Zlatomira Georgieva Ilchovska & Jennifer Culbertson

2019

The Learnability of Evidential Systems in the Case of L1 Bulgarian and L2 English

In

A. Gattnar, R. Hörnig, M. Störzer & S. Featherston (Eds.)

Proceedings of Linguistic Evidence 2018: Experimental Data Drives Linguistic Theory

Tübingen: University of Tübingen

https://publikationen.uni-tuebingen.de/xmlui/handle/10900/87132

In A. Gattnar, R. Hörnig, M. Störzer & S. Featherston (Eds.) (2019). *Proceedings of Linguistic Evidence 2018: Experimental Data Drives Linguistic Theory*. Tübingen: University of Tübingen. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/87132

The Learnability of Evidential Systems in the Case of L1 Bulgarian and L2 English

Zlatomira Georgieva Ilchovska[1,2] & Jennifer Culbertson[3]

[1]University of Oxford, [2]University of Birmingham, [3]University of Edinburgh
z_ilchovska@abv.bg, jennifer.culbertson@ed.ac.uk

1 Introduction

Evidentiality, or the grammatical category conveying the source of information of a statement, is present in about a quarter of the world's languages (Matsui & Fitneva, 2009). Evidentiality systems are diverse, and range from those distinguishing only two categories of source information, to those distinguishing five or more. Nevertheless, there are commonalities underlying these systems (Aikhenvald, 2004; Matsui & Fitneva, 2009). In this paper, we investigate the connection between universal features of evidentiality systems and cognitive biases which may affect their acquisition. More specifically, the current study will test the hypothesis that some evidentiality patterns are rare (or unattested) because they are not as learnable as more frequent ones. We do this using artificial language learning, a technique which has been used to investigate the relationship between typological frequency and learning in a number of different domains (e.g., word order Culbertson et al., 2012; case marking Fedzechkina et al., 2012; phonological alternations White, 2014).

The study was carried out with adult native speakers of Bulgarian who were also fluent in English. English does not encode evidentiality grammatically, but Bulgarian has a type of evidentiality system which is highly frequent in the world languages (Aikhenvald, 2004). We hypothesize that the learnability of different two-category (Aikhenvald, 2004) evidential-marking systems will correspond to their typological frequency. We also predict that knowledge of the Bulgarian evidential system will facilitate learning: the system most closely resembling the Bulgarian one (*non-first-hand versus "everything else"* – see Aikhenvald, 2004) will be the easiest to acquire, and the more the systems differ, the less successful their learning will be. To summarize, our findings suggest that the learnability of new evidential systems is largely influenced by whether participants already have acquired a grammatically evidential system in their native language rather than by the typological frequency of these systems.

2 Theoretical background

2.1 The concept of evidentiality

Aikhenvald (2004) states that every language is capable of expressing evidentiality but not every language has evidentiality encoded into its grammatical system. When viewed as a grammatical category, rather than as a purely conceptual and semantic one (Dendale & Tasmowski, 2001; Lazard, 2001), evidentiality refers to the grammatical marking of the source of information for a given statement (Aikhenvald, 2004; DeLancey, 2001, Dendale & Tasmowski, 2001). Such a marking indicates how the speaker has come to learn the information, and is usually conveyed through a morpheme. In comparison, some languages use lexical strategies. In English, markers of various statuses (adverbs, verbs or introductory clauses) like *I guess, I hear, the alleged, reportedly, it seems to me that*, etc. can serve this

function (Aikhenvald, 2004). For example, in English, if the speaker directly witnesses a man committing a crime, they might use a sentence like (1). However, if they did *not* directly witness this, they can indicate this using sentences like (2a) or (2b).

(1) **Direct context:** That was the man who committed the crime.
(2) **Indirect contexts**
  (a) That was the man who apparently committed the crime.
  (b) That was the man who reportedly committed the crime.

Although expressions like (1) and (2) can be used to express evidential meanings in English, according to Aikhenvald (2004), they are not evidentials because they are not obligatory and do not constitute a grammatical category. There are languages in which evidentiality is grammaticalised, but not obligatory (Lazard, 2001, 1999): the speaker has the choice between neutral or unmarked forms (omitting any commentary on the events) and marked forms. In most of these languages, evidentiality is expressed using perfect forms of the verbs, and Bulgarian is categorised as such a language (Lazard, 2001, 1999). Example (3)-(5) below illustrate how evidentiality is marked in Bulgarian.

(3) **Direct context**
Scenario: You see your friend Petya from far away, talking to another person on the street. As you approach the two, you hear your friend speaking French to the new person. Later on, you comment to your colleague:

Petya *govori* *frenski.*
Petya speak.3SG.PRS French.
'Petya speaks French.'

(4) **Inferential context**
Scenario: You visit your friend Petya and while in her house and looking at her bookshelf, you see a few books in French. Thus, you infer that Petya speaks French. Later on, you comment to your colleague:

Petya *govorela* *frenski.*
Petya speak.3SG.IPF.INDF.PST French.
'Petya speaks French, as I inferred.'

