

What can different types of linguistic data teach us on evidentiality?

Seppo Kittilä, Lotta Jalava and Erika Sandman

University of Helsinki

In this chapter, different types of data for evidentiality studies are discussed. We first discuss reference grammars, which are necessary for any cross-linguistic study of evidentials. This is followed by a discussion of (different types of) usage-based data, as well as natural language data and stimulus-based data collection used in linguistic fieldwork. We end our discussion by examining data collected by means of questionnaires and acceptability judgements. It is shown that all the types of data discussed are relevant for a full understanding of evidentiality, but they differ in their contribution and complement each other. For example, usage-based data is necessary for studying frequencies in languages, while reference grammars reveal what is common across languages.

Keywords: usage-based data, stimulus-based data collection, questionnaires, reference grammars

1. Introduction

Evidentiality comprises a range of different aspects. First of all, languages differ according to whether they have evidentiality as a grammaticalized category or not. Examples of languages with grammaticalized evidentiality include, e.g, Wutun, Lhasa Tibetan, and Tariana. Most Indo-European languages lack evidentiality as a grammatical category. Languages with

grammaticalized evidentiality have a varying number of evidential markers (ranging from two to five), while languages with non-grammaticalized evidentiality can be further subdivided according to the nature of the elements which are used for expressing evidentiality. In these languages, evidentiality may be expressed by evidential strategies (such as using certain tenses like compound tenses in Finnish), and/or by making use of dedicated lexical items for this function (see, e.g. Jaakola, this volume for a detailed discussion of some Finnish evidential adverbs).

Second, even though most of the earlier research on evidentials has focused on the source of information the speaker has for his/her statement (see Willet 1988; Plungian 2001; Aikhenvald 2004), evidentials can also express an array of other functions (see e.g., Curnow 2001, 2003). Examples are the expression of (decreased) volitionality and the use of evidential markers as discourse particles. Moreover, the interplay of evidentials with intersubjectivity has been a topic of numerous studies in evidentiality recently. Evidentials may be used to attend to distributions of knowledge among participants in interaction, such as expressing shared vs. non-shared information or accepting/declining invitations (see e.g., Fox 2001; Michael 2008; Gipper 2011; Hayano 2013). These studies have broadened our understanding of evidentiality in relation to other domains.

One of the consequences of the variation described above is that evidentiality can be studied from a variety of perspectives, employing many types of data. This chapter gives an overview of the applicability of these types of data to the study of evidentiality, and discusses what different approaches have taught us on evidentiality. This is achieved by discussing some central data collection methods in light of recent studies of previously lesser-studied aspects of evidentiality. Questionnaires can also be used as part of fieldwork, but here, we handle questionnaires as an individual data collection method. The goal of this chapter is to study the types of data used in evidentiality studies as exhaustively as possible. We also aim to combine

cross-linguistic studies and studies of individual languages. Therefore, we need methods for collecting data from a large number of languages (reference grammars) and from individual languages (corpora, field methods, and questionnaires). The examined types of data include reference grammars, usage-based data and corpora, questionnaires and different kinds of linguistic field methods (including elicitation, staged communicative events, collecting and analyzing texts and participant observation). Although questionnaires can also be used in fieldwork, here we will treat questionnaires as an individual data collection method. It is important to bear in mind that studying evidentiality differs from many other domains of grammar, due to its intimate connections with pragmatic and intersubjective aspects. This makes evidentiality much more genre- and context-specific than, for example, the studies of argument marking.

As noted, the goal of the chapter is to discuss different types of data used for studies on evidentiality, and what they can teach us about the concept of evidentiality. Due to limitations of space, we can only scratch the surface in many cases, but we nevertheless hope to discuss the most central aspects of all the examined types of data. Even though the different types of data are discussed separately below, it should be borne in mind that each of these types have their strengths and weaknesses, and they should be seen as complementary. All types of data should be considered if we wish to arrive at a comprehensive understanding of evidentiality. Therefore, the interplay between different types of data is also scrutinized. We start the discussion by examining reference grammars, followed by usage data, different kinds of field methods and questionnaires.

2. Reference grammars and the typological perspective

Our understanding of evidentiality as a linguistic category is to a large extent based on typological studies (nowadays also called diversity

linguistics). During the last three or four decades, typological studies have shown us what kind of evidential systems occur cross-linguistically. These studies have also shown how the semantics of evidentials attested in the world's languages depend on whether the evidential system of a given language is large or small. For example, in a system non-firsthand vs. everything else, the former may include hearsay evidence and also inference, while in larger systems these are explicitly distinguished (Aikhenvald 2004: 186). Typological studies also show that large evidentiality systems (with five or more categories) are rather rare across languages. Typological approach is very clear in Chafe and Nichols (1986), Willet (1988), Aikhenvald (2004), all of which serve as a basis for many recent studies of evidentiality.

First, reference grammars constitute the most important source of data for any cross-linguistic study on any topic in linguistics, including evidentiality. Consulting reference grammars is the only possible way of acquiring an overview of the cross-linguistic variation attested for in the studied phenomenon, whether this is argument marking, phonology, evidentiality, or any other aspect of grammar. Of course, this could also be achieved by consulting informants, but the number of languages for which informants are available is limited. Reference grammars are thus an essential part of any study whose goal is to arrive at a comprehensive understanding of evidentiality. A brilliant example of this kind of study is provided by Aikhenvald's (2004) cross-linguistic study of evidentiality that is based on the reference grammars of over 500 languages. The study does not only provide us with a typology of evidentiality systems, but many other aspects of evidentials (e.g., their development) are also discussed. Typological studies reveal what kind of evidentials are common across languages and to what extent their semantic values are comparable across languages, but they do not necessarily tell what is common within a single language. (Large corpora would be needed for these, but are especially hard

to construct for lesser-studied languages.) Moreover, reference grammars are focused on what is grammatically obligatory in the described language, which has the consequence that certain evidential types are prevalent in language descriptions, while, e.g., lexical evidentials are completely neglected. As a very illustrative example of this kind of approach we can mention Aikhenvald's seminal study of evidentiality.

The prevalence of grammaticalized evidentials may have the consequence that grammar writers only look for the categories that have already been shown to exist in other languages neglecting non-grammaticalized evidentials. This may give a false picture of what actually occurs in the world's languages, because certain novel categories may never be recognized. For example, the effects of intersubjectivity on the use of evidentials have been recognized only recently (Bergqvist 2017 and this volume). From this it follows that data for these phenomena cannot be found in older grammars (or other studies), simply because the authors have not looked for this phenomenon. Similar problems apply to older grammars; they don't have any references to evidentiality, because the notion has not been known to exist, and/or it may have been discussed under a different topic. Instead, evidentials may have been discussed under modalities, moods, or event tense, e.g., because they may occupy the same slot or they may be parts of the same paradigm (see Boas 1911 for Kwakiutl and Tsimshian). In general, newly discovered categories lag behind and they become a part of grammatical descriptions only later.

