

## Information source and evidentiality: what can we conclude?

Alexandra Y. Aikhenvald

Every language has an array of ways of referring to information source: this may be accomplished with verbs referring to reports, claims, or opinions, with adverbs, with parentheticals, prepositional phrases or with particles. In about one quarter of the world's languages, marking information source is obligatory. These languages have a grammatical category of evidentiality. Other languages have evidential extensions of non-evidential categories – such as perfect in Georgian and participles in Lithuanian. Just like non-grammatical expressions of information source, evidential extensions of non-evidential categories (known as evidential strategies) share the evidential meanings and not infrequently give rise to grammatical evidentials.

The term “evidential” primarily relates to information source as a closed grammatical system whose use is obligatory. The term “information source” relates to the corresponding conceptual category. This is akin to the distinction between the category of ‘tense’, as grammaticalized location in time, and the concept of ‘time’. Expressions related to information source are heterogeneous and versatile. They include closed classes of particles and modal verbs, and an open-ended array of verbs of opinion and belief. The term ‘lexical evidentiality’ is misleading in that it obscures these vital differences.

Extra-grammatical ways of marking information source may allow more detailed specification of various degrees of assumption, inference, opinion than do grammatical evidential systems, and often reliability, and speaker's evaluation of information. Recurrent patterns of expressing information source demonstrate parallel development of European languages discussed in the volume.

### *1. Information source and evidentiality*

Every language has a way of saying how one knows what one is talking about, and what one thinks about what one knows. But the ways in which the information source can be expressed vary. Languages differ not in what one can say, but in what one must say – as stated by Boas (1938), one of the founders of modern linguistics and of the study of grammatical expression of information source: “grammar [...] determines those aspects of each experience that must be expressed” (Boas 1938:132). One language may have a two-term gender system, while another has five genders and a third makes no gender distinctions at all in its grammar – though ‘man’ and ‘woman’, and males and females are clearly distinguished in other ways. Along similar lines, some languages have grammatical tense, and others do not. But in any language, one can talk about time.

In about one quarter of the world's languages, marking a limited selection of information sources is a must. In Tariana (Arawak), Matsigenka (Panoan), Makah (Wakashan), Hup (Makú), Quechua and Aymara, a clause without a marker of information source would not be acceptable to a native speaker. This is grammatical evidentiality – a brief summary is in §2.

Verbal categories whose main meanings do not reflect information source can acquire evidential extensions. Giacalone-Ramat & Topadze (§2.1, this volume) show how perfect aspect in Georgian regularly extends to cover evidential meanings associated with inference. This is an example of an evidential strategy – see the summary in §3.

Every language – no matter whether it has a fully grammaticalized evidential system, or evidential extensions of other categories – has an array of further ways of describing how one knows things. This is what most papers in this volume are about – see §4.<sup>1</sup>

Terminological clarity is essential in any branch of linguistics: the importance of distinguishing information source and grammatical evidentials is highlighted in §5. §6 is a brief summary.

## 2. *Evidentiality as a closed grammatical system: a bird's eye view*

As a category in its own right, evidentiality is a relatively recent 'arrival' on the linguistic scene – in contrast to other categories such as person, gender, number and tense which have been household concepts in linguistics for thousands of years (see Robins 1967). This may well be the reason why the proper limits of evidentiality are still debated by some.

The idea of obligatory marking of information source goes back to Boas, and his sketch of Kwakiutl (1911:443; 496). "The source, or nature, of human knowledge (known by actual experience, by hearsay, by inference)" is listed by Sapir (1921:108-109) alongside other grammatical concepts, such as person, modality, number and tense. Since Boas's work, the notion of grammatical evidentiality has made its way into many grammars of North American Indian languages. But for grammarians of European languages it remained largely unknown.

The term "evidential" as a label for the grammatical category of information source was first introduced by Jakobson in 1957; and became established by the mid-60s (see Jacobsen 1986:4-7; Aikhenvald 2004:10-17). That is, Dendale & Van Bogaert's (this volume) assertion that "a semantic study on opinion verbs published

by Oswald Ducrot more than thirty years ago (Ducrot 1975)”, “was years before the notion of evidentiality was introduced and studies on the subject became widespread” needs to be qualified. Lazard (1957) was among the first French linguists to have discussed evidential meanings (“inférenciel”), based on the material from Tajik, an Iranian language.

In languages with obligatory evidentiality, a closed set of information sources has to be marked in every clause – otherwise the clause is ungrammatical, or the speaker incompetent, or even not quite right in his mind (Weber 1986:142). Evidentiality is a category in its own right, and not a subtype of modality, tense, or mood.<sup>2</sup>

Languages with grammatical evidentials divide into a number of types depending on how many information sources are assigned a distinct grammatical marking. Semantic parameters employed in languages with grammatical evidentiality cover physical senses, several types of inference and of report. The recurrent terms in the systems are:

- I. VISUAL covers evidence acquired through seeing.
- II. SENSORY covers evidence through hearing, and is typically extended to smell and taste, and sometimes also touch.
- III. INFERENCE based on visible or tangible evidence or result.
- IV. ASSUMPTION based on evidence other than visible results: this may include logical reasoning, assumption or simply general knowledge.
- VI. REPORTED, for reported information with no reference to who it was reported by.
- VII. QUOTATIVE, for reported information with an overt reference to the quoted source.