(5) **Reportative context**
Scenario: You speak to a friend of yours – Kalin. He tells you that he saw your mutual friend – Petya, doing a simultaneous interpreting for a French tourist group. Later on you comment to your colleague:

Petya *govorela* *frenski.*
Petya speak.3SG.IPF.INDF.PST French.
'Petya speaks French, as I heard.'

To summarize, evidentiality is not obligatorily marked in Bulgarian, but a certain proportion of the sentences in Bulgarian contain a grammaticalised marker of evidentiality (Matsui & Fitneva, 2009).

2.2 Types of evidentiality systems

The set of main abstract semantic categories which have been used to classify evidential information in previous literature (Dendale & Tasmowski, 2001; Matsui & Fitneva, 2009; Peterson, 2016; Plungian, 2001) is illustrated in Figure 1. The major division is based on whether the information conveyed is witnessed first-hand (direct) or not (indirect), however there are additional divisions within this. The current study will focus on *direct visual*, where direct information of the event is obtained through sight, *indirect inferential*, where the speaker only has access to another situation that points to the actual event of interest, and *mediated reportative*, where knowledge is received through somebody else.

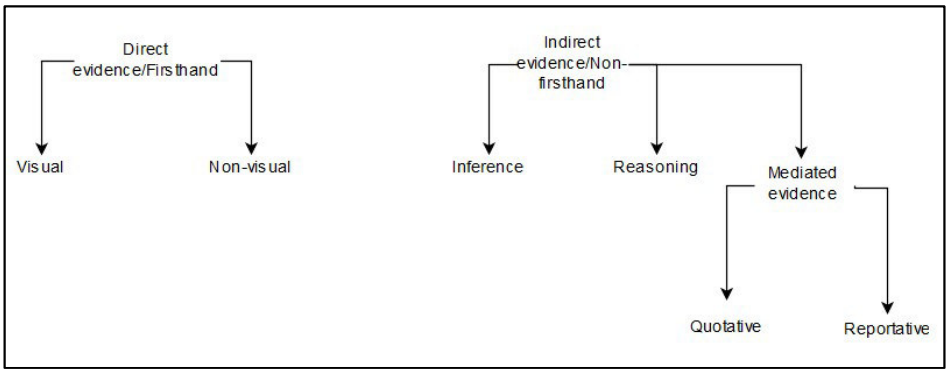

Figure 1. Major features used to classify evidential systems. Figure is made on the basis of Plungian (2001) and Peterson (2016)

Aikhenvald (2004) presents perhaps the most complete classification of evidential systems, according to the number of information-source types that they mark grammatically (and, therefore, that a speaker has to choose from). In the current work, we target the two-choice systems, which distinguish two types of marking of evidential knowledge, or alternatively have one category of knowledge source marked and the other functionally unmarked. There are three subtypes of the two-choice systems that are widespread. The first, which Aikhenvald (2004) calls *firsthand versus non-firsthand* (found in Cherokee, for example), marks directly-attested information differently from information received in *any* of the indirect manners. The second two-choice system, *non-firsthand versus "everything else"* (as in Turkish, for instance), uses an overt marker to indicate indirect information only, with all other sources of evidence being neutral (unmarked). The third system, *reported versus "everything else"* (for example in Estonian), only uses overt marking to indicate reportative indirect knowledge, with all other ways of obtaining the information being unmarked. Bulgarian could potentially be classified as either a *first-hand versus non-first-hand* system, or a *non-first-hand versus "everything else"* system, as the distinction involves present vs. imperfect marking (illustrated in (3)-(5) above). We will treat Bulgarian as *non-firsthand versus "everything else" system*, under the assumption that first-hand evidential forms in present tense are functionally unmarked (Aikhenvald, 2004; Friedman, 1999).

2.3 Typological frequency of evidentiality

According to the Universals Archive, there is an implicational hierarchy which describes the typology of evidential systems (Faller, 2002; Universität Konstanz, 2009):

(6) visual (direct) > non-visual > inferential > quotative/reportative (indirect)

According to (6), the direct visual source of information is treated as the most basic category. Importantly, if a language marks this basic category on the left side of the hierarchy, then it will also have markings on the non-basic categories to the right, but not vice versa. Comparing this hierarchy with Aikhenvald's (2004) evidentiality classification, it can be seen that the three wide-spread two-choice systems (including the one that Bulgarian represents) all conform to this hierarchy.

