MULTIMODAL COMMUNICATION IN THE 21ST CENTURY: PROFESSIONAL AND ACADEMIC CHALLENGES. 33rd Conference of the Spanish Association of Applied Linguistics (AESLA), XXXIII AESLA CONFERENCE, 16-18 April 2015, Madrid, Spain

Evidentiality as conversational implicature: Implications for corpus annotation

Marta Carretero\textsuperscript{a}, Juan Rafael Zamorano-Mansilla\textsuperscript{a}\textsuperscript{*}

\textsuperscript{a}Universidad Complutense de Madrid, Departamento de Filología Inglesa I, Ciudad Universitaria, Madrid 28040, Spain

Abstract

This paper discusses a number of issues involved in the annotation of evidentiality communicated as a conversational implicature in authentic written texts. As a pilot experiment, evidentials were annotated separately by two experts in sample texts from the MULTINOT corpus, which consists of English and Spanish comparable and parallel texts from different registers. The results of this annotation proved that 1) evidentiality in English is most often expressed by pragmatic means, and 2) these means easily provoked interannotator disagreement. Some types of these pragmatic evidentials are specified, together with the implications for the design of an annotation system for evidentiality.

© 2015 The Authors. Published by Elsevier Ltd.

Keywords: evidentiality; annotation; conversational implicature

1. The concept and scope of Evidentiality

Evidentiality could be broadly described as the source of information behind our message. This reference to our ‘source of information’ is present in most definitions of evidentiality in the literature (Boye 2010: 1). Aikhenvald (2007) establishes a distinction between ‘evidential’, which refers to an obligatory grammatical category found in some languages, and ‘information source’, which includes the “corresponding conceptual category” (2007: 209). In this paper we use the term ‘evidentiality’ to refer to the general conceptual category, which we define as the

\* Corresponding author. Tel.: +34-91-3945382.  
\textit{E-mail address: mcarrete@ucm.es; jrzamora@ucm.es}
linguistic expression of the kind, source and / or evaluation of the evidence that the speaker/writer (sp/wr) has or claims to have at his / her disposal, for or against the truth of the proposition.

In line with Boye & Harder (2009), we conceive evidentiality as a functional-conceptual substance domain “without interference from structural criteria associated with different forms of coding” (Boye & Harder 2009: 10). That is to say, the conceptual scope of evidentiality covers all means of expression, which may be grammatical, semantic or pragmatic. In fact, it is our claim that, since English lacks a grammaticized expression of evidentiality, the evidential interpretation often emerges from the pragmatic interpretation of a wide range of expressions. Incidentally, this also means that the annotation of evidentiality in English texts is particularly subject to individual variation, as will be shown below.

Let us consider now some examples of how an expression conveys evidentiality in English because of a pragmatic interpretation. Unless something else is explicitly said or obvious from the context, if a speaker utters (1),

(1) I’ve heard 22 Jump Street is a good film.

s/he is transmitting the following information: a) s/he hasn’t seen it, so the information s/he has is not first-hand; b) the information s/he is providing has been obtained from others.

The close connections between epistemic modality and evidentiality have been often discussed due to their similarities (De Haan 1999). Although we believe that these are distinct conceptual categories, it is interesting to notice how epistemic uncertainty is a necessary component for the evidential interpretation to obtain. Consider (2), a variation of (1):

(2) I’ve heard 22 Jump Street is a good film, but I don’t think so.

Here the expression in bold cancels the epistemic uncertainty of the message: the hearer assumes that the speaker has indeed seen the film, so s/he has first-hand information about it. Once epistemic uncertainty is missing, the expression I’ve heard no longer leads to an evidential interpretation. That is to say, the evidentiality of I’ve heard has a typical property of conversational implicatures: it can be cancelled, and the hearer is then forced to produce a more viable interpretation for the message (something like ‘Some people believe it’s a good film, but I disagree’).

Alternatively, an evidential interpretation can arise from the schemata associated to an entity. For instance, words such as rumour, study or book activate the notions of communication from third parties, opening the possibility of an evidential interpretation in appropriate contexts.

