Introduction

Evidentiality: How do you know?

Ad Foolen, Helen de Hoop and Gijs Mulder
Radboud University Nijmegen

1. Introduction to the concept of evidentiality

Evidentiality refers to the linguistic encoding of the speaker’s information source for a particular statement, a proposition. This statement can be seen as the primary information a speaker wishes to convey, to which the evidential expression is added to encode the type of evidence the speaker has to support their claim. Take for example the English utterance *John killed his wife*. The speaker of this utterance presents the information that John killed his wife as a fact, i.e., a true statement. However, a hearer may wonder: *How do you know?* Did the speaker witness the killing, did they read about it in the local newspaper, or did they infer it from the fact that John’s wife was found dead in suspicious circumstances? In order to answer that implicit question, the speaker may add an evidential expression, e.g., *John is said to have killed his wife* or *John killed his wife, I think*. Three types of evidential expressions can be distinguished, either conveying that the speaker witnessed the event (called direct evidentiality), was told about it, (reportative evidentiality) or conjectures in the here and now that it happened (inferential and assumed evidentiality). The latter two types of evidentiality are indirect since they express that the information is not first-hand.

Because indirect evidence entails that speakers do not have ‘hard evidence’ for their statement, there is overlap between the category of indirect evidentiality and epistemic modality (cf. van der Auwera and Plungian 1998; De Haan 2001; Cornillie 2009). The speaker suspects or believes that John killed his wife, but is not as certain as when they would have witnessed the murder with their own eyes. Boye and Harder (2009: 28) note that markers of direct evidentiality express a higher degree of certainty than markers of indirect evidentiality. Also, in languages with obligatory evidentiality marking, the verb in its unmarked form often expresses direct evidentiality (Aikhenvald 2004: 73).
Diverging subdivisions and some additional types of evidentiality have been suggested, in line with the distinctions that specific languages make. For example, inferential and assumed evidentiality can be distinguished, since inferences can be based on general or stored knowledge about the subject’s habits, i.e. indirect evidence, or on perceptual, especially visual (i.e., direct) evidence (cf. Hengeveld and Dall’Aglio Hattnher 2015). Aikhenvald (2004:2–3) argues that the difference between assumed and inferred evidentiality lies in the access to visual evidence in the case of the latter. An inferential evidential is “based on obvious evidence which can be easily observed (even if the event itself was not seen)” (Aikhenvald 2004:3). Also, she notes a difference in the degree of reasoning involved. That is to say, “the more the speaker has to rely on reasoning based on knowledge or on common sense, the more chance there is that the assumed evidential will be used” (Aikhenvald 2004:3). Müller (2013:212) notes that inferential evidentiality can be distinguished from direct visual evidentiality because in the case of inferential evidentiality the event itself is not directly witnessed. Thus, in English, when a speaker finds John’s wife’s body, an inferential evidential expression can be used, such as John must have killed his wife. Also, the speaker may actually see the result of an event, and utter I see that John cleaned up the place. In this case I see is not an expression of direct but of indirect evidence, as it is the result of the event that is witnessed, but not the event itself. Gipper (this volume) argues that Yurakaré -shi has in fact properties of both direct and indirect evidentiality, whereas Zeisler (this volume) argues that the notion of ‘direct’ knowledge for immediately observed facts is not exactly applicable to the Tibetic marker ḡ ḡ, for which reason she prefers to qualify it as ‘weakly indirect’. Bergqvist (this volume) stresses the importance of access to the evidence. As he points out, not only the speaker’s access plays a role in the meaning of evidentials, the hearer’s access can play a role as well. In this view on evidentiality, notions such as ‘subjectivity’, ‘intersubjectivity’ and ‘perspective’ come into the picture in a natural way.