2.4 Acquisition of evidentiality in the lab

A number of previous studies have attempted to provide behavioral evidence for a link between constraints on learning and frequency of certain grammatical patterns in the world languages. The basic idea is that typological tendencies or universals reflect features of language which facilitate learning (e.g., see Culbertson, 2012). Bartell & Papafragou (2015) consider the typological prevalence of evidential systems in order to investigate whether this is related to the ease of acquisition of evidential systems. The authors test the three main information

sources listed in Figure 1: direct knowledge, inferential knowledge, and reportative knowledge within four possible evidential systems. What they term *System 1* and *System 2* corresponds to Aikhenvald's (2004) *Reported versus "everything else"* and *Non-firsthand versus "everything else"*, both of which are typologically widespread. These involve a single overt marker of indirect sources of evidence. In the case of *System 1*, the marker is used on reportative sources of evidence only, while *System 2* marks both the inferential and the reportative evidential sources with the same marker. Their *Systems 3* and *4* depict unattested evidentiality systems violating the hierarchy in (6). *System 3* marks direct and inferential information with the same morpheme and doesn't mark reported information. *System 4* carries an evidential marker only on direct information, and is thus arguably even worse from the perspective of (6). Bartell & Papafragou (2015) train English monolinguals on the distribution of an evidential marker *ga* across three types of scenarios concerning direct, inferential, and reportative knowledge. These scenarios are embedded within each of the four evidential systems and are marked (or not) according to the specific system's evidential structure. Their results showed that the English monolinguals learn *System 1* (*reported versus "everything else")* with the highest accuracy, but they find no difference in performance across the other three systems. However, the overall level of performance is notably low, suggesting that English speakers have difficulty learning such systems in the lab.

Following Bartell & Papafragou (2015), we will further investigate biases in the learning of evidential systems using participants who are already familiar with the notion of grammatical marking of evidentiality: native speakers of Bulgarian. Note that Bulgarian corresponds to *System 2* in Bartell & Papafragou (2015). These participants will be tested on how well they learn a novel pattern corresponding to *System 1, 2* and *4*. These three conditions were chosen because they allow us to test the main questions of interest. Although *System 3* presents an interesting opportunity to test the learnability of a non-existing evidential system that violates the universal hierarchy explained, it has not been included in the current study, as it is not directly relevant to the Bulgarian evidential structure and the research question of possible L2 transfer versus learning facilitation based on the typological frequency. *Systems 1* and *2* conform to the universal evidentiality hierarchy in (6). *System 4* does not conform to the hierarchy, but it presents the same cognitive distinction, namely direct versus indirect, as in Bulgarian.

Based purely on the typological frequency differences described in Section 2 above, *Systems 1* and *2* would be predicted to be the easiest to learn, while *System 4* would be predicted to be the hardest (as in 7).

(7) *System 1 = System 2 > System 4* (prediction based on typological frequency)

However, already having a grammatical evidential category system in one's native language is expected to facilitate the acquisition of other evidential systems, even if they are different in structure (Papafragou et al., 2007; Robinson, 2009). Research on first language acquisition of evidentiality suggests that formation of evidential categories happens early in childhood, and may have lasting effects on the cognitive processes related to knowledge and reasoning about sources (Aksu-Koç, 2000; Matsui & Fitneva, 2009; Papafragou et al., 2007; Robinson, 2009). We therefore expect that participants' native knowledge of Bulgarian will influence the acquisition of the three presented evidential systems. However, it could also be that their L2 knowledge of English will influence how they learn a third (albeit constructed) language in the lab. There are a number of theories regarding how L1 and L2 experience might influence learning of third-language systems (e.g., see Rothman et al., 2010 for a review). At least one of these emphasizes the role of the first (L1) language as the one providing the main morpho-syntactic-transfer material for the third language system (Na Ranong & Leung, 2009). Under this view, evidentiality systems resembling Bulgarian should be easier to learn, and English knowledge should not exert any additional influence. This leads to an alternative prediction regarding the ease of acquisition in our experiment, with *System 2* predicted to be the easiest to acquire, as it maps most closely onto the Bulgarian evidential system. *System 4* could po-

tentially also be facilitated by experience with the Bulgarian system, since it makes the same distinction (direct versus indirect information), even though it violates (6) by overtly marking *direct* information. By contrast, *System 1* would be expected to be hardest since it is typological common, but involves a distinction not used in Bulgarian (*Reported versus "everything else"*). This alternative prediction is summarized in (8).

(8) *System 2* > *System 4* > *System 1* (prediction based on L1 transfer)

3 Methodology

3.1 Participants

Participants were 36 Bulgarian-speaking adults (22 female; mean age 25 years). They had graduated or were currently studying in an English-speaking institution, within an English-speaking country (all participants had at least one year of education successfully passed). Data from the LEAP-Q (Marian et al., 2007) language-background questionnaire identified participants as Bulgarian native speakers, fluent in English.[1] Bulgarian was the first language by order of acquisition for all participants. Participants were randomly assigned to one of three experimental conditions, with 12 participants per group.

3.2 Materials

The experimental materials were 63 written-text scenarios in English, based on the pictorial scenarios used in Bartell & Papafragou (2015). These scenarios were equally divided into three categories: direct, inferential, and reportative (Bartell & Papafragou, 2015). The general structure of the three types of scenarios was as follows (with example stimuli that have been used):

(6) **Direct type**[2]
[Actor 1] enters the room. [Actor 2] enters the room. [Actor 1] [manipulates a smaller object in the room in relation to a larger object in the room]. [Actor 2] sees [the smaller object after it has been manipulated] and leaves the room to meet [Actor 3]. [Actor 2] tells [Actor 3] about the situation. [Actor 2] says: '[Actor 1] [smaller object] [larger object, with its location markers and prepositions] [action (the bare infinitive of the verb)] [an optional evidential marker *ga*]'.