The maximum number of evidential terms in a system appears to be five. Recent studies in grammatical evidential systems have revealed the existence of further terms. For instance, Yongning Na (Mosuo), a Tibeto-Burman language (Lidz 2007), has a special term in its evidentiality system which covers exclusively general knowledge. No spoken language has a special evidential to cover smell, taste, or feeling (not so in sign languages: Catalan sign language is reported to have a special evidential marking smell: Sherman Wilcox, p.c.).

Semantic parameters group together in various ways, depending on the system. The most straightforward grouping is found in three-term systems – where sensory parameters (I and II), inference (III and IV) and reported (V and VI) are grouped together, as in Quechua, Shilluk, and Bora (Aikhenvald 2004:145-146; 159-166). Numerous

languages of Eurasia group parameters (II-VI) under a catch-all non-firsthand evidential, as does Abkhaz and Yukaghir.

Evidentials may or may not have epistemic extensions, to do with probability and speaker's evaluation of the trustworthiness of information.<sup>3</sup> Visual evidential in Quechua can refer to information the speaker vouches for. Not so in Tariana or Tucano.

Reported evidential in Estonian has an overtone of doubt. This is akin to how 'they say' in English may imply that the speaker does not really believe what is being reported, or to how the ubiquitous *dizque* has overtones of doubt in many varieties of South American Spanish (Kany 1944:171; Travis 2006; and Olbertz, this volume). In contrast, in Quechua, Shipibo-Konibo and Tariana, the reported evidential does not have any such overtones. As Valenzuela (2003:57) remarks for Shipibo-Konibo, the selection of reported evidential over the direct evidential "does not indicate uncertainty or a lesser degree of reliability but simply reported information". These languages have a plethora of other categories which express doubt, belief, disbelief and so on. See Chapter 5 of Aikhenvald (2004) for a survey of epistemic extensions, or lack thereof, in the grammatical evidentials in the world's languages.

Just like most other grammatical categories, evidentials interrelate with moods (that is, clause types: Lyons 1977). The maximum number of evidential specifications tends to be distinguished in declarative main clauses. The most frequent evidential in commands is reported ('do what someone else told you to').

Future and non-indicative modalities – conditional, dubitative and so on – (not to be confused with moods) may allow fewer evidential specifications than the indicative. The maximum number of evidential specifications is found in past tenses.

When used with a first person subject, the non-visual, non-firsthand evidentials and reported evidentials in systems of varied types may acquire additional meanings to do with lack of intention, control, awareness and volition on the part of the speaker. Verbs covering internal states may require obligatory evidential choice depending on person. As a result, evidentials may acquire the implicit value of person markers.

Evidentials typically have sentential, or clausal, scope. Dependent clauses usually cannot have an evidential value different from that of a main clause. Having an NP within the scope of a grammatical evidential is highly unusual: the only example known so far comes from Jarawara (an Arawá language from South America: Dixon 2003; Aikhenvald 2004:88).<sup>4</sup>

### *3. Evidential strategies*

Meanings to do with how people know things may be expressed without developing a dedicated form whose primary meaning is information source. Non-evidential categories frequently acquire evidential extensions. A conditional mood, or a perfect, or a passive can develop an evidential-like meaning as a “side effect” (see the discussion in Lazard 1999).

One of the best-known examples is the conditional in French (known as “conditionnel d’information incertaine”) used to relate information obtained from another source for which the speaker does not take any responsibility – this is mentioned in §1 of Dendale & Van Bogaert, this volume. Conditional in Italian also extends to cover reported information (see §2 of Pietrandrea, this volume, and also Squartini forthcoming).

An extension of the perfect aspect to express non-firsthand evidential meanings in Georgian is masterfully summarized by Giacalone Ramat & Topadze (§2.1, this volume) (also see Hewitt 1995:259; 293). This development is shared with many Iranian, Turkic, and Northeast Caucasian languages (also see Comrie 1976:110; Aikhenvald 2004:289-296).