The grammatical differences between languages are also worth considering in this respect. In some languages, information source is expressed formally by the same constructions that are also used to express epistemic modality, management and access to knowledge (such as perspective, e.g., egophoric systems of Tibetan languages) and clause type (e.g., Ecuadorian Siona, see Bruil 2014). If a typology of evidentials were based on data from these kinds of languages, our understanding of the

notion could be very different from how evidentiality has been traditionally defined.

Moreover, due to limitations of space, reference grammars usually focus on the basic cases and the most frequent uses of the discussed evidentials, which means that they cannot provide us with a very broad picture of various semantic and pragmatic functions of evidentials. Descriptive grammars are based on small sets of data, and the nature and size of data vary dramatically, depending, e.g., on how the material has been collected. For example, using folklore differs greatly from using naturally occurring conversation as data. Besides, usually not all possible uses of evidentials are discussed due to limitations of space, or lack of suitable data. However, grammars give us information on what the role of a specific evidential is in the evidential system of a language; a general hearsay evidential in a small system may also be used for any kind of second-hand information, while a larger system usually has specific evidentials for other types of second-hand information. The terminology used in different grammars also varies, making it hard to pinpoint the exact meaning of a given evidential in a language. e.g., the terms *hearsay* and *reported*, and the terms *visual* and *direct* may refer to the same category depending on the language and grammar (see Keinänen 2017 for a more detailed discussion).

As noted, reference grammars usually focus on what is obligatory in a language, which has the unfortunate effect that optional means of expressing evidentiality are not necessarily discussed in them. From this it follows that reference grammars only provide us with (cross-linguistically reliable) data on the kinds of evidentials languages have grammaticalized, while their full evidentiality potential may not be revealed. For instance, the use of hearsay evidentials as discourse particles may not be discussed due to limitations of space. Moreover, the (hypothetical) picture that languages tend to have general hearsay evidentials, but especially smaller systems lack quotatives may be distorted by the fact that quotatives may be attested in languages

only as a lexical element. Put another way, grammars do not usually render it possible to study languages where evidentiality is not obligatorily expressed.

To sum up. Despite their limitations regarding the actual use of evidentials and the potential lack of any discussions of lexical evidentiality, reference grammars do give us invaluable information on what kind of evidentials exist in the world's languages. Moreover, reference grammars, and the overview of cross-linguistic variation they provide can serve as a basis for further studies on evidentiality in individual languages. For example, does cross-linguistic frequency correlate with a high frequency in an individual language? Hearsay evidentials are cross-linguistically very common, and there are many languages whose only evidential is some kind of hearsay evidential. However, whether this actually correlates with how frequent hearsay evidentials are in a given language cannot be reliably studied based on grammars only. Closely related to this, cross-linguistic studies may aid us in finding novel categories in languages where evidentials are not grammatically obligatory, and they may also aid field linguists to search for these categories in the languages of their expertise. Cross-linguistic variation can also be regarded as a window to how language works. Needless to say, mere linguistic evidence is not enough to say anything definite about psychological reality and about how humans actually process the information they have available for their statements, but broad cross-linguistic studies of any topic in linguistics may serve as a starting point for psycholinguistic studies of the same topic for verifying whether the hypotheses made actually hold.

3. Usage-based data

Data that represent authentic language use in real situations are necessary for arriving at a complete picture of the use of evidentials in any given

language, as the researcher's and/or consultant's own intuition always has its limits and we all may interpret utterances differently. This means that not everything ends up being discussed in a reference grammar. Studying usage-based data both qualitatively and quantitatively provides us with insights into how evidential expressions are used in different types of spoken and written language, different genres, and contexts. Closely related to this, corpora consisting of natural language data are necessary for presenting any frequency information on evidentials (for example, the differences between grammaticalized evidentials and evidential strategies used to express similar functions). Large-scale corpora usually contain enough attestations to argue for the rare or frequent nature of the studied evidentials.

It is, however, important to note that usage-based data is far from being a homogeneous concept. There are several types of data available including corpora of written language such as newspaper, literature, and translated literature, corpora of spoken language representing various situations of language use such as conversation and interview, and corpora for different varieties of a given language, among others (see also Section 4.1 for data collected in the field). It is also important to note that there is not just one way of analyzing real usage data, but there are different empirical approaches, such as quantitative corpus analysis based on large amounts of data, or very detailed conversational analysis (CA) of limited sections of discourse. Especially in CA, it is typical to study listeners' reactions, corrections, or gestures, while in other approaches usage patterns of grammatical constructions are more important. Moreover, different approaches, such as sociolinguistics and anthropological linguistics put weight on the socio-cultural aspects of empirical data in different ways.

Usage data also serve the purposes of semanticists, pragmatists, interactional linguists and descriptive grammarians. During the last few decades in particular discourse and corpus analyses have contributed to our understanding of evidentiality and related phenomena. Many corpus-based

studies (on European languages) have shown how evidentiality is expressed in languages without obligatory, grammaticalized evidentiality (see, e.g. Cornillie 2010). Studies of social interaction, in turn, have shown that the source of information is related to such notions as epistemic authority, epistemic status, epistemic or epistemological stance, and knowledge asymmetries (see e.g. Mushin 2001; Heritage and Raymond 2005; Heritage 2012; Bergqvist 2016, 2017).

Discourse and corpus analyses can reveal whether evidentials are typical of certain genres only and/or whether their use varies depending on the context in which they are used, which is important for arriving at a comprehensive picture. This is naturally relevant to our overall understanding of evidentiality, but the variation in the available data also has its drawbacks. The most obvious is that different evidentials may be typical of certain genres only. In Finnish, for example, the use of (optional) evidentiality particles is not felicitous in newspapers, which means that this kind of data cannot be used for any study concerned with evidentiality. Also, folklore in some languages may be characterized by certain evidentials, which are infrequent or used differently in other genres. For instance, evidentials in Brazilian Nheengatú and Ecuadorian Quichua are used for marking a specific poetic genre of story telling (Floyd 2005). Therefore the results of a specific study depend on the type of corpora used, because they usually provide us with information on the use of evidentials in the specific genre, variety, speech situation, and period of time of the corpora.