The ways in which language users can talk about the kind, source and / or evaluation of the information behind their message are rather varied. Linguists have defined different labels to classify evidential meanings, some of them hierarchically (Willett 1988), others explaining the oppositions found in cross-linguistic studies (Aikhenvald & Dixon 2001). Although the names employed differ slightly, the notions are quite similar. These are the main types of evidentiality used in our study:

- Reportative. The speaker indicates that their information is based on what someone has communicated.
- Perceptual. The speaker indicates that their information is based on some sensory perception.
- Inferential. One piece of information is presented as being guessed or logically inferred from another piece of information for which there is epistemic certainty.

2. Annotating Evidentiality: general results

This paper describes the results of a pilot experiment in which two experts separately annotated a corpus for the category of evidentiality. The purpose of the experiment was to test the reliability of the definition of evidentiality and its replicability when used to tag English texts.

In this initial stage of the experiment, both experts tagged all the expressions they identified as conveying an evidential meaning in four English texts randomly selected from the MULTINOT corpus. This corpus, compiled as part of the MULTINOT research project, consists of English and Spanish comparable and parallel texts from different registers. Each of the texts used in this initial experiment contains about 1,000 words, and their registers were as follows: two popular science texts, a tourism leaflet and a political essay on economics.
Comparing the agreement rate between both annotators, it is possible to spot weaknesses in the current linguistic characterization of evidentiality. This information can be used to refine the linguistic definition of evidentiality and increase the replicability of the results obtained by two independent annotators.

The general results of this initial experiment are as follows: 4,000 words of English, provided 78 cases of evidential expressions. From these, 44 were identified by both annotators; 26 were identified only by annotator A, and 8 were identified only by annotator B. Table 1 summarizes these results.

One of the problems posed by the method of annotation employed in our experiment is that it is not obvious how these results are to be interpreted. Statistical tests devised to test interannotator agreement, such as Cohen’s Kappa, work on the principle that annotators must classify a set of objects using a previous defined classification. But in our experiment annotators were asked to identify elements that may or may not be present in a text. For this reason, at this point we can only offer a percentage of the number of cases that caused complete interannotator agreement, which was set at 56.41%. A qualitative analysis of the cases that provoked agreement and disagreement is given in the following sections.

Table 1. Interannotator agreement in the annotation of evidentiality.

| Identified by both annotators | 44 | 56.41% |
| Identified by annotator A      | 26 | 33.33% |
| Identified by annotator B      | 8  | 10.26% |
| Total number of evidential expressions | 78 |

3. Expressions unanimously marked as evidentials

The expressions that both annotators marked as evidentials were mostly reportatives. Some examples are *it is often said that*, *it is agreed by most demographers that*, *Wendy Baldwin from the Bureau says that*, *she explains*, *some estimates from the Middle Ages were…*; *Keynesian economists predicted that*, *Every (advanced) country has realized that…*

Other expressions were quotation marks and the verb *seem*, both of which, as will be specified in Section 5 below, were a source of disagreement in other cases. It is also worth mentioning verbs of thinking with third person subjects, such as *recognized* and *understood* in (3):

(3) Seventy years ago, at the end of World War II, the Allies *recognized* that Germany must be given a fresh start. They *understood* that Hitler’s rise had much to do with the *unemployment* (not the *inflation*) that resulted from imposing *more* debt on Germany at the end of World War I. (EO_ESSAY_001) (original italics)

4. Evidentials that provoked disagreement

A small number of evidentials were marked as such only by one of the annotators, due to an error of the other. This occurred in (4) with *obviously*, one of the few English expressions whose evidentiality can be considered as context-independent and therefore semantic (Carretero & Zamorano-Mansilla 2013: 345-348):

(4) *Obviously* the cuisine of Castilla y Leon is by no means confined to these succulent dishes. There are the juicy veals of Avila and Aliste (Zamora), exquisite Zamora cheeses, the marvellous hams of Guijuelo (Salamanca) […] (ETRANS_TOU_006)

However, most cases of disagreement correspond to the expression of evidentiality by pragmatic expressions, which in certain contexts fit squarely within the definition of evidentiality specified above and in other contexts are non-evidential. In this section we describe a number of different types of these expressions, and the solution proposed in each case.
4.1. Expressions that most often express evidentiality (as a GCI), but in some contexts do not

