Are Cross-Linguistically Frequent Semantic Systems Easier to Learn? The Case of Evidentiality

Dionysia Saratsli (Dsaratsl@Udel.Edu)
Department of Linguistics and Cognitive Science, 125 Main Street
Newark, DE 19702 USA

Stefan Bartell (Sbartell@Udel.Edu)
Department of Linguistics and Cognitive Science, 125 Main Street
Newark, DE 19702 USA

Anna Papafragou (Papafragou@Psych.Udel.Edu)
Department of Psychological and Brain Sciences, 105 The Green
Newark, DE 19702 USA

Abstract

It is often assumed that cross-linguistically more prevalent distinctions are easier to learn (*Typological Prevalence Hypothesis* - TPH). Prior work supports this hypothesis in phonology, morphology and syntax but has not addressed semantics. Using an Artificial Language Learning paradigm, we explore the learnability of semantic distinctions within the domain of evidentiality (i.e. the linguistic encoding of information sources). Our results support the TPH, since the most prevalent evidential system was learned best while the most rare evidentiality system yielded the worst learnability results. Furthermore, our results indicate that, cross-linguistically, indirect information sources seem to be marked preferentially (and acquired more easily) compared to direct sources. We explain this pattern in terms of the pragmatic need to mark indirect, potentially more unreliable sources over direct sources of information.

Keywords: evidentiality; artificial language learning; learnability; semantics; information sources

Learnability and the Typological Prevalence Hypothesis (TPH)

It is often assumed in the literature that linguistic distinctions that are encountered more frequently across different languages share some characteristics that make them easier to learn than others (Jacobson, 1971; Rosch, 1972; Clark, 1976; Pinker, 1984). This idea has been captured effectively by Gentner and Bowerman's (2009, p.467) *Typological Prevalence Hypothesis* (TPH): "All else being equal, within a given domain, the more frequently a given way of categorizing is found in the languages of the world, the more natural it is for human cognizers, hence the easier it will be for children to learn". Gentner and Bowerman (2009) tested this hypothesis within the spatial domain, comparing how English-speaking and Dutch-speaking children acquire their native language's support prepositions. English and Dutch differ in the number of prepositions they use to express spatial support: Dutch utilizes three different prepositions (*op, aan, om*) to express the same meanings that English encodes with the single preposition *on*. Importantly, these two support systems differ in their typological prevalence, with the English preposition system being more typologically common. The TPH therefore predicts that the English preposition system should be more easily learned than the Dutch system. Gentner and Bowerman's results support this prediction. One issue with this conclusion, however, is that the slower acquisition rate could be due to the increased number of subcategories found in Dutch compared to English as opposed to an inherent learnability asymmetry of semantic categories per se. This language asymmetry complicates the interpretation of Gentner and Bowerman's results and hence the evidence in favor of TPH.

In this paper, we offer a new test of TPH using an Artificial Language Learning Paradigm. This type of experimental design often requires participants to learn different versions of a target language that differ minimally from each other in terms of a grammatical or lexical feature (see Folia, Uddén, de Vries, Forkstam, & Petersson, 2010 for a review). Typically, this design includes an initial learning phase in which learners are exposed to the grammar/lexicon of the artificial language, usually with the help of visual stimuli. The learning phase is followed by a test phase in which the extent to which participants learned the linguistic target is assessed. This paradigm offers a unique opportunity to explore the participants' learning process in relation to a specific linguistic feature of interest (Fedzechkina, Newport & Jaeger, 2016). By having participants learn minimally different versions of the same artificial language, one can bypass the role of frequency in the learnability of attested systems in individual languages, such that any learnability pattern that surfaces can be more directly tied to the inherent characteristics of the cross-linguistic distinction that is being explored. Moreover, it is possible to have adults learn the target artificial language which in turn eliminates the possibility that any learnability patterns observed could be due to cognitive-developmental limitations in the learners themselves.

Previous studies using an Artificial Language Learning paradigm have confirmed that cross-linguistically common distinctions are learned more easily than less common ones in the domains of syntax, phonology and morphology

(Newport & Aslin, 2004; Wonnacott, Newport, & Tanenhaus 2008; Merkx, Rastle, & Davis, 2011; Culbertson, 2012; Tabullo, Arismendi, Wainselboim, Primero, Vernis, Segura, Zanutto & Yorio., 2012; Culbertson & Newport, 2015;). Nevertheless, within the domain of semantics (which was the main focus of TPH), this hypothesis remains to be tested systematically. Here we address this open issue. We focus on a semantic domain that is not grammaticalized in English and can be taught to adults within an Artificial Language Learning paradigm without native language interference: the domain of evidentiality, i.e., the linguistic encoding of information source.