Next, different types of usage-based data are very suitable for studying the expression of evidentiality in languages where it is not an obligatory category. This opens new insights into how languages take account of information source, when this is not grammatically required. This kind of interactional approach is natural for corpora, because they are mostly available for languages that lack grammatical evidentiality (such as English, French, Finnish and German). Two recent studies of this kind are provided

by Diewald and Smirnova (eds., 2010) and Diewald and Smirnova (2010). The former is a collection of papers dealing with evidentiality in European languages, including studies on verbs of perception in German and English (Whitt 2010) and studies of information source in Basque and Spanish (Alcázar 2010). Diewald & Smirnova (2010), in turn, is a study of evidentiality in German focusing on lexical evidentials such as *werden* ('become') and *scheinen* ('seem') both from a synchronic and diachronic perspective.

Corpora also make contrastive studies of evidentiality possible, as shown above (Alcázar 2010; Whitt 2010; see also Mushin 2001). Another interesting example of a contrastive study is provided by Jaakola (2012) who studied the textual and intersubjective meanings of Estonian and Finnish epistemic-evidential particles *teatavasti* ('as is generally known') and *tieltävästi* ('as far as is known') in newspaper data (see also Jaakola, this volume). The study shows that the Finnish particle refers more often to the reported nature of the information, whereas the Estonian particle usually denotes certainty. Jaakola's (2012) study also underlines the relevance of contrastive linguistics, because two closely related languages nevertheless display differences in the use of etymologically related particles.

Closely related to contrastive studies are studies based on parallel corpora. Typical examples include books that have been translated into many languages. Similar to any large-scale corpora these are mainly available for largely studied languages. Another shortage is that the translator has to make choices when translating. For evidentials, this means that s/he has to take account of the information source and other related functions, when this is made explicit in the source language. The choices made by the translator may also reveal something important about the status of evidentiality in the target language; is it always necessary to take account of the information source, when this is done in the source language? An example of this kind of study is Helin (2004), a contrastive study of German

and Finnish, both of which lack evidentiality as a grammatically obligatory category. The study shows that the German *Konjunktiv* corresponds to the Finnish pluperfect in the expression of second-hand information. In other words, both languages can refer to second-hand information by an evidentiality strategy, but the mechanism used varies.

Corpora of spoken language (along with participant observation) also constitute an important source for unexpected examples of different uses of evidentials, and of evidential functions of other categories and constructions. An individual language user usually cannot think of all possible uses of evidentials, and usage data is therefore absolutely necessary for arriving at a fuller picture of the actual potential of the studied evidentials. Many corpora of spoken language have been collected in the field, and they can be seen as annotated and recorded attestations of participant observation scenarios, which will be discussed in more detail in Section 4. One potential problem related to the uses of evidential expressions is provided by subjectivity. Ultimately, the researcher is responsible for his/her judgements; we may never be sure whether this is actually what the speaker has meant, and in case of corpora of natural language we cannot check this.

Discourse and conversation data represent natural language use, usually including the contexts. Therefore, they constitute the only source of data that renders it possible to study such aspects as epistemic stance and status, epistemic authority, different perspectives etc. On the other hand, information source itself that can be more easily studied, for instance, by elicitation. One important aspect in this respect is illustrated by whether the information sources employed are private or shared, which has obvious consequences for how evidentials are used (see, Howard 2012). For example, Hayano (2013) has studied Japanese evidential particles by using video-recorded face-to-face conversations and audio-recorded telephone conversations. Her results show that the particles *yo*, *yone* and *ne* are used

for managing knowledge distribution between speech act participants. The particle *yo* is used to express the epistemic authority of the speaker, while *ne* indicates shared information between the speaker and the hearer. This kind of information about the use of evidentials cannot be gathered by elicited examples or by consulting grammars only; authentic examples from corpora are necessary for this.

Despite the obvious advantages of usage-based data, they also have a number of problems and limitations. First of all, larger corpora enabling reliable studies of the natural use of evidentials are available for only a handful of languages, many of which lack evidentiality as an obligatory category, as noted above. This also has the consequence that corpora of different languages are not directly comparable with each other (with the possible exception of translational corpora), which renders large comparative cross-linguistic studies less reliable or even impossible. Moreover, most corpora can be used only if the researcher knows the language(s) s/he is studying and is capable of assessing the use of the studied evidentials in question. This dramatically decreases the number of languages a corpus-based study can take into account. Finally, the researcher is always responsible for how the examples from corpora are interpreted, which may lead to incorrect interpretations in some cases. This is especially evident for various pragmatic uses of evidentials in discourse. We cannot be sure whether it was the source of information, other intentions related to the discourse context such as encoding epistemic authority, or these combined, that affected the speaker when she chose to use a certain evidential expression.

One further thing that needs to be accounted for is presented by the reliability of the used corpora. First, as noted, corpora differ from each other dramatically as regards their applicability to the studied problem. Genres, in turn, differ with respect to the nature of evidentials occurring in them and also in the functions in which evidentials are used in different genres. This

may mislead us to think that the kinds of evidential (form or function) attested in our data are in general more common, even though they may be typical of a certain genre only. Second, there are differently annotated corpora, and they may vary according to the information given for contexts and the producers of the utterances.

Especially the usefulness of the Internet as data source has been under debate (see, e.g. Kilgarriff 2007), for example, because it may be hard to justify the authenticity of the employed data and anyone has access to the Internet, including non-native speakers of any given language. However, the most obvious advantages of using Internet searches as an alternative to formal corpora is that gathering data is very easy, and a Google search may give us examples on evidential uses that have very low frequency in corpora. For example, evidential particles in Finnish are not very widely attested in certain genres, including newspaper texts that lack hearsay evidentials altogether. Therefore, it may be difficult to gather enough examples for any relevant study of evidentiality by using existing corpora based on newspaper texts or daily conversations. In some cases like this, it may be justified to use Internet searches for gathering preliminary data for testing hypotheses, and at least for English, it appears that even valid frequency counts have been produced by using internet searches in linguistic research (Mittelberg et al. 2007: 42–43).

However, the language used in social media and different discussion forums in the Internet is often close to spoken language, which renders it possible to find spoken language data – like uses of evidentials in written form. Moreover, data obtained from Internet can be fruitful for studying pragmatic uses of evidentials, including humor, irony, sarcasm and (im)politeness. For example, de Hoop, Foolen, Mulder and van Mulken, and Mulder (this volume) have used Twitter as a data source for their studies. de Hoop et al. were able to pinpoint different uses of two Dutch particles *ik geloof* ‘I believe’ and *ik denk* ‘I think’. They can both be considered

expressions of inferential evidentiality, but they also have differences. The use of *ik geloof* less rational, more impressionistic and less inferential, and it also has a mirative reading in certain contexts.