Although some renowned specialists in the field have argued that only obligatory grammatical markers of evidentiality fall into the category of evidentiality in language (e.g., Aikhenvald 2004), many linguists nowadays take a broader perspective and also take lexical expressions into account. We concur with Marín Arrese’s (2015:212) assessment that “if we restrict the notion of evidentiality to cases of obligatory grammatical marking, we miss out on the expression of evidentiality in a significant number of languages, and we fail to adequately characterize and delimit the conceptual domain of evidentiality”. The aim of the present volume is to provide empirical evidence for the existence of evidentiality across the board: not just in a subset of languages, but presumably in all languages; not just in specific grammatical markers or designated lexical expressions, but also in grammatical categories that do not have the expression of evidentiality as their primary function,
such as future tense in French (Mari, this volume). Several chapters in the book demonstrate that the distinction between grammatical and lexical evidentials cannot be clear-cut. On the one hand, some of the evidential expressions discussed in this volume are partly grammaticalized. Prime examples are cognitive verbs that are frequently used with an evidential meaning, like Dutch *denken* ‘think’ (de Hoop et al., this volume) and Spanish *creer* ‘believe’ (Mulder, this volume). In first person singular present tense, these expressions exhibit clear signs of grammaticalization, i.e., they behave like a kind of adverbial expressions or particles, and do not constitute primary information in the sentence. On the other hand, even in grammaticalized form these evidential expressions retain aspects of their lexical meanings. The lexical verb *think* can be considered an agentive verb, while the lexical verb *believe* is characterized as an experiencer verb. This difference between the two lexical verbs in Dutch is still reflected in their evidential uses, as argued by de Hoop et al. (this volume). Another example is Gipper (this volume), who shows that the suffix *-shi* in Yurakaré has developed from a primary similarity meaning to an (uncertain) visual evidentiality meaning, and from there to an inferential evidentiality meaning. All these functions, including the primary meaning, are still present in Yurakaré. This is reminiscent of the semantic development of the Dutch evidential verb *schijnen* ‘seem’ as presented by de Haan (2000). De Haan argues that the meaning of *schijnen* ‘seem’ developed from the primary meaning ‘shine’ to the extended meaning ‘to become visible’ (uncertain visual evidentiality), and from there to indirect evidentiality meaning (hearsay). Whereas the auxiliary verb *schijnen* ‘seem’ in its evidential use is the most grammaticalized one, e.g., *Jan schijnt zijn vrouw vermoord te hebben* ‘John seems to have killed his wife’, the full lexical meaning is still present in modern Dutch, as in *De zon schijnt* ‘The sun is shining’. On the basis of this and other examples, de Haan (2000: 84) concludes that “[e]videntiality is a more widespread phenomenon than generally assumed. It (…) in fact can be found in most language families around the globe”. The chapters in this book further contribute to this insight. They present evidence for evidentiality in at least ten different languages of the world, viz. Dutch, Estonian, Finnish, French, German, Ika, Khalkha Mongolian, Spanish, Tibetan, and Yurakaré.

Although the evidence in the chapters comes from various unrelated languages, they sometimes show remarkably similar patterns. For example, several studies suggest that evidential expressions have unreliability or uncertainty as part of their meaning, or both (Tibetan, Yurakaré, Estonian, Spanish), which is of course not surprising in view of the overlap in functions between evidentiality and epistemic modality. Indeed, both concepts relate to the way languages express aspects of the ‘knowledge status’ of an utterance: Is what a speaker says private or rather shared knowledge, is it observed by the speaker or known on the basis of hearsay, is the speaker convinced that it is true what he says or are there reasons for doubt?
This domain, also labeled ‘epistemics’ (Heritage 2012; Lynch and Macbeth 2016; Lindström et al. 2016), covers phenomena that are referred to as evidentiality, modality, mirativity, egophoricity, epistemic attitude, (inter)subjectivity, etc. Besides epistemics, ‘stance’ is regularly used as a cover term, as for example in Bergqvist (this volume). The focus of the present volume is on evidentiality, but unavoidably, discussions regarding the connections between evidential and ‘neighboring’ meanings are part of the story.

In this short introduction, we do not endeavor to summarize the state of the art in evidentiality research. There are several excellent recent overviews, for example Aikhenvald (2014), Cornillie et al. (2015), and Marín Arrese et al. (2017). Below, we present an overview of the chapters in this volume, and link them to other recent literature when relevant. The present volume brings together a variety of methods to study evidentiality. Contributors make use of intuitions, corpus data, grammars, historical sources, and general linguistic knowledge. For example, Brosig (this volume) explores subtle meaning differences in a corpus of spoken conversation, Jaakola (this volume) examines hearsay evidentiality in Internet discussions on a newspaper site, and Tamm et al. (this volume) conducted an experiment with Estonian children to study their interpretation of the evidential morpheme -vat.