Example stimulus
Sophie enters the room. Mina enters the room. Sophie places a book in the basket. Mina sees the book and leaves the room to meet Quinn. Mina tells Quinn about the situation. Mina says: 'Sophie book in the basket place (ga).'

(7) **Inferential type**
[Actor 1] enters the room. [Actor 1] [manipulates a smaller object in the room in relation to a larger object in the room]. [Actor 2] enters the room. [Actor 2] sees [the smaller object after it has been manipulated] and leaves the room to meet [Actor 3]. [Actor 2] tells [Actor 3] about the situation. [Actor 2] says: '[Actor 1] [smaller object] [larger object, with its location markers and prepositions] [action (the bare infinitive of the verb)] [an optional evidential marker *ga*]'.

[1] Some participants had native or non-native knowledge of other languages as well, however we disregard that here. Importantly, no participants spoke any language other than Bulgarian with grammatically encoded evidentiality.
[2] The square brackets indicate elements that change for every scenario and are replaced by the appropriate word. The evidential marker *ga* is either placed, or omitted, depending on the experimental condition.

Example stimulus
Samantha enters the room. Samantha places an apple on the board. Lucia enters the room. Lucia sees the apple and leaves the room to meet Curtis. Lucia tells Curtis about the situation. Lucia says: 'Samantha apple on the board place (ga).'

(8) **Reportative type**
[Actor 1] enters the room. [Actor 2] enters the room. [Actor 1] [manipulates a smaller object in the room in relation to a larger object in the room]. [Actor 2] sees [the smaller object after it has been manipulated] and leaves the room to meet [Actor 3]. [Actor 2] tells [Actor 3] about the situation. [Actor 3] says: '[Actor 1] [smaller object] [larger object, with its location markers and prepositions] [action (the bare infinitive of the verb)] [an optional evidential marker *ga*]'.

Example stimulus
Boris enters the room. Damian enters the room. Boris leaves an orange in the fridge. Damian sees the orange and leaves the room to meet Ali. Damian tells Ali about the situation. Ali says: 'Boris orange in the fridge leave (ga).'

The uniformity of the scenarios outlined above meant that the three different evidentiality types strongly resembled each other in structure. The scenarios were divided into a training set (27 scenarios) and a test set (36 scenarios). The language used to convey the utterances in each scenario was semi-artificial: comprised of English lexical items and a non-word evidentiality marker, *ga*. The decision to use a semi-artificial language was motivated by the aim to avoid the cognitive load of learning completely new words. We used semi-artificial English rather than Bulgarian to avoid actively confusing participants by omitting or inserting morphology on L1 words. As described above, the design was between-subjects, with three conditions. The conditions differed only in which scenario types used the evidential marker at the end of the actor's utterance. Each participant thus saw the same 63 scenarios, but depending on the condition that they were assigned into, the scenarios marked with *ga* varied. Recall that the original study (Bartell & Papafragou, 2015) had four conditions, while we use only three of them. In *System 1*, only the reportative cases are marked with *ga*; in *System 2* both inferential and reportative scenarios are marked with *ga*; in *System 4*, only the direct scenarios are marked with *ga*. To be consistent with the original study, we retain these condition labels.

3.3 Procedure

The experiment took place in a lab in the University of Edinburgh. Participants were first given the LEAP-Q, and then invited to complete the experimental part of the procedure, which was presented using E-Prime (Psychology Software Tools Inc., 2016). The experiment was divided into two parts: training and testing. In both parts, stimuli were randomized for each participant. Participants were given unlimited time to read the 27 training scenarios, and were asked to read carefully and try to figure out in which cases *ga* is used at the end of the actors' utterances. Then, during the testing phase, participants were instructed to respond by pressing 'Yes' or 'No' buttons to whether the *(ga)* in parenthesis at the end of each scenario utterance should or should not be used, based on the training they had just received.

Figure 2A-B shows an example of a training and a testing trial. More specifically, during the training, both slides remain for as much time as the participant needs them, until they press the 'b' key. However, during the testing, the first slide remains for as long as the subject needs it, until 'b' is pressed. The fixation cross remains for 200 milliseconds, and then the utterance with *(ga)* appears. Then participants must respond 'Yes' or 'No' within a 5-second time window.