On the other hand, discussions collected from the Internet represent a certain genre, which is regulated by social norms different from face-to-face conversation, and the data obtained from the Internet does not necessarily give reliable information on how evidentials are used in other kinds of daily conversations. Because discussion in some Internet forums makes it possible to express views and opinions anonymously and without seeing other participants in the conversation, people may be more prone to aggressive verbal behavior than in face-to-face conversation, and a large number of insulting comments obtained from the Internet may distort the data. We may also say that these kinds of data broaden our perspective into evidentiality, because such data may be hard to find in normal face-to-face conversation.

We can then conclude that the Internet may be a valuable source for studying rarely attested and non-evidential uses of evidentials, but the data obtained from the Internet may be biased, and the use of evidentials on the Internet is conditioned by different social norms than in face-to-face conversation. Therefore, to get a complete picture of evidentiality in a single language, data from the Internet must be complemented by corpora of spoken language. If evidentials are not widely attested in conversations, one alternative would be to use carefully designed questionnaires and staged communicative events to complement the data collected from Internet with spoken language data.

The results of usage-based studies may also serve as a basis for other kinds of studies of evidentials. First, the study of lesser known languages can be informed by research on evidential expressions in better documented languages for which corpora are available. They may, for example, make it possible to search for previously unknown categories in lesser-studied

languages. Natural language data render it possible to approach evidentiality from various, potentially novel, perspectives, which opens new insights into evidentials in languages for which corpora are lacking; we may find categories that were not known to exist in the language under study. This results in a more comprehensive overview of evidentials and their use in the lesser studied languages. Second, usage-based data may serve as a basis for detailed studies based on questionnaires or grammaticality judgments. A very nice example is provided by Brosig (this volume). Brosig used about 10 hours of relatively free conversational data from unscripted television programs, which were transcribed and annotated. This was followed by judgments by native speakers, who were asked not only what the utterances mean, but also whether another evidential would be appropriate and if so, what the difference would be. Natural language data provide us with a very good overview of how certain evidentials are actually used, but they do not make it possible to test the grammaticality of the attested examples, which is very helpful for understanding the full evidentiality potential of a given language.

As noted, large-scale corpora are mostly available for languages where evidentiality is not an obligatory grammatical category. Therefore, if our aim is to compare the use of evidentials in languages where expressing evidentiality is obligatory, and in languages where the use of evidentials is optional, it is necessary to combine information derived from grammars, corpora and linguistic fieldwork. Grammars are usually biased towards the description of grammaticalized evidentials, while corpora with natural language data are needed for understanding evidentiality in languages that lack grammaticalized evidentials, as well as for evidential extensions of non-evidential categories (e.g. tense in Uralic languages). These kinds of study are of the utmost importance for arriving at a more comprehensive understanding of evidentiality. However, large-scale comparisons of this kind are still lacking. Moreover, as noted above, interpretations about

natural language data made are at least to some extent based on the researcher's intuition. Without a direct access to the speech situation and the intentions of the speakers, which are practically never available for corpus-based data, we may never be sure whether our interpretations are actually correct. The motivations of the speaker may be different from what we expect them to be. Consequently, a very fine-grained analysis is not necessarily possible if the researcher is not able to control the discourse context (see also 4.2. for staged events and stimuli-based techniques).

4. Field linguistics

To begin with, it is important to note that field linguistics cannot be viewed as a homogeneous notion. Linguistic fieldwork includes several different methods: grammaticality judgment, elicitation by translations, sentence interpretation tasks or by setting a discourse context for a given utterance, stimuli-based techniques, collecting narratives, and participant observation (see also Himmelmann 1998). Also, dialogues and conversations of multiple participants can be studied in the field. The different field methods will be discussed separately in what follows, even though they complement each other, and a meaningful distinction between them may be difficult to make.

First of all, the most evident merit of field linguistics (regardless of the examined topic) is that it aids us in collecting valuable data from lesser studied languages. Descriptive grammars of previously undescribed languages are based on fieldwork, which means problems related with field linguistics are present also in typological studies dealing with evidentials. Moreover, consulting native speakers is necessary when we want to deepen our understanding of the evidentiality system or the semantics and pragmatics of evidentials in a given language. Native speakers' intuitions are also necessary when our goal is to find out what strategies the language under investigation uses in situations where some other languages typically

use evidentials. The different methods in the field enable us to focus on the kind of data we are most interested in.

4.1. Naturally occurring speech

By doing linguistic fieldwork, we often aim to collect data of natural language. However, the field linguists' paradox is that recording a situation always affects the language use of the participants. Consultants may want to prepare themselves and plan what they want to say, or they might, perhaps unconsciously, accommodate their speech to the standard language which may be very different from their own idiolect. Nevertheless, spontaneous speech usually consists of grammatical sentences, even if the speakers may prefer different style and variety in their speech because of the fact that the situation is observed, and the chosen utterances do not necessarily represent their everyday language use. If a certain evidential, for example, is seen as a sign of a low register (even when they are frequent), the speakers may avoid their use, which distorts our picture of the evidentials in the studied language, in this case giving the false picture that these evidentials are rare. Moreover, the audience of the speech situation may affect the language use dramatically. If there are other community members present in addition to the researcher, the speaker may use different perspective-indexing forms when she assumes the audience to share the narrated information with her. (See Berqvist 2016 for encoding the speaker and addressee-perspective in Kogi (Arwako).)

When studying evidentials, not only the immediate linguistic context is important, but also a larger context of the whole communicative event. For instance, not only the speaker's information is important, but also the hearer's perspective. However, it might be difficult to find spontaneously produced examples of certain expressions related to evidentiality, especially when evidentiality is not an obligatorily used category in the language. In

Tundra Nenets (Samoyedic, Uralic, Northern Russia), the so called reputative mood encodes reported information of something that the speaker does not commit herself to, or, retelling something that the speaker considers not to be true (roughly in English ‘as if’). Example (1) was used in a personal narrative that a consultant told about her childhood. She refers to a situation where she felt that she was expected to know something she did not know:

Tundra Nenets (Uralic, Jalava 2012: 135, modified)

(1) *xə-ńaəd=ći ʔeńewə-wna-w?*

what-ABL=CL know-**RPT**-1SG>SG

‘How would I know that?’ (as I clearly didn’t know)

Because the reputative form can be used in restricted contexts only, it is not very common in speech. For the same reason, it is a form that is not easy to elicit new examples of from the language consultants without creating and explaining a specific context.