Or the choice of a complementizer or a type of complement clause may serve to express meanings related to how one knows a particular fact. In English, different complement clauses distinguish an auditory and a hearsay meaning of the verb *hear*: saying *I heard John cross the street* implies that I did hear John stamping his feet, while *I heard that John crossed the street* implies a verbal report of the result. That is, a *that*-clause with perception verbs can refer only to indirect knowledge (see a concise analysis of complement clauses with verbs of perception in English in the context of complementation in general, by Dixon 2005:270-271).<sup>5</sup>

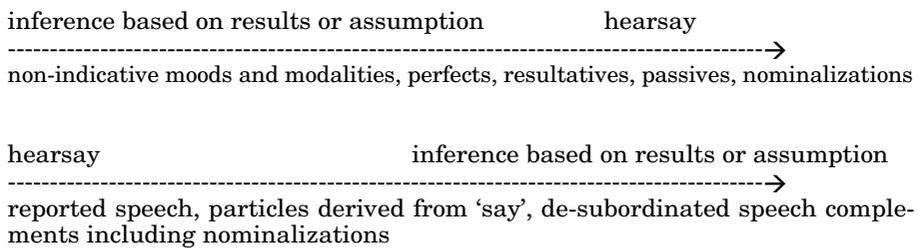
Nominalizations and participles often develop connotations similar to non-firsthand evidentials. In his discussion of Lithuanian, Wiemer (§4.2, this volume) mentions participles “as semi-grammaticalized means of indicating hearsay or inferential meanings”. This is consistent with Gronemeyer’s (1997) and Timberlake’s (1982) analysis of Lithuanian passive participles. According to Mathiasen (1996:134-5) and Gronemeyer (1997) active participles as heads of predicate in Lithuanian have been reinterpreted as reported evidentials.

Marking of assertion and speaker’s authority correlates with speaker’s attitudes to information and – indirectly – to its sources. As shown by Pusch (this volume), enunciative particles in Gascony

Occitan mark speaker's assertion intertwined with meanings related to expression of the ways in which information was acquired. "Assertivity" in Gascony Occitan is not an evidential system; in Pusch's words, "it oscillates between some kind of modality [...] and evidentiality proper".

Evidentiality strategies typically develop a range of meanings characteristic of reported and non-firsthand evidentials: they combine reference to inference and to verbal report. And they are not averse to having epistemic extensions to do with probability, and also expressing speaker's attitudes to the veracity of what is being said. The range of meanings of evidentiality strategies is given in Scheme 1.

**Scheme 1.** The semantic ranges of evidentiality strategies



No language has been found to have a special evidentiality strategy for each of the evidential meanings which can be expressed (I-VI in §2). Most of the features outlined for grammatical evidentials in §2 are not characteristic of evidential strategies. They qualify as 'incipient' evidentials-in-the-making which tend to grammaticalize into a closed system of evidentials.

Over time, an evidential overtone of a non-evidential category may conventionalize as its major meaning. In other words, evidential strategies may develop into grammatical evidentials – as was the case in Lithuanian (Gronemeyer 1997). A future tense can give rise to a dedicated non-firsthand evidential, as happened in Abkhaz (Chirikba 2003:262-264). And the enunciative particle *que* in Gascony Occitan (Pusch, this volume) has the potential of developing into an evidential marker.

Not every extension to do with information source is an evidentiality strategy. Pietrandrea (§2, this volume) shows that the "epistemic future" in Italian is not really an evidentiality strategy: its connections with information source are tenuous, and its evidential extensions are context-dependent.





this volume), Italian *sembra* (§3.4 of Giacalone-Ramat & Topadze, this volume), *dice* ((43) of Pietrandrea, this volume); and French *dit-on* and *paraît-il* (§2.1 of Dendale & Van Bogaert, this volume). A parenthetical – defined as “a word, phrase, or sentence which interrupts a sentence and which bears no syntactic relation to that sentence at the point of interruption” (Trask 1991:199) – expresses more than ‘source of evidence’: it is a way of referring to one’s opinion, judgement, belief, inference, assumption, doubt, attitude and more (see Urmston 1952 and Ifantidou 1993, on their varied semantic effects).

Parentheticals in English are an open class: Dixon (2005:233-238) demonstrated that verbs of ATTENTION and THINKING semantic types can all be used as parentheticals (provided they can take a *that*-complement clause and have a positive meaning). So can many other verbs, and predicative adjectives.

The meanings of parentheticals – just like with lexical verbs – are broader than those of grammatical evidentials. This is what one expects of an open class. Whether parentheticals are universal remains an open question. No such phenomenon has so far been documented for Amazonian, Australian or Papuan languages.

#### 4.3. *Modal verbs and particles*

So-called “modal verbs” frequently combine reference to information source with whatever other meaning they have. In agreement with Dixon’s (2005) classification, modal verbs express secondary concepts, “those providing semantic modification of some other verb, with which they are in a syntactic or morphological construction” (p. 96). In many languages they are a closed subclass. Secondary verbs of the same semantic group as *seem* (pp. 203-205), and verbs of obligation and permission often extend to cover probability, inference, and assumption.

In Dixon’s (2005:204) words,

*seem* is used when the Arbiter is not fully certain whether the adjectival description is appropriate, or whether the statement of the complement clause in a construction like *It seems that Mary found the body* or *Mary seems to have found the body* is correct – perhaps when there is not quite enough evidence. *Appear* has the same syntactic possibilities and a very similar meaning, but may imply ‘can be observed by me’ in contrast to *seem* ‘can be inferred by me’.

A link with information source is obvious – yet information source is an overtone of *seem*, rather than its only meaning.

