPORTED/NON-REPORTED system of evidentiality developed just once, either in Panoan or in one of the South Arawak languages, and that the category then diffused into nearby languages.

Languages of the Panoan family are adjacent to Kulina, the westernmost member of the Arawá family. However, evidentiality has not been reported in Kulina, only in Madi, further to the east. Add to this the fact that the Madi specification is basically EYEWITNESS/NON-EYEWITNESS, and it appears unlikely that the category of evidentiality could have diffused between areas (3) and (4).

Turning now to the northern region, it appears that evidentiality must have developed independently in at least two places:

(5) Yanomami. The languages of this family are very closely related (being almost dialects). It is likely that proto-Yanomami was spoken in a more restricted area, and that there has been recent expansion and split. However, this matter has not received serious study, and there has been no reasoned argumentation as to where the proto-Yanomami homeland may have been.

(6) Tucanoan. The Tucanoan people have legends that they originated in the west, and there is clear ethno-historical and linguistic evidence that they moved eastwards, into the Vaupes region, within the last 500–1000 years (Nimuendaju, 1982: 170; Aikhenvald, 1996a). That is, Tucanoan probably came into contact with Yanomami within this period. It is highly unlikely that either group borrowed the category of evidentiality from the other. Rather, it seems that evidentiality developed independently in Tucanoan and in Yanomami, but that the number of evidentiality choices in Yanomami languages closest to Tucanoan then increased, presumably due to diffusional pressure from Tucanoan.

As mentioned in Section 3.2 four geographically separate North Arawak languages—Resigaro, Piapeco, Baniwa of Içana and Achagua—have a REPORTED/NON-REPORTED system, but with different markers. All of these languages are in the same general area as Tucanoan languages and could have been in contact with them at some time in the recent past. The category of evidentiality could have diffused from Tucanoan into some or all of these languages; or evidentiality could have evolved independently—in one or two or three or all four cases.

We stated in Section 3.2 that there is a REPORTED/NON-REPORTED evidentiality system in Arabela and a VISUAL/REPORTED NON-VISUAL/OTHER NON-VISUAL system in Bora. These languages are spoken in the same region as Tucanoan languages and evidentiality is likely to have spread by diffusion.12

There is no doubt that Dâw, a Makú language spoken in the Vaupes region, developed an evidential marker under areal influence from Tucanoan. And that Tariana, the only North Arawak language in this region, was also operating under diffusional pressure from Tucanoan when it replaced the original optional reported clitic *pida (which is maintained in its close genetic relative Baniwa of Içana) by an obligatory four-term system, realised in portmanteau form with tense and mood.

The Tucanoan family has already been discussed. Eastern languages have a four/five-term system while western languages simply distinguish EYEWITNESS/NONEYEWITNESS. There is no doubt the category of evidentiality was innovated at some time and at some place within the Tucanoan family. It could have been in proto-
Tucanoan, or in proto-East-Tucanoan, or in one modern language, with diffusion of categories and in this instance also of forms.

In summary, the evidence points to evidentiality, as a grammatical category, having been independently innovated in at least six places: (1) Kamaiurá; (2) Nambiquara; (3) Madi; (4) Panoan/South Arawak; (5) Yanomami; (6) Tucanoan. There may well have been more than six innovations—there could have been several in the Panoan/South Arawak area, and there could have been some in North Arawak languages. There might even have been more than one innovation within Tucanoan.

Working with the minimal list of six places of innovation, the category of evidentiality must in some way have diffused (a) between Panoan and South Arawak; (b) from Tucanoan into North Arawak (to establish a reported marker); (c) from Tucanoan into Makú; (d) from Tucanoan into Yanomami (to augment the system); (e) from Tucanoan into Tariana (replacing an optional reported marker by an obligatory and intricate grammatical system).

In Section 1 we noted that the Andean linguistic area shows a significantly different typological profile from the Amazonian area. However, one feature that is found in both areas is evidentiality. The Andean languages do not have an obligatory morphological system, but there are optional clitics that provide information about evidence. Peruvian Quechua has three: -mi for ‘firsthand information’ (where the speaker is convinced of what they are saying), -sh(t) ‘hearsay, reported’, and -chi ‘conjecture’ (Adelaar, 1977: 79, Weber, 1986). There are slightly different systems in Ecuadorian Quechua (Cole, 1982: 164–5) and in the Aymara family (Hardman, 1986).

Amazonian languages with evidentiality are not in geographical contact with Andean languages. But the Tucanoans state that they came from further to the west, and there could conceivably have been contact in the not-too-distant past. The most likely scenario is that evidentials evolved in the Andean area independently of what was happening in Amazonia. But it is not absolutely impossible that there was some category diffusion (in either direction).