Observing communicative events is undoubtedly the best method for examining pragmatic extensions of evidential expressions in spoken language, such as irony and politeness. In recent years, several studies have also discussed the role of evidentials in interaction and different kinds of communicative practices (see e.g. Gipper 2011; Nuckolls and Michael 2012). Studies based on recorded natural speech, both conversation and narratives, have also shed new light on communicative functions and categorical status of evidentials in lesser studied languages. Using a multi-genre corpus including natural conversations, Frog-story retellings, narrative texts and elicitation, Gipper (2011) has observed that speakers of Yurakaré (unclassified, Bolivia) use hearsay evidentials to express both reported evidence and epistemic possibility. Hearsay evidentials in Yurakaré also have many uses not directly related to the information source, but to

interaction between the speech act participants, like expressing agreement and non-personal wishes (Gipper 2011: 70). Moreover, based on data consisting of natural conversations between speakers of Nanti (Arawak language family, Peru), Michael (2012) has concluded that in Nanti, lexical and grammaticalized quotatives have different interactional functions. Furthermore, also analyzing recordings of spoken language, Bruil (2014) argues that in Ecuadorian Siona, the reportative is synchronically a distinct clause type in which the epistemic authority lies outside of the communicative event and its participants. In many cases, as is generally known by field linguists, the description is based only on one idiolect, and we cannot be sure of its representativeness. On the other hand, sociolinguistic variation may open new perspectives into the grammaticalization of evidentials. Gipper (this volume) has compared varieties of Yurakaré spoken in three villages, and she shows that in one of the varieties, some of the younger speakers use the evidential *-shi* in a new meaning, which points to a change in meaning in that variety.

In her seminal work on evidentiality, Aikhenvald (2004) concludes that reliable information on evidentials can be obtained only by studying spontaneous conversations, which may mean, as we have shown, many different things in different studies. Any study on the evidentiality system of a little documented language should start with collecting and analyzing texts of various genres and the evidentiality patterns found in texts should then be confirmed and systematically studied by participant observation. The fieldworker should concentrate on extensive work on spontaneous dialogues, including gossips, casual remarks and overheard conversations, and avoid direct elicitation (Aikhenvald 2004: 385–386).

Without denying the importance of the naturally occurring spoken conversation in studying evidential categories, it should be noted that the participant-observation technique also has its limitations. Vokurková (2008: 13–14) has noted in her study on epistemic modalities and evidentials in

spoken Standard Tibetan that some of the verbal endings were so rare in spoken language that it was impossible to gather adequate information on them by studying only dialogues, and using systematic elicitation and techniques based stimuli was necessary to establish all the evidential paradigms. The study of Yurakaré evidentials by Gipper (2011) further highlights the importance of multi-genre corpus and the interplay of different methods in the study of evidentiality. The hearsay evidential in Yurakaré can express both reported evidence and epistemic intersubjectivity, with reported use being more common in narrative texts and epistemic uses dominating in conversational data. Furthermore, elicitation was necessary in collecting complete evidential paradigms in Yurakaré, as well determining which combinations of evidentials not attested in the corpus were ungrammatical (Gipper 2011: 18).

Participant observation and recording spontaneous speech provides us the specific context in which a particular evidential expression is used. However, the specific functions of evidential expressions and their frequency in data might vary prominently, depending on the nature of the communicative event. Evidentials are, for example, often used differently in data resulting from monologues than in data that represents dialogues.

4.2. Controlled and staged events

Participant observation as well as reference grammars and text collection may give the researcher important clues on what kind of evidential expressions exist in the language under study. The ways of using evidentials in different communicative contexts, as well as limits of use of the given expressions may then be systematically tested by using context-specific elicitation tasks and staged communicative events. It is important to note that elicitation by translations generally suits rather bad for evidentials, since source of information is not a part of the event itself that we are

describing, its contribution is very different. However, recent studies have shown that carefully designed stimulus-based tasks and staged communicative events constitute a very important tool for studying evidentiality and they make it possible to study spontaneous discourse in contexts where evidentials are common (see San Roque et al. 2012; Silva and AnderBois 2016).

The importance of context for eliciting evidentials is illustrated by example (2) from Wutun. In Wutun, third-hand information is expressed by using a combination of an evidentiality strategy (reported speech construction) and a grammaticalized reported evidential, both of which are based on the verb *sho*, ‘to say, to speak’. Example (2) was used by the speaker when she was asked to think about different ways of telling that she had just heard that her friends were leaving:

Wutun (Sinitic, Sandman 2016: 345):

(2) *gu-jhege qhi-zhe sho-di-li sho-li*
3-PAUC go-PROSP say-PROGR-SEN.INF REP-SEN.INF
‘I have heard from somebody that they say they will go.’

Example (2) was not directly elicited by asking the speaker to translate sentences, but instead the researcher provided the communicative context and let the speaker think about different expressions that could be used in that particular communicative context. Expressing third-hand information is not very common in texts and conversations, but by using the staged communicative events technique it is possible to elicit examples of some of the less commonly attested uses of evidentials and gather information on how the expression changes if different evidentials are added.

Eliciting a discourse context for clauses with evidential marking or testing the grammaticality and meaning of different evidentials in various contexts might provide us with important information on not only what the

consultant considers possible or common in the language, but also on what is impossible. Of course, playing with the language and creating new discourse contexts require good cooperation between the researcher and the consultants. In field linguistics, stimuli-based techniques such as culturally relevant video-clips, animated scenarios, pictures, drawings, and comics as well as pantomime or movable objects are often used in order to avoid the interference of the meta-language. But do they help us to avoid the fact that consultants may use a form that they think they should use? However, for studying evidentiality, visual stimulus materials might not always be as useful as for instance for research on spatial relations or concrete actions and events.

An example of testing the applicability of both visual vs. non-visual stimuli is Vokurková's (2008) study on epistemic modalities and evidentials in spoken Standard Tibetan. Visual stimuli included photographs of unknown people and landscapes that the speakers were asked to talk about, while non-visual stimuli included different types of covered objects. Speakers were first allowed to observe the shape of objects, and then touch them from outside and inside the cover. It turned out that when the speakers were asked to talk about the photographs or judge objects merely by sight, they used primarily epistemic verbal markers. However, when they were allowed to touch the objects, they used evidential verbal endings. Non-visual stimuli therefore turned out to be more useful in eliciting evidentials than visual stimuli.

Furthermore, applicability of different stimuli-based techniques depends on the evidentiality systems of individual languages as well as the cultural context. This is why all materials are not useful for every language. For instance, Foley (2003) emphasizes the importance of cultural context in language documentation by giving an example of his fieldwork with speakers of Watam (Papua New Guinea). He used a story *Frog, Where are You?*, in which the storyline is presented by pictures, and asked the

consultants to narrate the story based on what they saw in the pictures. Foley noticed that the structure of the language used in the narrative prompted by the Frog-story differed in many ways from the the discourse profile of a traditional Watam narrative.