Evidentiality and TPH

Languages differ in the way they encode evidentiality: some languages like English make use of lexical means such as verbs (e.g., *see, hear, infer*) or adverbs (e.g., *allegedly, reportedly*) to mark information sources. Other languages use a set of grammatical morphemes to indicate information sources in an utterance. There are three common types of evidential morphemes depending on which information source is marked: Visual (firsthand/perceptual evidence), Inferential (inference based on evidence), and Reportative (hearsay) (Willett, 1988; Papafragou, Li, Choi & Han, 2007; deHaan, 2013b; Aikhenvald, 2018). In the Wanka Quechua examples below, -*mi* in (1) marks the speaker's direct visual experience of the event, -*chr*- in (2) marks an inference drawn by the speaker and –*shi* in (3) marks another person's report about what happened (Aikhenvald, 2004):

(1) Chay-chruu-mi achka wamla-pis walashr-pis alma-ku-lkaa-ña.
this-LOC-DIR.EV many girl-TOO boy-TOO bathe-REEL-IMPF.PL-NARR.PAST.
'Many girls and boys were swimming' (I saw them).

(2) Daañu pawa-shra-si ka-ya-n-chr-ari.
Field finish-PART-EVEN be-IMPF-3-INFR-EMPH.
'It (the field) might be completely destroyed' (I infer).

(3) Ancha-p-shi wa'a-chi-nki wamla-a-ta.
too.much-GEN-REP cry-CAUS-2 girl-1P-ACC.
'You make my daughter cry too much' (they tell me).

Across languages that grammatically mark only <u>one</u> type of information, evidential systems that involve only Reportative morphemes are the most widespread ones; systems that use an indirect morpheme to mark inference or reports are less frequent (Papafragou et al., 2007; deHaan, 2013a; Aikhenvald, 2004, 2018; Ünal & Papafragou, 2018;). Evidential systems that only have Visual morphemes are rare (Aikhenvald, 2018). The reasons for this asymmetry have not been discussed extensively but might be connected to the pragmatic need to mark indirect, probably unreliable sources but not direct/perceptual, and hence more reliable, experience (Dancy, 1985; and discussion below).

Here we used an Artificial Language Learning paradigm to compare the learnability of three evidential systems (see Table 1): 1) a system in which a grammatical morpheme is used only when the speaker has full direct visual access to what happened (*Visual* System), 2) a system where a grammatical morpheme is used only when the speaker infers what happened based on some visual cues (*Inferential* System), and 3) a system in which a grammatical morpheme is used only when the speaker obtains information by another person (*Reportative* System). Based on the typological frequency patterns for evidential systems reviewed earlier, the TPH predicts that the Reportative system should be the most learnable and the Visual system the least learnable (with the Inferential system falling somewhere in-between). The experiment that follows tested these predictions.

Table 1: Evidential Systems.

Evidential System	Speaker's Information Access		
	Visual Perception	Inference	Report
Visual	morpheme		
Inferential		morpheme	
Reportative			morpheme

Experiment

Our experiment consisted of two phases following the general Artificial Language Learning experimental design: a Training Phase and a Testing Phase. In the Training Phase, participants were exposed to one of the three evidentiality systems in Table 1 and had to figure out when the evidential marker was used. In the Testing Phase participants were evaluated on how well they had learned the target evidentiality system through a Production and a Comprehension Task.

Participants. We recruited 101 participants between the ages of 18 and 22. All participants were undergraduate students at the University of Delaware and were enrolled in an Introductory Psychology course that awarded credit for their participation.

Stimuli and Procedure. For the Training Phase, we filmed 21 videos in three versions each, with each version corresponding to a type of information access (Visual Perception, Inference, Report). In each video there were three characters; across videos, they were played by the same three female undergraduate research assistants. The roles of these characters were consistent across the videos: one of the characters (henceforth the "Agent") performed an event using some materials and then put these materials away. The second character accessed the event in one of several ways and would later describe the event (henceforth the "Speaker"). A third character manipulated the Speaker's access to the event (e.g., either allowed the Speaker to watch

Figure 1: Sample screenshots from one Training Phase video shown in 3 versions corresponding to Access types: (A) Visual Perception, (B) Inference, (C) Report. Across Access types the video ended with the Speaker producing a sentence (Panel 5) that either included or omitted an evidential (e.g., "She drawing copiedga", "She drawing copied").

the event or blocked her visual access for the complete duration or part of the event). The setting was identical for all the videos: the Agent and the Speaker were sitting on different sides of a table while the third character stood behind them in full view of the table. Each video was approximately 15 seconds long. At the end, the Speaker turned to the camera and described what happened. At that point, the video stopped and a speech bubble appeared with an artificial language sentence, and stayed there for 7 seconds before the next video began.