Even apparently obvious evidential expressions, such as the verbs *seem* and *appear*, are not evidential in certain contexts, since the proposition is not true and the verbs only concern appearances, as in (5):

(5) "Critical flicker fusion frequency" - the point at which the flashes *seem* to merge together, so that a light source *appears* constant - provides an indication of time perception. (EO_POPSCI_006)

This paragraph describes the way some animals perceive light. In truth, the flashes do not merge together and the light source is not constant. However, one of the annotators mistakenly annotated *seem*, (but not *appears*).

4.2. Expressions that may have an evidential conversational implicature but it is not always relevant

The word *tangible* means ‘able to be touched’, in a real or metaphorical sense. A state of affairs may be characterized as tangible in order to qualify the proposition as true, thus having similar effects to ‘clear’. This is the case of (6), from the British National Corpus, which concerns tactics for keeping forests in tropical countries:

(6) Show that forests have a *tangible* value kept as forests, for example for timber, and countries will manage them for a sustained cash yield. This top-down approach is favoured by the two major institutions charged with saving the tropical forests, the International Tropical Timber Organisation, and the World Bank and UN backed Tropical Forestry

However, this adjective is less clearly evidential in (7), from one of the texts analyzed, because the truth of the proposition is not challengeable, there thus being no need to insist that it is true. One annotator marked it as evidential though.

(7) The bustle of the university is *tangible* in Salamanca’s local Carnival where boisterous festivity mingles with events of a more serious tone (ETRANS_TOU_006)

This is a complicated issue for annotation, since there will be cases in which the truth of the proposition does not clearly have a challengeable or non-challengeable status.

4.3. Clauses that are not evidential per se, but have evidential effects in discourse

Interannotator disagreement also occurred in a few clauses that do not have evidential value by themselves, but commonly acquire it in discourse. For example, in (8) ‘Records were kept’ is used to state that these records provide reliable information about population growth in the modern period.

(8) Population growth has mostly happened in the modern period, she says, when *records were kept*, so if estimates for the early period are slightly out, this will not drastically change the overall ratio of “ever lived” to “living” (EO_POPSCI_006)

Similarly, in (9) ‘We hardly needed another test’ is used to evaluate the evidence provided earlier and later about the truth of the proposition, mentioned earlier, “that the austerity that was being imposed to Greece and the other crisis countries would fail”.

(9) *We hardly needed another test*. Austerity had failed repeatedly, from its early use under US President Herbert Hoover, which turned the stock-market crash into the Great Depression, to […] (EO_ESSAY_001)

5. Conclusions

The annotation discussed in this paper of evidential expressions in the MULTINOT texts leads to conclude that evidentiality in English is mostly expressed by pragmatic means, which pose difficulties for annotators and, not surprisingly, often provoke interannotator disagreement. A collateral finding is that this research hints that the frequency of the evidential conversational implicature varies with the expressions. For example, verbs such as *appear* or *seem* are evidential unless otherwise indicated or obvious from the linguistic or situational context. Other expressions communicate evidentiality in contexts where they evaluate evidence for (or against) non-obvious truth of propositions, as is the case of the adjective *tangible*. Another kind of expressions, such as the clauses ‘records
were kept’ or ‘we hardly needed another test’, are not evidential per se but have evidential effects in discourse. Further research on evidentiality as a functional-conceptual domain in English (and other languages with non-grammaticized evidentiality such as Spanish) will shed light on these issues. Such research will also uncover the evidential or non-evidential value of different constructions with expressions often considered as evidential in the literature, and will ultimately contribute to the creation of a reliable system for the annotation of evidentiality.

Acknowledgements
This research has been carried out as part of the MULTINOT Project, financed by the Spanish Ministry of Economy and Competitiveness (MINECO) under the I+D Research Projects Programme (reference number FFI2012-32201). As members of the team, we gratefully acknowledge the support provided by the Spanish Ministry and also the BSCH-UCM grant awarded to the research group to which we belong.

References