Different sources are accessed in different ways and are more or less reliable, leading to certain claims, inferences and assumptions. If challenged, an important way of defending oneself as a researcher is to explicate the reliability of the access to the data. The same holds in everyday language use, when confronted with the question How do you know?

This question is indeed ambiguous, relating to the phenomenon of evidentiality itself as well as to the researchers who study evidentiality. Foregrounding the aspect of access to the evidence combines the theoretical question of what type of linguistic expressions belong to the cognitive-functional domain of evidentiality with the researcher’s access to phenomena of evidentiality across languages.

2. The chapters in this volume

Apart from the introductory chapter, this volume collects eleven more chapters related to the question How do you know?, addressing various issues of evidentiality. The first part of the book addresses the question What do we know? It investigates the relation between knowledge and evidence from both the speaker’s and the hearer’s perspective. The first chapter by Henrik Bergqvist outlines a general perspective on evidentiality and the relation between evidentiality and epistemic modality. Epistemic modality is used to weaken the factuality of a proposition, which is not necessarily the case for evidentiality. This relates to the second chapter by Benjamin
Brosig, which focuses on the difference between factuality and evidentiality in terms of established and non-established knowledge in Khalka Mongolian. It is debatable whether an expression of established knowledge such as the evidential *sang* in Khalkha Mongolian can be considered an evidential or not. By contrast, both *-laa* and *-jee* are true evidentials, and refer to the accessibility of non-established direct or indirect evidence. Brosig emphasizes that the evidence in case of these markers has to be newly acquired; it cannot be common knowledge or assimilated information (as in the case of *sang*).

Evidential expressions based on or derived from verbs of cognition such as *think*, *believe*, and *know* are the topic of Chapters 3–5. Chapter 3 by Helen de Hoop, Ad Foolen, Gijs Mulder, and Vera van Mulken, and Chapter 4 by Gijs Mulder zoom in on inferential evidentiality of expressions such as *I think* and *I believe* in Dutch and Spanish. While the evidential expressions *I think* and *I believe* are clearly subjective, Minna Jaakola’s chapter on Finnish evidential adverbs in argumentative texts analyzes the meaning of three adverbs in her native language Finnish: *tietääkseni*, *tiettävästi* and *tietysti*, including a synonym of the latter adverb *tietenkin* and its negative form *eitietenkään*. All these forms are derived from the verb *tietää* ‘to know’.

The second part of the book, entitled *When do we know?* investigates the accessibility of evidence as well as the development of evidentiality and extensions of evidential meanings in time. The sixth chapter, i.e., the first chapter of the second part, by Anne Tamm, Reili Argus, and Kadri Suurmäe, deals with the acquisition of a marker of evidentiality in Estonian. Experimental research shows that children initially interpret the hearsay evidential as a generic marker. One explanation put forward by the authors is that the experimental context does not provide the children with any clues that the new morpheme they encounter expresses hearsay evidentiality. Because the children do not have access to evidence for its status as a marker of hearsay evidentiality, they interpret the morpheme as a marker of general knowledge.

The link between modality and evidentiality is a recurring question in the literature on evidentiality. Although the categories are clearly distinguishable in and of themselves, they are often expressed by the same forms in languages. This is illustrated by the deontic-modal and the hearsay-evidential uses of the verb *sollen* in German, investigated by Jeroen Vanderbiesen in the seventh chapter.

Alda Mari in Chapter 8 examines the evidential use of future tense in French, and discovers that it occurs even in cases when the speaker knows that the proposition is true, as long as the evidence can only be accessed by the hearer at a future point in time relative to some other point in time. Whereas hearsay evidence as well as evidence indicated by verbs of cognition such as *believe*, *think*, and *know* is clearly subjective, i.e., inaccessible to hearers, direct evidence is typically intersubjective,
and accessible to hearers. Direct evidential expressions often develop from verbs of perception, e.g., see, hear, smell, but some go back to other notions, such as existential verbs in Tibetan, as discussed in Chapter 9 by Bettina Zeisler, and simulative 'like' in Yurakaré, as discussed in Chapter 10 by Sonja Gipper.