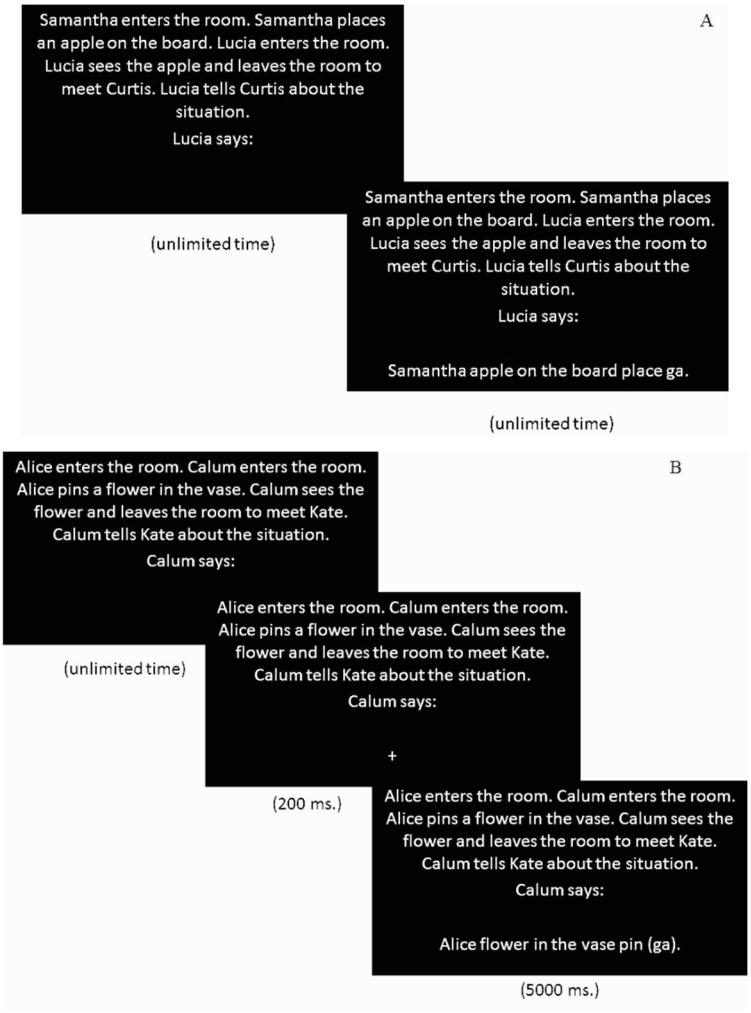

Figure 2: A. An inferential training scenario from *System 1* (where inferential and reportative situations are marked with *ga*). **B.** A direct testing scenario

4 Analysis

4.1 Accuracy

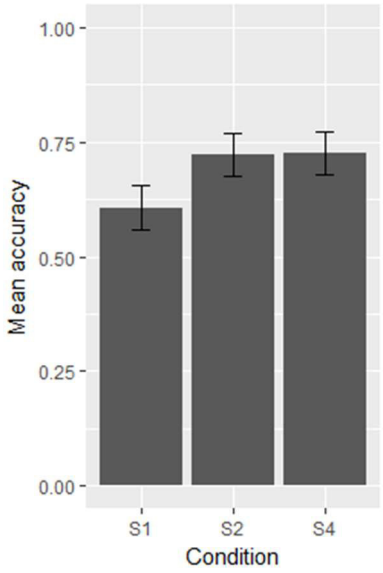

Figure 3. Mean accuracy of responses by condition. *S1, S2* and *S4* stand for the three semi-artificial language systems – *System 1, System 2* and *System 4*. Error bars represent the standard error of the mean

We first analysed whether participants in the three experimental conditions showed any differences in the extent to which they learned when the evidential marker *ga* was used in the language. This was calculated based on the accuracy of their button-press responses. Figure 3 shows the accuracy performance of participants by condition.

As shown in Figure 3, the differences between the performance in *Systems 2* and *4* were minimal. However, the performance in *System 1* was numerically less accurate. The participants' performance was therefore worse when they were presented with a typologically common evidentiality system marking only reportative information sources with *ga*.

Figure 4 presents the accuracy scores broken down according to the type of scenario (direct, inferential or reportative). Recall that in *System 1*, only the reportative cases are marked with *ga*; in *System 2*, both inferential and reportative scenarios are marked with *ga*; in *System 4*, only the direct scenarios are marked with *ga*.