However, promising examples of stimuli-based techniques applicable to different languages and environments have been published recently. An illustrative example is provided by the *Family Problems Picture Task* (San Roque et al. 2012). This method was designed as a field elicitation tool for recording language material especially rich in social cognition content. It is mainly meant for exposition and discussion that includes reporting of speech and thought, using of “cognitive categories” such as evidentiality and mirativity, and references to emotion. The task consists of collaborative narrative problem-solving and retelling by a pair or small group of language speakers. It allows different cultural groups to integrate their own experiences, concerns and conventions with the pictures/story. It may also allow speakers to use constructions that have not been studied extensively before. The problem-solving nature of the task stimulates participants to express their own observations, inferences, and discoveries and also judgments of doubts or certainty.

Another interesting example of staged communicative events applied to research of inferential evidentials was introduced by Silva and AnderBois (2016). They used the logic board game *Mastermind* as a stimulus for collecting natural data on evidential expressions in Desano (Eastern Tukanoan). Asking the consultants to play the game, the researchers created a situation in which the participants had to work collaboratively and discuss the possible solutions based on logical reasoning. This way the researchers were able to record naturally occurring examples of evidential markers and miratives which rarely occur in Desano narratives. In contrast to the Family Problems Picture Task (San Roque et al. 2012), Silva and AnderBois (2016)

focus on a context where inference, reasoning and mirativity are strongly present.

In conclusion, knowing the context of the communicative event is crucial for studying evidential expressions in any language. Using context-dependent elicitation tasks, stimuli-based techniques and experiments may be crucial in order to ‘force’ the speakers to use evidentials that are too rare to be recorded in naturally occurring speech, but it requires mutual understanding of the researcher and the consultant about the communicative context. Naturally occurring speech is the most reliable material for studies of evidential expressions in any language, but it does not always suffice for acquiring the data we need for our study. Moreover, as noted above, we cannot be sure of the actual intentions of the speakers when studying naturally occurring speech. Consequently, staging and guiding are necessary, because it gives the researcher more certainty of the speakers’ intentions, even though it decreases the authenticity of the data (see also Himmelmann 1998; Lüpke 2006). Moreover, well designed stimulus-based tasks can be repeated for many languages, which makes contrastive studies possible.

5. Questionnaires and acceptability judgments

Questionnaires have many different functions in linguistics. They have been used, e.g., for grammaticality judgements, the use of certain constructions, or for choosing the most appropriate construction in a specific context. They are also useful for field linguistics, and specific designed questionnaires may aid the fieldworker in covering as many areas of grammar as possible. Eurotyp TMA questionnaires illustrate a very good example of a questionnaire used for surveying contexts in which tense and aspect markers are used (Dahl 1985: 198–206). We must, however, bear in mind that such

questionnaires still lack for evidentials, and the techniques used are rather stimuli-based (see 4.2.). Here, however, we understand questionnaires as an individual data collection method. Such questionnaires are used for gathering bigger samples than what usually is possible in fieldwork understood as language documentation. It is typical that they are filled in by the informants themselves.

Questionnaires constitute an important data collection method that can be used for complementing results gathered by other methods (see Brosig, this volume); they can also be used as a basis for individual studies (Mari, this volume) or for gathering data on well but also lesser studied languages. They are easily distributed, which means that the number of participants for a study can be very high. Moreover, the researcher is free to choose the target group for his/her study, which means that questionnaires make sociolinguistic studies possible and we may study the differences between, e.g., sexes, generations and social groups (see e.g. Cheung, Leung, Yang, Xing and Tse 2010 for a study on variation of the use of multiple hearsay and inferential evidentials in Japanese depending on speakers' age). Also cross-linguistic studies are possible using questionnaires, because the same questionnaire may be filled in by speakers of many languages. Closely related to questionnaires are translation corpora consisting of several individual translations of the same source text to the same target language. This kind of data may reveal which strategies individual translators use in the target language to interpret evidentiality expressed in the source language, or, what implications in the source language direct them to use evidentials in the target language (see Helin 2006).

The most obvious advantage of questionnaires is that they enable us to systematically gather data on exactly those features we are interested in. Like elicitation and staged communicative events techniques, questionnaires also render it possible to study the limits of a particular construction, which is not possible by using corpora or grammars alone. The study of

ungrammatical or marginal uses of a construction (e.g. studying situations where certain evidentials are not possible, or are marginal at best) broadens our horizon and aids us to better understand the studied phenomenon. In this sense, questionnaires resemble corpora, but they enable us to search for the examples we are most interested in. On the other hand, such systematicity is lacking for corpora, and we can never count on finding the data needed for our study.

Situations where multiple kinds of evidence are available are hard to find in corpora, but they can be easily created for questionnaires. The following are genuine examples from a questionnaire for the study of hierarchy between different sources of information in Finnish (both the original stimulus and an English translation are given below). The questionnaire was filled in by 134 native speakers of Finnish attending Seppo Kittilä's course 'Introduction to Linguistics' in the fall of 2011 (*näköjään* is the sensory evidence/inference evidential of Finnish, while *kuulemma* codes hearsay evidence):

Tunnet olevasi tulossa kipeäksi. Ystäväsi mittaa kuumesi ja sanoo, että sinulla on kuumetta 37,6, mutta itse et jaksa katsoa mittaria. Soitat kaverillesi ja sanot, ettet varmaankaan pääse huomenna yliopistolle:

Moi, mä en pääse tulemaan, koska olen _____ tulossa sairaaksi

You think that you are getting sick. Your friend checks your temperature and says that it is 37,6, but you are too tired to look at the thermometer. You call your friend saying that you probably won't make it to the university tomorrow:

Hi, I won't be able to come, because I am _____ becoming sick.

Kuulemma 10

Blank 3

Olet Naga Morichin syönnin maailmanmestari ja ystäväsi kehuskelee laittaneensa ruokaan paljon chiliä, mutta et maista mitään. Kerrot tästä toisellesi kaverillesi:

Siinä ruoassa oli _____ aika paljon chiliä

You are the World champion in Naga Morichi eating and your friend boasts with the amount of chili s/he has put in the food, but you don't taste anything. You tell about this to another friend:

That food _____ had a lot of chili in it.