Whereas Mari’s study on the future in French examines the question at what time evidence is accessible for verification, Zeisler points out that existential notions like ‘limited stay’ versus ‘a more permanent way of being’ provide the basis for different ways of representing the truth: lasting acquaintance vs punctual observation yield different modes of access. Another issue that relates to type of access is the perception of the evidence, as studied by Gipper. Gipper shows that expressions of direct evidentiality can develop into expressions of indirect evidentiality, thereby weakening the factuality of the proposition, and thus they are related to epistemic modality. Collectively, the chapters in the second part of the volume illustrate the importance of the notion of access to the evidence in the study of evidentiality. The final chapter by Seppo Kittilä, Lotta Jalava and Erika Sandman highlights the need for applying various methods to get access to the evidence for the existence of the category of evidentiality across languages.

2.1 Part I: What do we know? Knowledge and evidence

Henrik Bergqvist’s chapter “Evidentiality as stance: Event types and speaker roles” considers evidentiality in the context of a model of the speech situation. The main question is, which dimensions of the speech situation have to be postulated in order to build a model that can accommodate the big variety of phenomena of evidentiality in the languages of the world. Bergqvist selects challenging cases from the existing literature to argue for a sophisticated model of the speech situation. He shows that besides the perspective of the speaker, that of the hearer is needed as well, because some evidentiality distinctions take the two different perspectives into consideration. Moreover, Goffman’s (1981) distinction between author, animator and principal is considered relevant as well as Jakobson’s (1957) notion of shifter. The model offers a framework for (re)considering the relation between evidentiality and its close companions, epistemic modality and egophoricity (see also the discussion in Section 6.2 in San Roque et al. 2018). Bergqvist shows that in cases where more than one of these categories are present in the same utterance, they take specific narrow and wide scope relations, which in turn provides arguments for their place in the model of the speech situation (for the relevance of scope phenomena as a source of evidence for evidentiality phenomena see also Hengeveld and Dall’Aglio Hattner 2015).
From this very succinct indication of Bergqvist’s paper, it will be clear that it well serves as a starting point for the present volume. This kind of modeling can only be done on the basis of detailed descriptions of evidentiality phenomena in languages of the world.

Benjamin Brosig in his chapter “Factual vs. evidential? The past tense forms of spoken Khalkha Mongolian” studies the evidential forms in spoken Khalkha Mongolian, focusing on the use of three evidential morphemes that are used in past tense and occasionally future contexts: -sang, -jee, and -laa. These forms have been studied before, but Brosig points out that the data in those studies were based on questionnaires or small corpora, often taken from specific genres like literary prose. Such data do not give a proper insight in the subtleties of use of these morphemes in free conversation, leading to distorted descriptions in existing grammars.

The methodological approach in the present study is an interesting one. Brosig used about 10 hours of relatively free conversational data from unscripted television programs, which were transcribed and annotated. In a following step, passages containing a given tense-aspect-evidentiality form were presented to native speakers. For each form the judgments of at least ten informants were collected, adding up to 285 hours of ‘data sessions’ for the entire finite inventory. The informants were not only asked what the utterances meant in the specific contexts, but also whether the evidential form could be replaced by some specific others, and if yes, what the difference in meaning would be. The meanings and strategic uses of the morphemes turn out to be very subtle, but on the most general level, Brosig summarizes the meanings of the three morphemes at the end of Section 1 as follows: “[T]he basic distinction is drawn between established knowledge (-sang), which may implicate a claim to validity, vs. non-established (mostly new) direct (-laa) or indirect (-jee) knowledge, which may relegate the evaluation to the addressee.”

The research presented in Helen de Hoop et al.’s chapter “I think and I believe: Evidential expressions in Dutch” also involves evidence that is newly acquired, more in particular what Plungian (2001) called ‘endophoric evidentiality’. Jing-Schmidt and Kapatsinski (2012: 348) characterize endophoric evidentials as “that kind of evidential in which a direct personal inner state constitutes the source of information.” The study also relates to what Thompson (2002) called “epistemic/evidential/evaluative formulaic fragments”, first person forms of verbs of cognitive attitude like think, believe, guess, mostly used in the present tense. They can take variable positions, at the beginning, middle or end of the clause that contains the main message. De Hoop et al. study the first person uses of the Dutch verbs geloven ‘to believe’ and denken ‘to think’, both in the present and past tense.