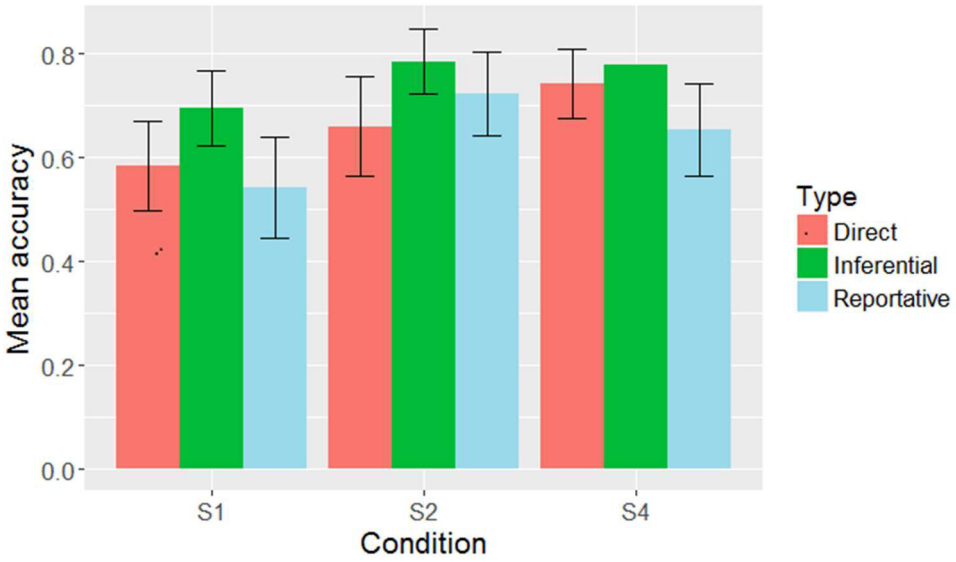

Figure 4. Mean accuracy of response by condition and scenario type. *S1, S2* and *S4* stand for the three semi-artificial language systems. Error bars represent the standard error of the mean

Figure 4 suggests that the performance on the three types of scenarios may differ to some degree across conditions. Although performance is lower in *System 1* for all three scenario types, the inferential scenarios were learned best across all the three conditions. The reportative type was relatively poor in both *System 1* (where only these scenarios carry a marker), and *System 4* (where they were unmarked).

The data were further analysed using a generalised linear mixed-effects model. The dependent variable *Accuracy* was modeled as a function of two fixed-effects: *Condition* and *Years of experience with English* (calculated by subtracting the age of acquisition of English for each participant from their current age, using the LEAP-Q data). Adding *Years of experience with English* was motivated by the possibility that the participants' length of use of English might affect their experimental results if extensive use of a second language without a grammaticalised evidential system might reduce participants' likelihood of inferring evidential marking in a new language. *Condition* was helmert coded, with *System 1* as the baseline level. In addition, random intercepts for *Subject* and *Item* were included in the model. The full model formula was therefore: *Accuracy ~ Condition + Condition * Years of experience with English + (1|Item) + (1|Subject)*. This model revealed no significant main effects or interactions, as shown in Table 1.

Table 1. Results for the fixed effects in the mixed model

Fixed effects	Estimate	Standard error	*z* value	*p* coefficient
Intercept (*System 1*)	1.99	1.86	1.07	.286
System 2 vs. *1*	-0.33	2.49	-0.13	.896
Mean of *System 1* and *2* vs. *4*	-1.15	1.18	-0.97	.330
Years of experience with English	-0.04	0.11	-0.40	.689
System 2 vs. *1* * Years English	0.03	0.15	0.23	.819
Mean of *System 1* and *2* vs. *4* * Years English	0.07	0.07	1.01	.311

4.2 Response time

Response time data was also collected and analysed to determine whether participants differed across conditions in the speed with which they decided whether *ga* was required. First of all, inaccurate and missing responses were removed from the data: 39 % for *System 1,* 28 % for *System 2,* and 28 % for *System 4* (out of a total number of 432 cases per condition). Figure 5 shows the response-time scores for participants by condition. According to those scores, subjects had the slowest response times in *System 1*, and the fastest in *System 4*, matching the accuracy results.

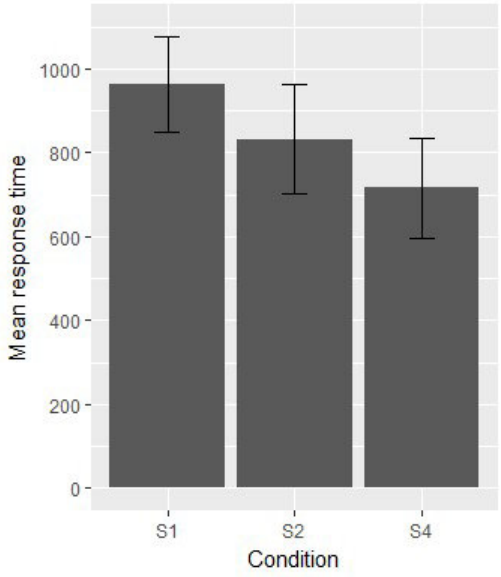

Figure 5. Mean response time by condition. *S1, S2* and *S4* stand for the three semi-artificial language systems. Error bars represent the standard error of the mean

Looking at the response times for each scenario type by condition (see Figure 6), the numerical results also match the general pattern seen in the mean accuracy scores per condition: participants in *System 1* were slowest in all cases, and reportative trials were slowest for both *System 1* and *4* participants.

These data were analysed using a linear mixed-effects model in which the dependent variable *Response Time* was modeled as a function of *Condition* and *Years of experience with English*. Again, random intercepts for *Subject* and *Item* were included. The final formula of the model was: *Response Time ~ Condition + Condition * Years of experience with English + (1|Item) + (1|Subject)*. As with the accuracy data, no significant main effects or interaction were found (see Table 2).