Modal verbs may share syntactic features – such as raising – and thus form a syntactically defined subclass of verbs (see de Haan’s discussion of such ‘raising verbs’ in English, German, Dutch and Swedish; Cornillie’s contribution on *parecer*, and the discussions of modal verbs by Pietrandrea, Giacalone-Ramat & Topadze, and Dendale & Van Bogaert in this volume).

The choice of modal verbs in each language is limited – they are closed subclasses. This makes it difficult to treat them on a par with lexical expression of information source through verbs, and adverbs, or parentheticals (see arguments by Cornillie, Pietrandrea, and also Dendale & Van Bogaert). Modal verbs tend to have more restricted meanings, close to those of grammatical evidentials. As exponents of information source, they are akin to evidentiality strategies (see §3 above; and discussion in Aikhenvald 2004:147-148).

A plethora of particles referring to verbal report, or inference, or both may form a largish but closed class: see the discussion of over twenty-five particles referring to verbal report or inference in Lithuanian in §3.1-2 by Wiemer (this volume). None of them is obligatory. Many come from depleted reanalyzed verbs of perception, as does *girdi* (§3.1.2), literally ‘you hear’, used to mark reported information, or speech, as does *tariamai* (§3.1.5), a present passive participle of the verb ‘say, pronounce’. Their meanings tend to be much less fine-grained and less specific than those of members of open classes (see §4.2-3 above). This is another non-obligatory, and yet non-lexical, way of expressing information source. They can be considered a type of evidentiality strategy (see Scheme 1).

#### 4.4. *Speech report constructions*

Every language has a way of reporting what someone else has said. This can be cast as a direct, or an indirect speech report (see Aikhenvald forthcoming for a summary). Multiclausal speech report constructions can be viewed as lexical ‘paraphrases’ of meanings grammaticalized in closed evidential systems. And in many languages, speech reports acquire epistemic overtones – see §2.1 of Pietrandrea (this volume) (and Table 1), and §3.2 of Dendale & Van Bogaert (this volume). They are often used to transmit something one does not really believe (see, for instance, Dimmendaal 2001, on reported speech as a ‘hedging’ device).

And it comes as no surprise that a speech report construction is a universal source for developing reported evidentials. One such grammaticalization path involves reanalysis of a biclausal quotation

or reportative construction whereby the matrix clause with the verb *say* and a complement clause of this verb become a single clause via the loss or reinterpretation of the subordinator (Aikhenvald 2004:273-274; 281-283). This is what we see in marker of reported speech, *dizque*, in Mexican and in Colombian Spanish.

Grammaticalization is a gradual process. In Italian, grammaticalization of the speech verb has just started. As Giacalone-Ramat & Topadze (§3.4, this volume) put it, ‘the third singular form *dice* is frequently used in spoken Italian mostly as a marker of direct speech, but also of indirect speech, and is morphologically invariable and positionally mobile’. Parallels to this are found in Sardinian and Rumanian.

In Giacalone Ramat & Topadze’s words, “the parameters which allow us to describe a shift in the direction of a more grammaticalized category are: 1) decategorization (i.e. loss of inflectional distinctions), 2) positional freedom, 3) variability in scope (i.e. single constituent vs. entire clause scope), 4) semantic erosion”. And this is what we find in expressions involving *dizque* (literally ‘says that’) in Latin American varieties of Spanish (Travis 2006, for Colombian Spanish; Olbertz, this volume, for Mexican Spanish; Olbertz 2005 for Ecuadorian Spanish; evidence for the same phenomenon in Venezuela, Chile and Argentina come from Kany 1944), and in Brazilian Portuguese (Aikhenvald 2002, 2004).

*Dizque* (often written as one word) has a variable scope – as Olbertz (this volume) shows, the scope of *dizque* can be a main clause (as in her (4)), a subordinate clause (5), or a noun phrase in any syntactic function (examples (17-21)). A similar phenomenon in Colombian Spanish was described by Travis (2006). In both varieties *dizque* extends to cover doubt and (negative) attitude to the information and its validity – that is, it goes beyond simply reporting what someone else has said. The meanings of *dizque* overlap with some of the meanings of grammatical evidentials – they may involve reported speech, quotation, inference and assumption. *Dizque* is on the way to grammaticalizing into a marker of a category (in the sense of Heine & Kuteva 2002) – but this mechanism is very different from employing lexical items to refer to information source.

*Dizque* in Mexican Spanish has a greater syntactic freedom than the verb *decir* ‘say, speak’ it comes from, because *dizque* has been reanalyzed as a grammatical particle marking both speech reports and unreliable information. *Dizque* follows a common grammaticalization path, and is not exceptional in any way. A similar path has been documented for many other languages of the world, including

Georgian. In §2.3, Giacalone-Ramat & Topadze discuss the particles *metki* and *-tko* which mark reported speech; both result from the grammaticalization of the verb *tkma* ‘say’.