De Hoop et al. use a corpus of Twitter data to study two expressions of inferential evidentiality in Dutch, namely ik geloof ‘I believe’ and ik denk ‘I think’, and their inverted counterparts geloof ik ‘believe I’ and denk ik ‘think I’. These expressions do
not only have evidential interpretations, but they can also be interpreted literally in certain contexts, and think can have the additional interpretation of intentionality, as in I think I’ll have a beer. Although the two expressions can both be considered instantiations of inferential evidentiality in Dutch, they crucially differ as well. Whereas denken ‘think’ is an agentive (cognitive) verb that takes an Agent subject, geloven ‘believe’ is a truly cognitive verb, the subject of which is an Experiencer. This is argued to make the evidential use of ik geloof ‘I believe’ less rational, or more impressionistic, and therefore less inferential than ik denk ‘I think’. It seems that ik geloof ‘I believe’ has retained part of its literal meaning, and therefore the type of evidence that it refers to is more impressionistic than the type of evidence carried by the inferential evidential use of ik denk ‘I think’. Finally, de Hoop et al. argue that the difference in the semantic role for the subject can be used to account for the fact that denken ‘think’ can maintain its evidential reading in the past, as in ik dacht ‘I thought’, whereas this does not hold for geloven ‘believe’ in Dutch. That is to say, ik geloofde ‘I believed’ can only refer to an impression the speaker had in the past, but this cannot be put forward as inferential evidence at the time of utterance. A speaker cannot put forward an impression they had in the past, as evidence for a statement in the present, as this type of evidence is no longer accessible.

Whereas I think in Dutch is more frequent than I believe, this is the other way around in Spanish, as claimed by Gijs Mulder in his chapter “(Yo) creo que as a marker of evidentiality and epistemic modality: Evidence from Twitter”. Mulder does not compare Spanish creo ‘I believe’ to pienso ‘I think’, but instead focuses on creo ‘I believe’. As in de Hoop et al.’s chapter, data for the present study were taken from Twitter. Although tweets are often stand-alone utterances, they typically have a strong dialogic orientation, either being part of a ‘thread’ or reacting to an actual topic that is ‘in the air’.

Mulder based his analysis on 180 tweets in first person uses of creer ‘to believe’ in the present tense, both with spelled out first person yo, and with null subjects (for a specific study on the difference between these variants see, among others, Posio 2011), focusing on the evidential use of (yo) creo ‘I believe’ in Spanish. The ‘literal’ and other non-evidential uses of the verbs were sorted out manually.

Mulder shows that (yo) creo usually combines with a complement clause starting with the complementizer que ‘that’. The formulaic phrase also occurs in medial and final position as a parenthetical construction, which can be considered a clear sign of its grammaticalization. Mulder briefly investigates the uses of two other frequent evidential expressions in Spanish, pienso ‘I think’ and me parece ‘it seems to me’, but concludes that (yo) creo ‘I believe’ is by far the most frequent of the three evidential/epistemic expressions in Spanish. Whereas (yo) creo can be considered an instantiation of inferential evidentiality, it can also take up other non-literal functions, to wit epistemic modality, opinion, and intention. The fact that (yo)
creo can have other non-literal readings besides evidentiality can be considered an additional sign of its grammaticalization in the language. Mulder’s corpus study reveals that whereas the evidential interpretation is clearly dominant for creo que (without the subject pronoun), expressing an opinion is the most frequent use of yo creo que (with the explicit subject pronoun). Mulder explains this by arguing that in evidential uses speakers do not want to put themselves in focus, because evidentials are used to provide secondary information about the source of information. Clearly, the source of information can be the speaker, as in the case of inferential evidentiality. However, the focus of attention should be on the primary information, i.e., the propositional content that the speaker wishes to convey. If yo ‘I’ is added, the focus of attention may shift to the speaker, resulting in a subjective viewpoint, hence the expression of an opinion.

Minna Jaakola’s chapter on Finnish evidential adverbs in argumentative texts analyzes the meaning of three adverbs in her native language Finnish: tietääkseni, tiettävästi and tietysti, including a synonym of the latter adverb tietenkin and its negative form ei tietenkään. All these forms are derived from the verb tietää ‘to know’.