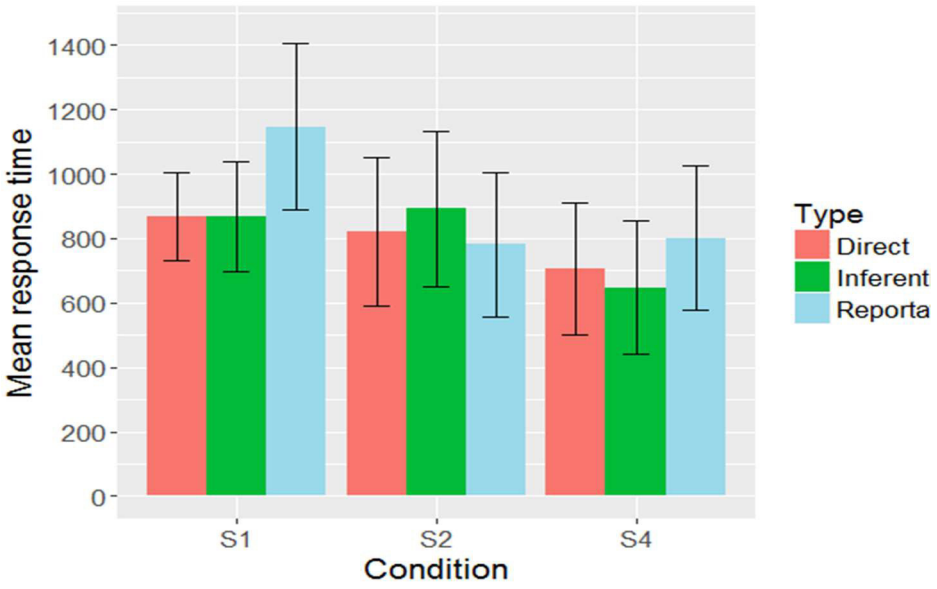

Figure 6. Mean response time by condition and scenario type. *S1, S2* and *S4* stand for the three semi-artificial language systems. Error bars represent the standard error of the mean

Table 2. Results for the fixed effects in the mixed model. As the *lmer* function in R does not directly provide *p* values, they were calculated using the two-tailed-test formula: *2 * (1 – pt (abs (t value), df))* (Baayen, 2008)

Fixed effects	Estimate	Standard error	*t* value	*p* coefficient
Intercept (*System 1*)	431.94	755.70	0.57	.567
System 1 vs. 2	1102.28	1009.08	1.09	.275
Mean of *System 1*, *2* vs. *4*	- 652.98	481.32	-1.36	.175
Years English	18.16	44.33	0.41	.682
System 1 vs. 2 * Years of English	- 65.85	60.21	-1.09	.274
Mean of *System 1*, *2* vs. *4* * Years English	33.96	27.51	1.23	.218

4 Discussion

We tested whether native speakers of Bulgarian were able to learn novel systems of evidentiality marking, and whether any differences in the learnability of such systems could be connected to typological frequency or L1 experience. While no significant effects were found, there are trends which differ relative to what Bartell & Papafragou (2015) found in their original study with monolingual English speakers, and which shed some light on the mechanisms underlying acquisition of evidentiality systems in the lab. Recall that, based on the typological frequency data, *Systems 1* and *2* were predicted to be easier than *System 4*. In Bartell & Papafragou (2015), participants were indeed most successful at learning *System 1* (only reportative marked with *ga*). In our study, by contrast, Bulgarian speakers clearly did not show any advantage for *System 1*, in fact, it was numerically the least successfully acquired evidential system. This suggests that similarity to the L1 is indeed playing some role here. We suggested that *Systems 2* and *4* might be relatively easy for Bulgarian speakers since they both make the basic-level distinction between direct and indirect information. Thus, even though these systems use marking in a different way, they both reflect a potentially salient cognitive and semantic distinction between direct and indirect information (Aikhenvald, 2004; De Haan, 2001; Faller, 2002; Papafragou et al., 2007; Universität Konstanz, 2009). Further research would be needed to verify whether this is purely the result of prior knowledge from Bulgarian, or represents a more general cognitive preference. The findings from Bartell & Papafragou (2015) suggest the former, however, as we noted previously, the overall success in learning in that study was very low (Note that while the monolinguals learned *System 1*

with approximately 78 % accuracy, *System 2* and *System 4* were acquired with, accordingly, 55 % and 51 % accuracy.). Further evidence for the role of L1 comes from Bulgarian speakers' apparent difficulty with the typologically common *System 1*. Under this explanation, this system is expected to be hardest because it involves a distinction not used in Bulgarian (*Reported versus "everything else"*).