Particles marking reported speech form a closed grammatical class. They cannot be subsumed under a broad umbrella of lexical expression of information source. Instead, they can be considered evidentials in the making – akin to evidential strategies.

#### 4.5. *Expressing information source: a summary*

Meanings associated with information source can be expressed with members of open classes. The range of meanings is wider and more fine-grained than that of grammatical evidentials. Closed classes of particles and modal verbs tend to share their meanings with evidential strategies.

The choice of a grammatical evidential often depends on mood or tense of the clause (see §2 above). The choice of a parenthetical or an adverb depends on what the speaker wants to say. A parenthetical, an adverb, or a modal verb can have an NP or a whole clause in its scope. For grammatical evidentials, these options are restricted. None of the means listed in §4.1-2 forms a paradigm of any sort. In contrast, grammatical evidentials do.<sup>6</sup>

What may justify putting various verbs, adverbs and parentheticals discussed in §4.1-2, together with modal verbs and particles is the fact that they all vaguely relate to the ways in which one knows things. All these devices for marking information source combine reference to inference, assumption, and often speech reports with increasing “subjectification” – a “historical pragmatic-semantic process whereby meanings become increasingly based in the speaker’s subjective belief state, or attitude toward what is said” (Traugott 1996:185) (also see the discussion by de Haan, this volume). This is what sets them apart from closed evidential systems – whose primary meaning has nothing to do with subjectification – and makes them similar to prototypical modalities.

### 5. *On terminological clarity*

Categories of grammar need to be distinguished from ‘real world’ notions. The expression of information source is akin to other conceptual notions in that they can be viewed differently, “in terms of their importance for the structure of the language” (Comrie 1985:8).

The first type is the set of grammatical categories. It is common knowledge that many languages have a closed set of grammaticalized expressions of location in time: these can involve present, past and future; or nonpast, recent past and remote past, etc. Along similar lines, languages with grammatical number typically distinguish singular and plural, or singular, dual and plural. These closed grammatical systems coexist with sets of lexical items which refer to location in time (e.g. *yesterday*, *today*, and so on) or to quantification (e.g. numerals and quantifiers). In addition, languages may have a potentially unlimited number of ‘composite lexical expressions’ for measuring time intervals, or for expressing quantification.

In other words, a closed grammatical system offers restricted options. This is in contrast to the lexicon where the choices are potentially open. So, for grammatical tense “even the maximal system would have at most tens of categories, rather than the several orders of magnitude more possible in the lexicon” (Comrie 1985:9). The analogy with number is even more instructive: in Comrie’s (1985:9) words,

English has grammatically only a two-way opposition (singular and plural); lexically there are around thirty items (excluding those restricted to mathematical or scientific contexts); while for many speakers the possibilities for lexically composite [number] expressions are infinite.

The terminological distinction between *time* as a notional concept and *tense* as grammaticalized location in time is very handy: it helps to keep apart a closed grammatical system of tenses and a potentially open pool of temporal expressions of other sorts.<sup>7</sup> Such terminological clarity helps understand the nature of tense and time, and sheds light on the differences between closed grammatical systems and potentially open categories of the lexicon.

Let’s follow the analogy of tense and time. In the same way as tense refers to closed grammatical systems, ‘grammatical evidentiality’ refers to a closed set of obligatory choices of marking information source. In the same way as time covers the potentially unlimited set of choices, ‘information source’ covers the rest. Is it appropriate to use one cover term, ‘lexical evidentiality’, to include the rest? The brief answer is no.

Cornillie (this volume) complains that “following this view would entail that evidentiality cannot be considered from a functional perspective and would imply that most European languages do not have an evidential category”, and that this “account seems to be based on a dichotomous view of grammar and lexicon, which contrasts with the

idea of a continuum argued for in the studies on grammaticalization of the last twenty years”.

But the accepted definitions of grammaticalization are anchored in opposing lexical and grammatical notions as basic building blocks. From a diachronic perspective, grammaticalization is conceived as “that part of the study of language change that is concerned with such questions as how lexical items and constructions come in certain linguistic contexts to serve grammatical functions or how grammatical items develop new grammatical functions” (Hopper & Traugott 2003:1; Brinton & Traugott 2005:23-25; also see Heine & Kuteva 2002:4-5). In order to understand the diachronic and synchronic dynamics and development of the expression of information source, we need to adhere to terminological and conceptual clarity – rather than using the umbrella term ‘functional’ as a cover-up.

Saying that English parentheticals, or adverbs like *surely* or *allegedly*, are evidentials is like saying that time words like *yesterday* or *today* mark tense. Any grammarian would dismiss this as unmitigated rubbish. But if we say that *yesterday*, *today*, *two minutes ago* and so on and so forth express time, and present and past express tense, everyone will agree.<sup>8</sup> Along similar lines, the term ‘evidential’ is best used for closed grammatical systems, and the term ‘information source’ for the vast body of other ways of referring to knowing things.