Jaakola selected a little over 700 texts from internet discussions on the website of a Finnish newspaper in which one of the three adverbs derived from know occurred. The question raised was which role these expressions play in the text. The analysis is inspired by ethnomethodology, in particular by ‘epistemics’ (cf. Heritage 2012). The analysis of the texts is interpretative, making use of cues in the context, like other adverbs and conjunctions, resulting in the claim that the three adverbs fulfil different functions: the core meaning of tietysti and its variants is ‘as is normal to know, of course’, and they are used in sentences that serve as justifications or hedges for the main claim in the text. In contrast, tiettävästi “expresses that the writer has access to information, and the reader can likewise have that access, or he understands how the writer uses the access” (Section 6.2). And the third form, tietääkseni, “emphasizes the writer’s own access to knowledge. The reader is not expected to have a similar access, but tietääkseni opens up the dialogic frame of negotiating different points of view” (idem). An advantage of the interpretative, ‘close reading’ approach of the data is that pragmatic functions of the use of these evidential markers in the text can be discovered, among them irony when using tietysti ‘of course’.

2.2 Part II: When do we know? Accessibility of evidence in time

Anne Tamm et al. begin their chapter “Uralic perspectives on experimental evidence for evidentials: Early interpretation of the Estonian evidential morpheme” with an overview of evidential systems in the Uralic languages. They point out that
there is a continuum from east to west, with the eastern having a richer and the western languages having a poorer evidential system. Estonian, belonging to the western group, has only one grammaticalized optional evidential morpheme, the hearsay evidential -vat.

In the second part of their chapter, Tamm et al. focus on the acquisition of -vat. Their method is experimental, the only one in the present volume. Most acquisition studies focus on the acquisition of obligatory evidentials (such as in Turkish, e.g., Ünal and Papafragou 2016), whereas -vat is optional. In Turkish, which has obligatory evidential marking in the past, correct production (around age 3) precedes correct comprehension (around age 6 or 7) (Ünal and Papafragou 2016). Estonian -vat is optional, and moreover it is extremely rare in conversations with children. Therefore, Tamm et al. expect its evidential interpretation to be acquired very late, not before the age of 6. In order to find out more about the development of learning this new grammatical item, subjects of their study are four- and six-year-old Estonian children. In an experiment, the effect of the use of -vat on their exploratory play with magnetic objects is investigated. The exploratory play was stimulated either with a ‘neutral’ utterance by an adult, meaning ‘The blicket (a block to play with) is magnetic’ or with the same utterance, but now containing the hearsay evidential morpheme -vat. The researchers found that the four-year-olds showed increased testing blickets for magnetic properties in reaction to the utterance with the evidential, whereas there was no significant difference between the two types of stimuli for the six-year-olds. Based on this difference in behavior, the authors consider two explanations. First, increased exploratory play by four-year-olds could be evidence that children initially interpret the evidential as a generic marker. Second, because the evidential morpheme is a ‘new’ grammatical morpheme to the four-year-olds, their increased exploratory behavior may indicate that they test out its meaning. However, in this experimental context they will not be able to figure out that the morpheme is used to express hearsay evidentiality, simply because they do not have the evidence needed to draw such a conclusion.

Hearsay evidence is also the topic of Jeroen Vanderbiesen’s chapter “Reportive sollen in an exclusively functional view of evidentiality”. Vanderbiesen analyzes 200 German text fragments, mainly taken from newspapers, in which the verb sollen has an evidential hearsay meaning. Sollen has also other meanings, most of them deontic, some of them intentional. Vanderbiesen selected the hearsay evidentials from all the occurrences of sentences with sollen in the texts. Some ambiguous cases, which can be interpreted in an evidential or intentional way, are discussed in the chapter.

In corpus linguistics, a distinction is made between corpus-based and corpus-driven methods. Vanderbiesen’s study belongs in the corpus-based category, in the sense that he uses corpus data to explore a research question, in his case the
question what the constructional environment of evidential uses of *sollen* looks like. It turns out that the subject is typically 3rd person and a source of the information is absent. If a source is mentioned, it is rather an institution or other 'abstract' source, not a specific person. Vanderbiesen takes the latter finding as support for his theoretical claim that a strict distinction should be made between hearsay evidentiality, which he calls ‘reportive’ (others use ‘reportative’, cf. Faller 2017; Guardamagna 2017) and quotatives, in which information is attributed to a specific source. The function of quotatives is attribution, the function of reportives is to provide justification for the information in the text. When *sollen* is used, the justification lies in the fact that ‘it has been heard, read’.