6. Some speculations

One of the most important (but seldom faced) question in linguistics is: why is the grammar and lexicon of a given language organised in the way that it is? One language has five past and three future tenses; another has no tense distinctions at all but shows a system of aspctual choices—why? In one language it is possible to negate a clause or any constituent of a clause; in another only a clause can be negated—why? One language has several kinds of complement clauses; another has none—why?

A number of factors combine to determine why a language is organised in the way that it is. These include:

1. Diffusional influences. It is undoubtedly the case that categories are more often diffused than independently innovated. The existence of noun classes among most of the languages of sub-Saharan Africa, and of classifiers among most of the languages of east and south-east Asia, for instance, is undoubtedly due to areal diffusion.

2. Genetic inheritance. The unmarked situation is for a language to have the same grammatical categories, organised in the same way, as its ancestor language. But some categories can be lost and others gained, through a variety of causes.
3. Typological possibilities. A certain type of grammatical system, \( X \), may tend to co-occur with another type of system, \( Y \), and be seldom found in the same language as a further type of system, \( Z \). If a language has \( X \) it is likely to develop \( Y \), and unlikely to develop \( Z \). If this language does adopt \( Z \), through areal pressure, it is likely to lose \( X \). As one example of this, switch-reference marking is only found in languages with accusative (not ergative) syntax.

4. Organisation of discourse. For instance: the genres available; use of direct and indirect speech; strategies for turn-taking; conventions for ellipsis and anaphora; whether narratives are generally told in third or in first person (see Dixon, 1977: 406–7 for an illustration of how this correlates with the type of syntactic pivot used in the language).

5. Factors relating to the speakers and their environment. There are a number of possibilities here:

(a) Geographical environment. For instance, some languages make a grammatical specification 'up'/'down' in their deictic system (and/or elsewhere in the grammar). This is only found in languages spoken in hilly country.

(b) Social organisation, lifestyle, political and economic system. The most complex pronominal systems—with a distinction between three numbers (singular, dual and plural) or even four (singular, dual, paucal and plural) and/or an inclusive/exclusive distinction for first person non-singular—tend to be found in small-scale language communities with a classificatory kinship system. Another example concerns systems of obligatory, grammatical indicators of politeness. These are only found in languages spoken by sizeable nations that have a strongly articulated social hierarchy (e.g. Japanese and Balinese).

(c) The beliefs, mental attitudes and behavioural conventions of the speakers. For example in the Vaupes area there are numerous prohibitions, e.g. one must not eat cold food after dark (because this could anger spirits who would then emerge from the jungle and eat the transgressor); a man should not think about women when on a hunting expedition or he will be unlikely to catch anything; a woman should not eat, drink, urinate, defecate or make love during the process of making a ceramic pot, otherwise the pot will break. Correlating with this, the Tucanoan languages of the Vaupes (and Tariana as well) have a plethora of imperative markings, including 'do it immediately', 'do it at some distant place or time', 'do it on someone else's orders', 'speaker asks addressee to do something which they know the addressee will be reluctant to do', 'speaker asks addressee to do something that will be for the speaker's personal profit'. There is also a set of 'frustrative' clitics, such as 'you did it but I told you not to and now you must take the consequences'.

Parameters such as \( S(c) \) have been scarcely studied. We believe that such study will be rewarding and will suggest many lines of explanation for why languages are the way they are, in certain respects. In particular, we feel this will shed light on why some language have grammatical specification of evidentiality, while most lack this.

The use of a grammatical system of evidentials must surely correlate with such matters as:

(i) Whether there is a convention that one should be as specific as possible when speaking, or whether a high degree of vagueness is a normal social expectation.

(ii) Attitudes to the communication of information—whether one should tell people what they want to know, or whether 'new information' should be regarded as prized
EVIDENTIALS AND AREAL TYPOLOGY

goods, only to be disseminated in exchange for some appropriate return (see Keenan and Keenan, 1979: 147–53 and Weber, 1986: 138).¹³

(iii) Attitudes to truth, e.g. whether or not telling lies is an accepted social practice.

As already said, this is a topic that requires careful study. Beyond suggesting some possible parameters we do not, at this stage, suggest how they might be likely to correlate with whether or not a language has an evidentiality system.

What we do suggest is that communities in the Amazonian linguistic area (and also those in the adjoining Andean area) in some way share a common set of beliefs, mental attitudes and behavioural conventions;¹⁴ and that these are compatible with the development of an evidential system. This would explain why evidentiality has independently evolved in at least six (possibly many more) places in Amazonia, and is then particularly susceptible to be diffused.

Because of the way in which these communities act and think, they welcome specification of the nature of the evidence for a statement. Their languages are thus open to the introduction of grammatical marking of this evidence.