The expression of information source which does not form a closed grammatical system and is not an extension of an existing category has been informally nicknamed “lexical evidentiality”. As we saw in §5, extra-grammatical means of expressing various overtones of information source and attitude to information cover a vast ground. Some of them are purely lexical, and some belong to closed subclasses. The term “lexical evidentiality” is misleading since it obscures the differences between the two types of extra-grammatical expression of information source outlined in §4: open choices on the one hand (§4.1-2) and closed classes (§4.3-4), on the other.

The term “lexical evidentiality” is confusing in yet another way. In languages with obligatory evidentials, lexical subclasses of verbs can require certain evidentiality choices. For instance, internal states and processes, felt rather than seen, are often cast in sensory (non-visual) or non-firsthand evidential. Such preferences may get lexicalized as restricted evidentiality choices for predicate types and construction types (this is a typical feature of Tibeto-Burman languages: see for instance Lidz 2007). Lexicalization of evidential choices is all too easy to confuse with “lexical evidentiality”.

6. *Envoi*

Evidentials as closed grammatical systems are different from information source marked in other ways (just like time, a real life concept, is different from tense, realized in grammar). Meanings related to information source may be expressed through open classes of verbs (of perception, opinion, speech and others), adverbs and parentheticals. These tend to be richer in their semantic range than closed systems of grammatical evidentials. Alternatively, information source may be expressed via a closed subclass of modal verbs, or via particles (often grammaticalized from verbs). These are much closer to grammatical evidentials in their nature, and their meanings.

Grammatical evidentiality is highly diffusable in language contact (see Aikhenvald 2006b, and §1.3 of Giacalone-Ramat & Topadze, this volume). And so are various other means which may involve information source. The languages of Europe – the subject matter of this volume – share a remarkable range of semantic extensions of various verbs, parentheticals and speech reports towards expressing how one knows things.

In Bolinger's (1991:26-7) words,

one of the happier results of recent turns in linguistics is the search for universals and the emphasis on parallel developments in various languages. [...] It is as if given certain elements from a common heritage, plus a need of communicate the same ideas, common solutions are going to be hit upon sometimes, though the element of chance still plays its part. The verb *parecer* in Spanish and the verb *seem* in English reveal just such a convergence. The etymological sources are quite different. Yet once set on a path toward the common meaning of that which is evident to the senses, their developing grammars grow more and more alike.

These striking parallel developments – the essence of Sapir's (1921:171-172) "parallelism in drift" – are what makes the typology of European languages so worthwhile.

*Address of the Author:*

Alexandra Y. Aikhenvald, Research Centre for Linguistic Typology, La Trobe University, Victoria 3086, Australia <a.y.aikhenvald@live.com>

*Bibliographical References*

- AIKHENVALD Alexandra Y. 2000. *Classifiers: a Typology of Noun Categorization devices*. Oxford: Oxford University Press.
- AIKHENVALD Alexandra Y. 2002. *Language Contact in Amazonia*. Oxford: Oxford University Press.
- AIKHENVALD Alexandra Y. 2003. *A Grammar of Tariana, from Northwest Amazonia*. Cambridge: Cambridge University Press.
- AIKHENVALD Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
- AIKHENVALD Alexandra Y. 2006a. Evidentiality in grammar. In BROWN Keith (ed.). *Encyclopedia of Language and Linguistics*. 2nd edition, Vol. IV. Elsevier: Oxford. 320-325 (article 0252).
- AIKHENVALD Alexandra Y. 2006b. Grammars in contact: a cross-linguistic perspective. In AIKHENVALD Alexandra Y. & R.M.W. DIXON (eds.) 2006. *Grammars in Contact: A Cross-Linguistic Typology*. Oxford: Oxford University Press. 1-66.
- AIKHENVALD Alexandra Y. Forthcoming. Semi-direct speech: Manambu and beyond. *Language sciences*.
- AIKHENVALD Alexandra Y. & R.M.W. DIXON (eds.) 2003. *Studies in Evidentiality*. Amsterdam: John Benjamins.
- BOAS Franz 1911. Kwakiutl. In BOAS Franz (ed.). *Handbook of American Indian Languages*. Part 1, Smithsonian Institution. *Bureau of American Ethnology Bulletin* 40. 423-557.
- BOAS Franz 1938. Language. In BOAS Franz (ed.). *General Anthropology*. Boston, New York: D. C. Heath and Company. 124-145.
- BOLINGER Dwight 1991. *Essays on Spanish: Words and Grammar*. Newark, Delaware: Juan de la Cuesta.
- BRINTON Laurel J. & Elizabeth C. TRAU GOTT 2005. *Lexicalization and Language Change*. Cambridge: Cambridge University Press.
- CHAFE Wallace L. & Johanna NICHOLS (eds.) 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex.
- CHIRIKBA Viacheslav 2003. Evidential category and evidential strategy in Abkhaz. In AIKHENVALD Alexandra Y. & R.M.W. DIXON (eds.) 2003. *Studies in evidentiality*. Amsterdam: John Benjamins. 243-272.
- COMRIE Bernard 1976. *Aspect*. Cambridge: Cambridge University Press.
- COMRIE Bernard 1985. *Tense*. Cambridge: Cambridge University Press.
- DE HAAN Ferdinand 1999. Evidentiality and epistemic modality: setting boundaries. *Southwest Journal of Linguistics* 18. 83-102.
- DE HAAN Ferdinand 2005. Coding of evidentiality. In HASPELMATH Martin, Matthew DRYER, David GIL & Bernard COMRIE (eds.). *World Atlas of Language Structures*. Oxford: Oxford University Press. 318-321.
- DELANCEY Scott 2001. The mirative and evidentiality. *Journal of Pragmatics* 33. 369-382.
- DIMMENDAAL Gerrit 2001. Logophoric marking and represented speech in African languages as evidential hedging strategies. *Australian Journal of Linguistics* 21. 131-157.