*Alda Mari*’s chapter “The French future: Evidentiality and incremental information” investigates the relation between modality and evidentiality in her analysis of the French future tense.

Mari uses constructed examples as data, consisting of a specific utterance in a sketched scenario. These data, however, were submitted to a panel of native speakers who evaluated the data on acceptability. As it turned out, the ratings were clear-cut, so that the data can be considered as realistically representing the use of the French future.

Mari proposes to analyze the French future tense as a modal-evidential that shifts the time of evaluation of the modal component forward. Crucial in this new approach is the notion of ‘future verification of evidence’. That is, if a customer asks for certain shoes, a French shop assistant may tell them that these shoes “*will be*” on the shelf in the back of the store. The future in this case does not indicate that the shoes will (only) be on the shelf at a later time, nor does it imply that the speaker is not certain about whether the shoes are on the shelf or not (the speaker knows they are). Rather, the French future tense indicates that the truth of the embedded proposition (“the shoes are on the shelf”) are to be verified by the hearer at a point in time after the time of utterance. Note that this reading is not available for the future tense in Italian, nor for the modal *devoir* ‘must’ in French. Mari argues that the French future always maintains its core meaning of future verification of evidence. However, the time with respect to which this time of verification is situated in the future, may differ. For example, if somebody utters “You must have left your watch in the hotel”, using a French future (instead of modal *must*), then the time of the event precedes the time of the utterance (it is in the past), while the time of verification is in the future of the utterance time. Thus, (only) at a future time with respect to the time of utterance it can be verified whether the addressee had left their watch in the hotel. The time at which the future verification takes place is not necessarily after the time of utterance, however. It can also coincide with the time of utterance. When somebody witnesses Nadal’s expected and predicted winning of the Roland Garros final, they may utter in French that Nadal “will have won”
again, meaning that Nadal has won indeed, not that he will have won at some time in the future, nor that the winning can only be verified at some later point in time. Instead, future sets the time of verification, which in this case is a future time with respect to a time in the past of the time of utterance, and which therefore can be (and indeed is) the time of utterance itself. Accordingly, in this case the type of evidence for verification is direct, as we witness Nadal’s winning.

Bettina Zeisler’s chapter “Evidence for the development of ‘evidentiality’ as a grammatical category in Tibetan” focuses on the question how the evidential system in Tibetan originated. The core of this system exists of two markers, ḥdug and yod. They both developed from copula verbs, the first of which had a semantic load of transitory duration, whereas the second rather implied punctual or general existence. The transitory meaning led to an evidential marker with a semantic load of uncertainty, whereas the punctual/general meaning was the basis for an evidential that is typically used in authoritative, generally applicable statements. It is interesting to see that the transitory meaning is associated with visual perception, which, like in Yurakaré (see the chapter by Gipper) is felt as ‘only seen’, thus ‘uncertain’. In other words, the association of visual evidentiality and modal certainty has no universal status.

The diachronic study of Tibetan evidentials is based on written texts, which are available from the seventh century CE onwards. Zeisler is well aware of the genre-specificity of evidentiality. Many Classical Tibetan texts are doctrinal in nature, whereas evidentiality typically occurs in dialogical genres, in which discussions and justifications play a role (cf. González et al. 2017). Two biographies, one from the thirteenth and one from the fifteenth century come closest to the required dialogic property. Zeisler’s analysis of these two texts shows that in the older one the evidential system started to develop and that it was not fully developed yet in the second text. A survey of the occurrence of (variants of) the two markers in different varieties of present day spoken Tibetan leads to the additional finding that the evidential system must have spread from east to west.