Acknowledgements—We are grateful to Terry Malone, Janet Barnes and Mengistu Amberber for their most useful comments on a draft of this paper.

NOTES

¹ The problems involved in distinguishing genetically inherited features from diffusional features have been discussed for American Indian languages north of Mexico by Sherzer (1976) and Bright and Sherzer (1976); and for Australian languages by Dixon (1980: 238–65). For this purpose, a very elaborate reconstruction on the level of individual subgroupings is needed, to really prove the point, as has been done with Indo-European languages in the Balkan linguistic area. The rather unsatisfactory state of historical linguistic study in South America (see Kaufman, 1990 for an overview) is a real problem in any attempt to distinguish between areal features and diffusional patterns. Difficulties associated with studying areal features in the Americas have also been pointed out by Thomson and Kaufman (1988: 113f).

² Tucanoan languages are entirely suffixing; the bound pronominal suffixes appear further from the root than suffixes that mark valency-changing derivations.

³ There are of course a few exceptions to these pan-Amazonian tendencies. For instance, Tupi-Guarani languages are the only languages in Amazonia which allow incorporation of unpossessed nouns; however, in modern languages this technique is falling out of use (Seki, forthcoming; Kakumasu, 1986). In Nadeh, a Makú language from the Middle Río Negro, incorporated adverbs and adpositions precede the verb root instead of following it. In Palikur, a North Arawak language from Brazil and French Guiana, incorporated body parts follow the root (Green and Aikhenvald forthcoming). There are subareas which are characterised by properties different from those just described. These properties can serve as diagnostic traits for determining regions of diffusion. For instance, gender assignment is not semantically transparent in a region of southern Amazonia centred on the Purús River basin (where Bolivia, Brazil and Peru meet) which includes languages from the Arawá and Chapacura families and the Peruvian and Pre-Andine subgroups of Arawak. Classifiers are used in different morphosyntactic contexts in different regions—for instance, they are used in possessive constructions in the Upper Río Negro region, but not in Peruvian Arawak languages.

⁴ The term 'verificational' is sometimes used in place of 'evidential'. There is an excellent summary of work on recognising this category, and naming it, in Jacobsen (1986). See Wierzbicka (1994; 1996: 427–58) for discussion of the semantics of evidentiality.

⁵ Languages in other parts in Amazonia lack evidentials—these include most languages from the Carib, Jé, Tupí and Makú families; and Pirahã.

⁶ Each past-tense/evidentiality suffix also carries information about the gender (masculine or feminine) of the pivot argument in the clause.

¹⁷ Dení, another language from the Arawá family, is reported by Koop and Koop (1985: 16) to have a number of optional evidential-like verbal suffixes, but with quite different form and meaning from those in Madi. They include 'deduction from seeing, hearing, feeling or smelling'; 'fact known to speaker but not to others spoken of'; 'fact which speaker knows at the moment of speaking'; 'fact which the speaker didn’t know at
the time referred to’ and ‘fact which the speaker now knows, but didn’t know at the time referred to’. These suffixes require further detailed study.
9 See West (1980) and Ramirez (1997) for information on Tucano; and Jones and Jones (1991) for Barasano.
10 Note that the East Tucano languages (together with one of the two languages in the Central subgroup) constitute one geographical block, whereas West Tucano languages are spoken in discontinuous areas to the south and west.
11 When Baniwa speakers tell stories in Portuguese they make heavy use of diz que (‘they say that’) as a reported marker, in much the same contexts as they would use puida when telling a story in Baniwa.
12 In Koreguaje (spoken in Colombia) eyewitness is left unmarked and an auxiliary verb is used to mark non-eyewitness (Cook and Criswell, 1993: 86–87). Malone (1988: 124) reports a eyewitness system for Secoya (spoken in Ecuador) but we have no further details. Information on Siona (spoken in Colombia and Ecuador) is unclear but there appear to be eyewitness and non-eyewitness evidentials, portmanteau with tense markers (Wheeler, 1987: 152–153).
13 It is quite possible that Resigaro borrowed the category of evidentiality from Bora, since they are spoken in adjacent communities (Allin, 1975).
14 In Amazonian society it is held that there is an explicit cause for everything that happens. So as not to be blamed for something that in fact they had no responsibility for, a speaker is careful always to be as explicit as possible about what they have done. This relates to the desirability of stating the evidence for everything that is said. (Based on an idea communicated by Terry Malone, p.c.). Other factors may also be relevant. For example, in Tariana there are different conventions concerning the use of remote past evidentials in different kinds of narrative—see Alkhenvald, forthcoming-c.

REFERENCES


Seki L. (Forthcoming) Gramática Kamaíurú. To be published by Editora da Unicamp, Campinas, Brazil.