- DIXON R.M.W. 2003. Evidentiality in Jarawara. In AIKHENVALD Alexandra Y. & R.M.W. DIXON (eds.). *Studies in evidentiality*. Amsterdam: John Benjamins. 165-187.
- DIXON R.M.W. 2005. *A Semantic Approach to English Grammar*. Oxford: Oxford University Press.
- GRONEMEYER Claire. 1997. Evidentiality in Lithuanian. *Working Papers* 46. 93-112, Lund University, Department of Linguistics.
- HEINE Bernd & Tania KUTEVA 2002. *World Lexicon of Grammaticalization*. Cambridge. Cambridge University Press.
- HEWITT G.B. 1995. *Georgian: a structural reference grammar*. Amsterdam: John Benjamins.
- HOPPER Paul J. & Elizabeth C. TRAUOGOTT 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- IFANTIDOU Elly 1993. Parentheticals and relevance. *UCL Working Papers in Linguistics* 5. 193-210.
- JACOBSEN William H. Jr. 1986. The heterogeneity of evidentials in Makah. In CHAFE Wallace L. & Johanna NICHOLS (eds.) 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex. 3-28.
- JAKOBSON Roman O. 1957. Shifters, verbal categories, and the Russian verb. Cambridge, MA: Russian Language Project. Harvard University. [Selected Writings 2. 130-147].
- JESPERSEN Otto 1924. *The Philosophy of Grammar*. London: George Allen & Unwin Ltd.
- KANY Charles 1944. Impersonal *dizque* and its variants in American Spanish. *Hispanic Review* 12. 168-177.
- KIRSNER Robert S. & Sandra A. THOMPSON 1976. The role of pragmatic inference in semantics: a study of sensory verb complements in English. *Glossa* 10. 200-240.
- LAZARD Gilbert 1957. Caractères distinctifs de la langue tadjik. *Bulletin de la Société de Linguistique de Paris* 52. 117-186.
- LAZARD Gilbert 1999. Mirativity, evidentiality, mediativity, or other?. *Linguistic Typology* 3. 91-110.
- LAZARD Gilbert 2001. On the grammaticalization of evidentiality. *Journal of Pragmatics* 33. 359-367.
- LIDZ Liberty A. (2007). Evidentiality in Yongning Na (Mosuo). *Linguistics of the Tibeto-Burman Area* 30.2. Special issue *Evidentiality*, edited by AIKHENVALD Alexandra Y. & Randy J. LAPOLLA.
- LYONS John 1977. *Semantics*. Cambridge: Cambridge University Press.
- MATHIASSEN Terjie 1996. *A Short Grammar of Lithuanian*. Columbus, OH: Slavica Publishers.
- MATTHEWS Peter H. 1997. *Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- OLBERTZ Hella 2005. *Dizque* en el español andino ecuatoriano: *Dizque* conservador e innovador. In OLBERTZ Hella & Pieter MUYSKEN (eds.). *Encuentros y conflictos. Bilingüismo y contacto de lenguas en el mundo andino*. Madrid: Iberoamericana. 77-94.
- PALMER Frank R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.

- PLUNGIAN Vladimir A. 2001. The place of evidentiality within the universal grammatical space. *Journal of Pragmatics* 33. 349-358.
- ROBINS R.H. 1967. *A Short History of Linguistics*. London: Longmans.
- SAPIR Edward 1921. *Language*. New York: Harcourt, Brace and World.
- SQUARTINI Mario. *Forthcoming*. Lexical vs. grammatical evidentiality in French and Italian. *Linguistics*.
- TIMBERLAKE Alan 1982. The impersonal passive in Lithuanian. *Proceedings of the Eighth Annual Meeting of the Berkeley Linguistics Society*. 508-523.
- TRASK R.L. 1991. *A Dictionary of Grammatical Terms in Linguistics*. London: Routledge.
- TRAUGOTT Elizabeth Closs 1996. Subjectification and the development of epistemic meaning: the case of *promise* and *threaten*. In SWAN Toril & Olaf J. WESTVIK (eds.). *Modality in Germanic Languages: Historical and Comparative Perspectives*. Berlin: Mouton de Gruyter. 185-210.
- TRAVIS Catherine 2006. *Dizque*: a Colombian evidentiality strategy. *Linguistics* 44. 1269-1297.
- URMSTON James O. 1952. Parenthetical verbs. *Mind* 61. 480-496.
- VALENZUELA Pilar 2003. Evidentiality in Shipibo-Konibo, with a comparative overview of the category in Panoan. In AIKHENVALD Alexandra Y. & R.M.W. DIXON (eds.) 2003. *Studies in evidentiality*. Amsterdam: John Benjamins. 33-62.
- VAN DER AUWERA Johan & Vladimir A. PLUNGIAN 1998. Modality's semantic map. *Linguistic Typology* 2. 79-124.
- WEBER David J. 1986. Information perspective, profile, and patterns in Quechua. In CHAFE Wallace L. & Johanna NICHOLS (eds.) 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex. 137-155.
- WILLETT Thomas 1988. A cross-linguistic survey of the grammaticization of evidentiality. *Studies in Language* 12. 51-97.