Näköjään 2

Kuulemma 132

In the first case, the speaker has both sensory evidence and hearsay evidence, of which the sensory evidence can be deemed more reliable. This has the consequence that the sensory evidence/inference particle *näköjään* is clearly favoured (121 times out of 134). In the second case, sensory evidence contradicts the hearsay evidence, and consequently the hearsay evidential becomes much more frequent (132 out of 134). These findings point to the fact that hearsay evidentials are used whenever the speaker has to rely more on other people's evidence. On the other hand, if the speaker has more direct evidence at his/her disposal, s/he rather resorts to direct evidentials. Corpora only rarely provide us with such reliable examples. Participant observation of course may sometimes produce some sporadic examples, but systematicity is lacking and broader generalizations are difficult, if not impossible, to draw. Questionnaires may also complement

preliminary corpus-based findings by enabling us to focus on problems that the corpora have hinted at without providing us with definite results.

Even though questionnaires have their evident advantages, they, as any other data collection method, have certain problems and limitations. The most obvious problem of questionnaires is to formulate the questions in the correct and most suitable way. In the case of evidentials, the scenarios in such a way that they are understandable and as unambiguous as possible to anyone partaking in the experiment. This problem is especially acute for larger, and thus more heterogeneous, groups of participants. We can never be sure whether participants have understood the questions correctly, which may affect the reliability of questionnaires, or the results. Moreover, in many cases the identity of the participants is anonymized, which makes it impossible to ask more specific questions later, and even if this were possible it would be too time-consuming. From this it also follows that we may become aware of the less than clear nature of the questionnaire only afterwards, which can be fully corrected only by running the test again. One option is to ignore unreliable/unclear answers, but it may not always be straightforward which answers are unclear enough to be ignored. As with corpora, we always use our judgements when interpreting the results, which may result in ignoring relevant data, because we cannot interpret it correctly.

It is worth noting that questionnaires concerned with acceptability judgements or fill-in tasks have features in common with experimental techniques that makes speakers interpret specific linguistic constructions. Tamm et al. (this volume) have studied the use of the Estonian hearsay evidential *-vat* with monolingual speakers of Estonian aged 4 and 6 years in different parts of Estonia in 2013. Results were analysed in two age groups for determining whether there are developmental differences in understanding the meaning of the Estonian *-vat* morpheme.

Finally, even though questionnaires render it possible to create suitable scenarios for studying whatever uses of evidentials (and also for testing

whether certain evidentials exist in a language), it must be borne in mind that the data we acquire may not give a real picture of how evidentials (or any other studied construction) are actually used, but that the data reflects how they are used in this experimental situation (see also Vanhatalo 2005: 33; Itkonen and Pajunen 2010: 93). The created scenarios are necessarily somewhat artificial and we cannot be sure whether the participants use the studied constructions in the same way when they have to make a decision in a normal speech situation. For example, in our questionnaire (see above) the participants were partly forced to choose between the two evidential particles of Finnish, and we may not be sure whether the participants would have actually used either particle in a normal speech situation. So, this kind of study may not necessarily tell us about how evidentials are actually used, but how they may be used. However, this is a type of uncertainty we have to deal with.

6. Conclusions

The primary goal of this chapter was to show that evidentiality is a broad notion, and that using an array of methods is necessary for arriving at a comprehensive understanding of it. All methods discussed have their advantages and disadvantages, but none should be excluded due to their potential problems. It also goes without saying that each of the examined methods suits best for a specific type of study. Grammars, for example, constitute the only way of acquiring a cross-linguistically relevant overview of evidentials, while corpora are needed for presenting any statistical data on evidentiality. However, the different methods rather complement each other. An example is the use of questionnaires for gathering more specific data for questions that corpora leave open. The relevance of this is demonstrated by Brosig (this volume). The data provided by corpora are random in the sense that we cannot control the data, and we may not be able

to find the very examples that are central to our study. Questionnaires, for their part, make it possible to collect exactly the kind of data needed for any study. The same advantage applies to carefully designed stimulus-based tasks (such as San Roque et al. 2012) as they make it possible to study naturalistic discourse in situations in which the researcher knows the context well.

Our paper includes references to studies that have applied different types of data to the study of evidentials. Moreover, it also includes some hypotheses and suggestions for future studies on evidentials. We have only touched upon the possibilities that questionnaires and staged communicative events provide for the subject at hand. It is therefore our hope that scholars working on evidentiality and related topics will put the discussed methods to the test in following studies on evidentiality.

The present chapter has shown that the study of evidentiality differs from other phenomena in linguistics because of their non-at-issue semantics. In other words, evidentials do not concern the contents itself, but they express how the speaker knows what s/he knows, and how other peoples' information affects their use. Consequently, traditional questionnaires, for example, do not necessarily capture the essence of evidentials, and a field linguist working on a given language may not be able to look for the right constructions, since the contexts studied may be wrong. On the other hand, studies of, e.g., argument marking or negation are easier, and it is possible to use elicited sentences for acquiring the needed information. As a result, different kinds of technique are necessary for evidentials. As the discussion above has shown, many techniques have been developed, and we know much more about evidentiality than we did a decade ago. Due to these evident differences, studies of evidentials may thus have a lot to offer to field linguistics methodology as well.

Abbreviations

1	First person
3	Third person
1SG>SG	First person singular subject > singular object (objective conjugation)
ABL	Ablative
CL	Clitic
PAUC	Paucal
PL	Plural
PROGR	Progressive aspect
PROSP	Prospective aspect
REP	Reportative
RPT	Reputative
SEN.INF	Sensory-inferential evidential
SG	Singular

References

- Aikhenvald, A. Y. 2004. *Evidentiality*. New York: Oxford University Press.
- Alcázar, A. 2010. Information source in Spanish and Basque: A parallel corpus study 1. In G. Diewald & E. Smirnova, (Eds.), *Linguistic realization of evidentiality in European languages* (131–156). Berlin: Mouton de Gruyter.
- Bergqvist, H. 2016. Complex epistemic perspective in Kogi (Arwako). *International Journal of American Linguistics*, 82(1), 1-34.
- Bergqvist, H. 2017. The role of “perspective” in epistemic marking. *Lingua*, 186, 5–20.