Sonja Gipper in her chapter “From similarity to evidentiality: Uncertain visual/perceptual evidentiality in Yurakaré and other languages” focuses on how evidentiality relates to perception of similarity, vision, impression. The Yurakaré morpheme -shi has three meanings: simulative ‘like’ (A is like B), uncertain visual/perceptual evidentiality and inferential evidentiality. While inferential evidentiality is a well-known category, uncertain visual/perceptual evidentiality is a less established type of evidentiality. Visual evidentiality is typically considered to go along with modal certainty, but apparently, there are also evidential markers that indicate that the evidence is based on visual (or other perceptual) input, but that the speaker is nevertheless uncertain about its really being the case.
Gipper’s main hypothesis is that the uncertain visual/perceptual and inferential meanings developed diachronically in this order. How to provide evidence for such a claim if no diachronic data are available for a language, like in the case of Yurakaré? Gipper provides circumstantial and converging evidence. The first type of evidence is sociolinguistic: looking at data from three different dialects, it turns out that the inferential use only occurs in one dialect and in particular in the language of younger speakers of that dialect. Secondly, the diachronic development of evidential meanings of the German verb *scheinen* ‘to appear’ is very similar to the scenario postulated for Yurakaré. Although the point of departure is different (the original meaning of *scheinen* is ‘to shine’), the evidential meanings are comparable to those of *-shi*, namely uncertain visual/perceptual and inferential. As Diewald and Smirnova (2010) have shown, these meanings developed successively in Middle and New High German. Thirdly, Gipper shows that in several other languages an evidential item can be identified with the same range of meanings, namely uncertain visual/perceptual and inference.

Seppo Kittilä et al., focusing on hearsay evidentiality, show in their chapter “What can different types of linguistic data teach us on evidentiality?” that different methods of searching evidence for evidentiality each have their own merits and disadvantages. Reference grammars, for example, provide information about grammatical (obligatory) evidentiality marking, but not about lexical means of evidentiality, nor about secondary functions of evidentiality markers. Corpora can inform us about how often grammatical as well as lexical evidentiality markers occur in a language, but corpus data are only production data, and we would still need native speakers’ intuitions to obtain information about interpretations. For comparative studies across languages, translation corpora can provide us with additional data. Such corpora may include work that has been translated from languages with grammatical evidentiality to languages with lexical evidentiality, and the other way around. Internet chat data can also provide us with more data on evidentials, since this type of language is closely related to informal spoken language. However, Kittilä et al. point out that the social norms that apply to chat language as compared to face-to-face conversation may differ, resulting in different patterns of evidentiality marking in different genres. Fieldwork is useful in collecting data from lesser-studied languages, and it includes different methods, such as elicitation by translation, recording of spontaneous conversations, collecting narratives, and elicitation of grammaticality judgments. Finally, questionnaires can be used for the study of certain characteristics of individual languages as well as for cross-linguistic comparison. A potential problem with using questionnaires is that they provide us with information about what is possible or not in a language, but this may deviate from actual language use. Kittilä et al. conclude that the different methods...
complement each other, and the best way to study evidentiality in language would be to combine them.

3. Conclusion and Outlook

The contributors to the present volume provide evidence for the pervasiveness of evidentiality in languages across the board. We have chosen to aim for empirical evidence, i.e., to focus on real language data rather than theoretical argumentation or modeling. We believe that reflection on methodological issues is especially important in the study of evidentiality and that advances in methodology constitute an important precondition for gaining new theoretical insights into the coding, meaning, and function of evidentiality across languages. In other fields of investigation, such as argument structure, semantic roles, the conceptualization of space and time, etc., access to reliable linguistic data is relatively straightforward, but the identification of evidential meanings is more elusive, both for the native speaker and the linguist. Intuitions are more difficult to explicate, and the use of evidential expressions strongly depends on the context. Certain genres focus on ‘news’, others on retelling general knowledge, or reporting personal experiences. Irrespective of whether one uses corpus or experimental data, questionnaires, or introspectively gained data, the contributions in this volume show that the study of evidentiality cannot be restricted to isolated utterances.

As pointed out above, analysis of evidentiality requires special methodological awareness on the part of the researcher. Indeed, the studies in this volume show that a range of methods should be used in the exploration of evidentiality phenomena. This range is as varied as the types of access, direct or indirect, that language users have as evidence for their statements, and as varied as the linguistic means, grammatical or lexical, that are available in natural languages to indicate the source of evidence. Future research on evidentiality should pay special attention to the methodological requirements outlined in this chapter.

References

Introduction


All rights reserved