## Notes

<sup>1</sup> This is the gist of the approach adopted in Aikhenvald (2004); also see a summary in Aikhenvald (2006a). The generalizations are based on the analysis of grammars of c. 600 languages (since the publication of Aikhenvald 2004, I have had access to further grammars). I avoid limiting myself to any artificially constructed samples of languages, since these are likely to engender skewed results.

<sup>2</sup> Statements to the contrary found in Palmer (1986), van der Auwera & Plungian (1998) and Willett (1988) are not borne out by the facts of languages, and are mistaken. See the arguments in de Haan (1999), Lazard (1999, 2001) and DeLancey (2001), and the general summary in Aikhenvald (2004:3-10). Some scholars whose experience is limited to a handful of familiar European languages tend to assume that evidentials are a kind of modal largely because of their absence in most major European languages, thus trying to explain an unusual category in terms of some other, more conventional, notion.

<sup>3</sup> The presence of such extensions does not make evidentials into 'modals' (contrary to some assumptions). This can be compared to the gender systems: in many languages feminine gender is associated with diminution, or endearment (see numer-

ous examples in Aikhenvald 2000), and masculine gender with augmentative; this however does not mean that gender is a type of diminutive or augmentative category. Readers should be warned against gratuitously dividing languages into those where evidentials have epistemic extensions, and those where they do not (as did Plungian 2001). As shown in Chapter 5 of Aikhenvald (2004), in the same language one evidential may have an epistemic extension, and another one may not.

<sup>4</sup> Contrary to assertions by Willett (1988) and others, an evidential may be within the scope of negation, as in Akha, a Tibeto-Burman language. An evidential can be questioned, as in Wanka Quechua. And the ‘truth value’ of an evidential may be different from that of the verb in its clause. Evidentials can be manipulated to tell a lie. One can give a correct information source and wrong information, as in saying *He is dead-reported*, when you were told that he is alive, or correct information and wrong information source, as in saying *He is alive-visual*, when in fact you were told that he is alive, and did not see him die. Two different information sources can be expressed within one clause (Aikhenvald 2004:93; 96-8). The grammatical category of evidentiality can be expressed through any of affixes, clitics, or auxiliary constructions. Linguists should be warned not to take seriously generalizations based on a limited sample such as those in de Haan (2005) which provides a highly inadequate coverage of formal means of marking evidentiality.

<sup>5</sup> Also see Kirsner & Thompson (1976) on a difference between “direct perception of a situation” and “deducing a situation” in their analysis of complements of sensory verbs in English. De Haan (§4, this volume) claims that “in perception verb constructions”, “combinations with infinitival constructions typically mean that the action is in the same deictic sphere as the speaker as in (a), while constructions with an embedded clause have no restriction on placement of action and speaker as can be seen in (b).

- a. I heard John cross the street.
- b. I heard, that John crossed the street.”

The reader should keep in mind that de Haan has overlooked a distinction between direct access to information source – actual hearing in (a) – and indirect, or hearsay, access to it.

<sup>6</sup> There appears to be a certain confusion concerning the term ‘paradigm’ as used by Pietrandrea (this volume). In conventional linguistic terminology (Matthews 1997:263), a paradigm is defined as “the forms of a given noun, verb, etc. arranged systematically according to their *grammatical* features” (italic mine: A. A.). The term “lexical paradigm” – as used by Pietrandrea, §1 of this volume, should be taken with a grain of salt, as an impressionistic way to refer to what are known as lexical sets, or semantic fields.

<sup>7</sup> For similar distinctions in terminological traditions other than the English, see Jespersen (1924:255).

<sup>8</sup> This is similar to how the linguistic literature on gender as a grammatical category does not discuss words for *man* and *woman*, or *bull* and *cow* in each particular language. Evidentiality in English has the same status as gender in Hungarian, or in Estonian. Of course, sex distinctions can be expressed in Hungarian and in Estonian if one wants to, but there is no grammatical category of gender. One can indicate information source in English, if necessary. But this is not grammatical evidentiality.