- Boas, F. (Ed.). 1911, *Handbook of American Indian Languages*. Part 1. Smithsonian Institution. Bureau of American Ethnology Bulletin 40.
- Bruil, M. 2014. *Clause-typing and evidentiality in Ecuadorian Siona*. PhD dissertation, Universiteit van Leiden.
- Chafe, W. L., and J. Nichols (Eds.). 1986. *Evidentiality: The linguistic coding of epistemology*. Advances in Discourse Processes 20. Norwood, NJ: Ablex.
- Cheung, J. C.K., Leung, M., Yang, A. Xing, D. & Tse, J. 2010. Variation in restrictions on multiple evidential markers in Japanese by speaker age. In T. Peterson, & U. Sauerland (Eds.), *Evidence from evidentials* (29-39).[University of British Columbia Working Papers in Linguistics. Vancouver, Canada.
- Cornillie, B. 2010. An interactional approach to evidential and epistemic adverbs in Spanish conversation. In G. Diewald & E. Smirnova (Eds.), *The linguistic realization of evidentiality in European Languages* (309–330). Berlin/New York: Mouton de Gruyter.
- Curnow, T. J. 2001. Evidentiality and me: The interaction of evidentials and first person. In C. Allen (Ed.), *Proceedings of the 2001 Conference of the Australian Linguistic Society*.
< <http://www.als.asn.au> >
- Curnow, T. J. 2003. Nonvolitionality expressed through evidentials. *Studies in Language*, 27(1), 39–59.
- Dahl, Ö. 1985. *Tense and aspect systems*. Oxford: Blackwell.
- Diewald, G., & Smirnova, E. 2010. *Evidentiality in German*. Berlin: Mouton de Gruyter.
- Diewald, G., & Smirnova, E. (Eds.). 2010. *Linguistic realization of evidentiality in European languages*. Berlin: Mouton de Gruyter.
- Floyd, S. 2005. The poetics of evidentiality in South American storytelling. *Santa Barbara papers in linguistics*, 16, 28–41.

- Foley, W. A. 2003. Genre, register and language documentation in literate and pre-literate communities. In Peter Austin (Ed.), *Language Documentation and Description*, 1, 85–98. London: Hans Rausing Endangered Language Project, School of Oriental and African Studies.
- Fox, B. 2001. Evidentiality: authority, responsibility and entitlement in English conversation. *Journal of Linguistic Anthropology*, 11(2), 167–192.
- Gipper, S. 2011. *Evidentiality and intersubjectivity in Yurakaré. An interactional account*. PhD dissertation, Radboud Universiteit Nijmegen.
- Hayano, K. 2013. *Territories of Knowledge in Japanese Conversation*. PhD dissertation, Radboud Universiteit Nijmegen.
- Helin, I. 2004. ... so der Wetterbericht. *Evidentialität und Redewiedergabe in deutschen und finnischen Medientexten und Übersetzungen*. Frankfurt am Main: Peter Lang.
- Helin, I. 2006. Implication of Evidentiality in Translation. *SKY journal of linguistics*, 19, 282–290. [In M. Suominen (Ed.), A man of measure: Festschrift in honour of Fred Karlsson on his 60th birthday, special supplement].
- Heritage, J. 2012. Epistemics in action: Action formation and territories of knowledge. *Research on Language & Social Interaction*, 45(1), 1–29.
- Heritage, J., & Raymond, G. 2005. The terms of agreement: Indexing epistemic authority and subordination in assessment sequences. *Social Psychology Quarterly*, 68(1), 15–38.
- Himmelman, N. P. 1998. Documentary and descriptive linguistics. *Linguistics*, 36, 161–195.
- Howard, R. 2012. Shifting voices, shifting worlds: Evidentiality, epistemic modality and speaker perspective in Quechua oral narrative. *Pragmatics and Society*, 3(2), 243–269.

- Itkonen, E., & Pajunen, A. 2010. *Empiirisen kielitieteen metodologia*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Jaakola, M. 2012. Displaying knowledge in journalistic texts: A contrastive analysis of an evidential particle in Estonian and Finnish. *Lähivõrdlusi* [Lähivertailuja 22] (43–70). Tallinn: Eesti Rakenduslingvistika Ühing.
- Jalava, L. 2012. Tempuksen ilmaiseminen tundranenetsin moduksissa. In L. Jalava, E. Sandman, J. Saarikivi, T. Hyytiäinen (Eds.), *Per Urales as Orientem: Iter polypronicum multilingue*, 131-144. Memoires de la Société Finno-Ougrienne 264, Helsinki: Suomalais-Ugrilainen Seura.
- Keinänen, S. 2017. Comparison of evidentiality terminology. A paper presented at the symposium ‘The expression of knowledge: epistemicity and beyond’, Helsinki, 23.8.2017.
- Kilgarriff, A. 2007. Googleology is Bad Science. *Computational Linguistics*, 33(1), 147–151.
- Lüpke, F. 2006. Small is beautiful: The contribution of small field-based corpora to different linguistic disciplines. In P. Austin (Ed.), *Language Documentation and Description*, 3, 75–105. London: Hans Rausing Endangered Language Project, School of Oriental and African Studies.
- Michael, L. 2008. *Nanti Evidential Practice. Language, Knowledge and Social Interaction in Amazonian Society*. PhD dissertation, University of Texas, Austin.
- Michael, L. 2012. Nanti self-quotation. Implications for the pragmatics of reported speech and evidentiality. *Pragmatics and Society*, 3(2), 321–357.
- Mittelberg, I., Farmer, T., & Waugh, L.R. 2007. They actually said that? An Introduction to working with usage data through discourse and corpus analysis. In M. Gonzalez-Marquez, M. Spivey, I. Mittelberg, & S Coulson (Eds.), *Methods in Cognitive Linguistics*, 19–52. Amsterdam/New York: John Benjamins [Human Cognitive Processing Series 18].

- Mushin, I. 2001. *Evidentiality and epistemological stance: narrative retelling*. Amsterdam: John Benjamins.
- Nuckolls, J., & Michael, L. 2012. Evidentials and evidential strategies in interactional and socio-cultural context. *Pragmatics and Society*, 3(2), 181–188.
- Plungian, V. 2001. The place of evidentiality within the universal grammatical space. *Journal of Pragmatics*, 33, 349–357.
- Sandman, E. 2016. *A Grammar of Wutun*. PhD dissertation, University of Helsinki.
- San Roque, L., Gawne, L., Hoenigman, D., Miller, J.C., Rumsey, A., Spronck, S., Carroll, A., & Evans, N. 2012. Getting the story straight: Language fieldwork using a narrative problem-solving task. *Language Documentation & Conservation*, 6, 135–174.
<http://hdl.handle.net/10125/4504>.
- Silva, W., & AnderBois, S. 2016. Fieldwork Game Play: Masterminding Evidentiality in Desano. *Language Documentation & Conservation*, 58–76.
- Vanhatalo, U. 2005. *Kyselytestit synonymian selvittämisessä. Sanastotietoutta kielenpuhujilta sähköiseen sanakirjaan*. PhD dissertation, University of Helsinki.
- Vokurková, Z. 2008. *Epistemic modalities in Spoken Standard Tibetan*. PhD dissertation, Université Paris 8.
- Whitt, R.J. 2010. *Evidentiality and perception verbs in English and German* [German Linguistic and Cultural Studies 26]. Bern: Peter Lang.
- Willet, T. 1988. A cross-linguistic survey of the grammaticalization of evidentiality. *Studies in Language*, 12, 57–91.