

Testing the constraints on evidentiality in English: A forced-choice acceptability judgement task

We present results from a forced-choice acceptability judgement task investigating whether English evidential constructions are constrained by parametrized evidential sub-dimensions, as argued by Matthewson (to appear). By manipulating the values of these sub-dimensions in the discourse context, we were able to elicit acceptability judgements which indicate that three of the sub-dimensions, namely, Evidence Type, Strength and Source - but not Location - all significantly predict acceptability of English evidentials. These results suggest that evidentiality is in fact grammatically encoded in particular syntactic constructions in English, and that its marking is obligatory (cf. Fox, 2001; von Stechow & Gillies, 2010; Matthewson, to appear).

Background In an evidential construction, the source of information that a speaker has for their assertion is grammatically encoded. While every language has the means to express evidential concepts (i.e., through lexical items, pragmatic implicature, etc.), an evidential is defined by obligatory grammatical marking (Aikhenvald, 2004). Most languages that mark evidentiality do so morphologically, and as English does not have any evidential morphemic markers, it has been argued that English does not incorporate evidentiality grammatically (Aikhenvald, 2004). However, Asudeh and Toivonen (2012) argue that evidentiality is in fact grammatically encoded in English. Specifically, they claim that English copy raising constructions (e.g., *John looks like he is cooking*) encode direct evidentiality, while unraised constructions (e.g., *It looks like John is cooking*) are unmarked for evidentiality, i.e., the speaker must directly perceive the subject of the event in order to use copy raising constructions (cf. Rett & Hyams, 2014). The former sentence above is only felicitous in a situation where John is observed in the kitchen, whereas the latter can be used when one simply sees pots boiling on John's stove. In an online felicity judgement task, Rett and Hyams (2014) confirmed this basic evidential pattern in English, finding that copy raising constructions were rated significantly lower than unraised constructions when the speaker could not directly perceive the sentential subject (*indirect* evidence). However, the criteria for distinguishing between direct and indirect evidence has remained unclear in the literature. According to Asudeh and Toivonen (2012), direct evidence must involve perception of the subject, while for other authors, it is the type of evidence that must be direct, e.g., sensory versus non-firsthand evidence (Aikhenvald, 2004). In a recent proposal, Matthewson (to appear) argues that evidentiality can be further refined into three different sub-dimensions: Evidence Type, Strength and Location, and that each of these sub-dimensions can be assigned a direct or indirect value (see chart below). Matthewson argues that languages differ in which dimensions they incorporate into their grammars. We have also added Asudeh and Toivonen's (2012) Evidence Source as a fourth possible sub-dimension.

Dimension	Direct	Indirect
Evidence Type	Sensory information.	Reports or reasoning.
Evidence Strength	<i>Best</i> evidence possible for event.	<i>Not best</i> : Lacking best possible evidence.
Evidence Location	Perceive event itself.	Perceive results of event.
Evidence Source	Perceive sentential subject.	Do not perceive sentential subject.

Goal Our study aimed to test and further refine the results from Rett and Hyams (2014) and examine which evidential dimensions are encoded in English evidential constructions. Specifically, we sought to answer: i) Which evidential dimensions have an effect on the acceptability of English evidentials?; ii) What are the values of the dimensions that predict whether a copy raising or expletive construction is preferred?

Experiment In order to facilitate the task of rating sentences that differ in a subtle way, participants were instructed that they were helping a non-native speaker, Idan, learn about the subtleties of English (cf. Rett & Hyams, 2014). In response to a context, participants were asked to rate the acceptability of evidential sentences as well as a declarative control on a scale of 1 (completely unacceptable, no native speaker of English would ever say this in response to that context) to 6 (completely acceptable, this is a perfectly normal response to that context). A sample context is found in (1). Each participant saw either the three expletive sentences in (1a) or the three copy raising constructions in (1b), plus the declarative control in (1c).

- (1) Context: Idan walks into the house and sees that his wife is in the kitchen. She is wearing an apron that is full of flour and has chocolate on her face. He thinks to himself:
- Unraised*: It seems/sounds/looks like she has been baking.
 - Copy raised*: She seems/sounds/looks like she has been baking.
 - Declarative*: She has been baking.

Contexts were manipulated to test for all possible combinations of evidential dimension and direct/ indirect values. For example, the context in (1) has direct evidence type (visual evidence), indirect evidence strength (does not have the best possible evidence for the assertion), indirect evidence location (does not perceive the actual event) and direct evidence source (perceives the sentential subject, his wife). Participants saw an equal number of copy raised and unraised trials but never within the same context manipulation.

Results By-participant *z*-scores of judgement ratings served as a dependent variable in a linear mixed-effects regression model, which included the interaction of Syntax (copy raised and unraised), Evidence Type (direct, indirect or no evidence), and Response Type (*looks*, *sounds*, *seems*). Note that the declarative control response type was not included in our model because it does not have a predictor for raised and unraised structures. A simple main effect indicated that unraised structures were significantly more likely to be rated as more acceptable than copy raised structures. Main effects also showed that participants were significantly more likely to rate sentences higher when Evidence Type was direct than when it was indirect or there was no evidence. Participants were also significantly more likely to provide higher ratings when Evidence Strength was *not best*, compared to when the context contained *best* evidence. Finally, when Evidence Source was direct, participants were significantly more likely to rate sentences higher compared to when it was indirect. Results show that there was a significant two-way interaction between Response Type and Evidence Type. The figure below graphs the mean raw *z*-scores of this interaction. Specifically, the interaction shows that higher ratings were more likely with the verb *seems* (green line) when the evidence type was indirect or there was no evidence (marginal) compared to the verb *looks* (red line) with direct evidence. If the verb was *sounds* (blue line), higher ratings were more likely with indirect evidence but lower ratings were more likely when there was no evidence, compared to the verb *looks* with direct evidence. The three-way interaction indicated that this pattern did not differ across copy raised and unraised structures, suggesting that the effects found in the two-way interaction are consistent across both types of evidential constructions. Interestingly, the predictor Evidence Location did not improve model fit, suggesting that it is not a dimension that is relevant for the use of English evidentials.

Theoretical implications By carefully controlling for the evidence available in the context, we were able to use linguistic judgements to uncover the parameters that are relevant for the encoding of evidentiality in English. Our results validate Matthewson’s theory of parametrized evidential dimensions, and provide further empirical support for the evidential nature of copy raising structures in English. More generally, this experiment demonstrates the importance of using experimental methodology in investigating subtle semantic questions.