Interestingly, there is evidence from L1 acquisition of evidentiality systems that children produce and comprehend direct evidentials earlier compared to indirect evidentials (Matsui & Fitneva, 2009). It is possible that the later development of metalinguistic awareness for indirect evidentials is related to a higher degree of abstractness of the inferential and reportative categories, while seeing is understood as a source of knowledge in early childhood. Following from this, marking the direct evidentials with *ga* in *System 4* could have made more cognitively and semantically salient an evidential category that is in any case the earliest and easiest to acquire. However, our study does not offer any clear support for this given that the lack of significant differences between conditions. In fact, direct evidential scenarios were not generally learned with higher accuracy; moreover, the inferential scenarios appeared to be the easiest to acquire in all three conditions. One possibility is that our results reflect the way that the stimuli were structured. While the difference between the direct and reportative was only at the end of the scenarios (in the identity of the actor uttering the comment), the inferential scenarios had a notably different structure in the order of its second and third sentence (see examples (6)-(8)). Consequently, it could be that this slightly different structure made the inferential scenarios easier to distinguish and remember. Finally, note that the reportative type was numerically most accurately acquired in *System 2,* where it was marked together with the inferential type, and hardest to acquire in *System 1*, where only the reportative was marked. Another factor that could have influenced the results might have been the perceived intentionality of the actions in the situations, as well as the types of prepositions that were used. As these factors were reported by a number of participants after the experiment, they might also need to be considered in future to assure no influence on participants' judgements on the marking. To summarize, the current study, along with previous work by Bartell & Papafragou (2015), fail to provide clear evidence for the role of learning in shaping the frequency of evidentiality systems cross-linguistically. Rather, we have found some evidence, though it remains weak, that the transfer from the L1 may be the main factor influencing learnability of these systems. Interestingly, our results point to the possibility that the presence of a grammatically encoded high-level distinction between direct and indirect (likely learned from the L1) facilitates learning regardless of the pattern of overt marking.

References

Aikhenvald, A. Y. (2004). *Evidentiality*. Oxford: Oxford University Press.

Aksu-Koç, A. (2000). Some aspects of the acquisition of evidentials in Turkish. In L. Johanson & B. Utas (Eds.), *Evidentials: Turkic, Iranian and Neighbouring Languages* (pp. 15-28). Berlin: De Gruyter.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R.* Cambridge University Press.

Bartell, S., & Papafragou, A. (2015*). Cross-linguistic variation and the learnability of semantic systems*. Presented at the '40th Boston University Conference on Language Development', November 2015, Boston, MA.

Culbertson, J. (2012). Typological universals as reflections of biased learning: Evidence from artificial language learning. *Language and Linguistics Compass, 6*(5), 310-329.

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition, 122*(3), 306-329.

De Haan, F. (2001). The place of inference within the evidential system. *International Journal of American Linguistics, 67*(2), 193-219.

DeLancey, S. (2001). The mirative and evidentiality. *Journal of Pragmatics, 33*, 369-382.

Dendale, P., & Tasmowski, L. (2001). Introduction: Evidentiality and related notions. *Journal of Pragmatics, 33*, 339-348.

Faller, M. T. (2002). *Semantics and Pragmatics of Evidentials in Cuzco Quechua*. Dissertation, Stanford University. http://personalpages.manchester.ac.uk/staff/martina.t.faller/

Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences of the United States of America, 109*(44), 17897-17902.

Friedman, V. A. (1999). Proverbial evidentiality: On the gnomic uses of the category of status in languages of the Balkans and the Caucasus. *Mediterranean Language Review, 11*, 135-155.

Lazard, G. (1999). Mirativity, evidentiality, mediativity, or other? *Linguistic Typology, 3*, 91-109.

Lazard, G. (2001). On the grammaticalization of evidentiality. *Journal of Pragmatics, 33*, 259-367.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940-967.

Matsui, T., & Fitneva, S. A. (2009). Knowing how we know: Evidentiality and cognitive development. In A. S. Fitneva & M. Tomoko (Eds.), *Evidentiality: A Window into Language and Cognitive Development* [Special issue]*, New Directions for Child and Adolescent Development, 125* (pp. 1-11). San Francisco, CA: Jossey-Bass.

Na Ranong, S., & Leung, Y.-K. I. (2009). Null objects in L1 Thai-L2 English-L3 Chinese: An empiricist take on a theoretical problem. In Y. I. Leung (Ed.), *Third Language Acquisition and Universal Grammar* (pp. 162-191). Bristol: Multilingual Matters.

Papafragou, A., Li, P., Choi, Y., & Han, C.-H. (2007). Evidentiality in language and cognition. *Cognition, 103*, 253-299.

Peterson, T. (2016). Mirativity as surprise: Evidentiality, information, and deixis. *Journal of Psycholinguistic Research, 45*, 1327-1357.

Plungian, V. A. (2001). The place of evidentiality within the universal grammatical space. *Journal of Pragmatics, 33*, 349-357.

Psychology Software Tools, Inc. [E-Prime 3.0] (2016). Retrieved from https://www.pstnet.com

Robinson, E. J. (2009). Commentary: What can we learn from research on evidentials? *Evidentiality: A window into Language and Cognitive Development* [Special issue]*, New Directions for Child and Adolescent Development, 125*, 95-103.

Rothman, J., Iverson, M., & Judy, T. (2010). Introduction: Some notes on the generative study of L3 acquisition. *Second Language Research, 27*(1), 5-19.

University of Konstanz (2009). The universals archive. https://typo.uni-konstanz.de/archive/intro/index.php

White, J. (2014). Evidence for a learning bias against saltatory phonological alternations. *Cognition*, *130*(1), 96-115.