Figure 1 shows a sample event in which the Agent copied a drawing (the Speaker is pictured in a blue shirt). In the Visual Perception version (series A), the Speaker had continuous direct visual access to the event (A1 began with occluded access to ensure that the hands-over-eyes would not be an easy-to-detect difference among access types, but the hands are removed from the Speaker's face immediately). In the Inference version (series B), the Speaker had visual access only for the beginning and the end of the event (panels 1 and 4), but her access was blocked for the middle portion (panels 2 and 3); therefore, she could infer what happened from the last stage of the event. In the Report version (series C), the Speaker's visual access was blocked throughout the event (panels 1-3); later (panel 4), the Speaker got a report about what had happened from the third character. All videos ended by displaying the Speaker's artificial-language description of what happened within a speech bubble (panel 5). The artificial language shared the same vocabulary with English (for simplicity's sake) but had a different syntactic structure (Subject-Object-Verb) and lacked function words. A novel verb-final morpheme, *ga*, appeared when appropriate as a marker for evidentiality.

We designed 3 evidential systems to be acquired (Visual, Inferential, Reportative) by having the Speaker describe only one type of Access with an evidentially marked sentence (e,g, *She drawing copiedga*, as in Figure 1) and include no marker for the other two Access types. For instance, for the Visual System, only the sentences in the Visual access versions included -*ga*. Then for each evidential system, we created 3 basic lists for the Training Phase (for a total of 9 lists): each basic list contained 21 videos, with 7 videos per Access type. Across lists, the videos rotated through each Access type. For instance, if the video in Figure 1 was shown in the Visual Perception version for list 1, then the same video was shown in the Inference version for list 2 and the Report version for list 3. The presentation order of the videos was randomized across lists.

We randomly assigned participants to one of 3 conditions depending on the System they were exposed to (n = 34 for the Inferential and Reportative System, and n=33 for the Visual System). Each participant was given one of the 9 stimulus lists. We tested participants in small groups, in a dimly lit, quiet room. Participants were told that they would watch some videos and one character would describe the videos in an "alien language". This language would share some words with English but would be different in several ways and would contain a special marker, *ga*. Their task was to pay attention to when *ga* appeared in order to try and figure out what it meant.

When the Training Phase was over, the Testing Phase began. Participants had to complete both a Production and a Comprehension task. For these tasks, we filmed new videos that were similar to those for the Training Phase (except for some features of the language in the event descriptions – see below).

For the Production task, we used 12 new videos, each filmed in 3 different versions corresponding to the 3 Access types. We arranged these stimuli into 3 basic lists, with each list containing 12 videos, 4 per Access type. As in the Training Phase, the lists were created by rotating each video through the three different Access types. For each basic list, three randomized presentation orders were created, resulting in 9 presentation lists in total. Within each condition, participants were assigned to one of these lists. As mentioned already, the structure of the videos in the Production task was identical to the Training Phase but when the speech bubble appeared at the end, the evidential marker was replaced by a gap next to the verb. Using an answer sheet, participants had

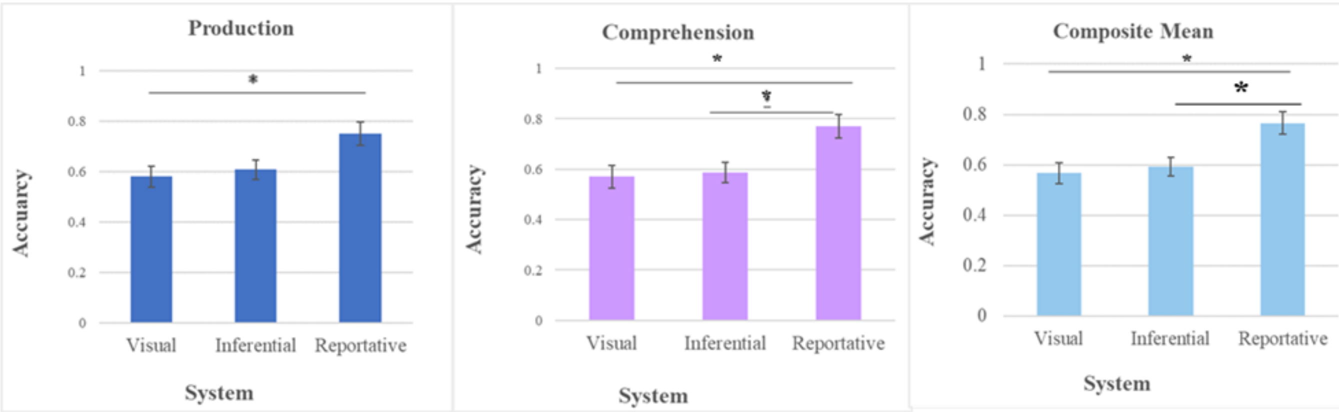

Figure 2. Accuracy Means Across Systems. The composite score represents a combined Production/Comprehension score. Error bars represent ± 1 S.E.

to write down the verb either with or without *ga* depending on whether they thought it was needed to correctly complete the character's phrase.

For the Comprehension task, we used 36 new videos, each filmed in 3 different versions corresponding to the 3 Access types. We arranged these stimuli into 3 basic lists, with each list containing 36 videos (12 per Access type) using the rotation method described above. Similarly to the Production task, for each of these basic lists, 3 lists with a unique randomized presentation order were created (9 lists in total). In half of the videos within each list (and within each Access type), the Speaker erroneously used the marker *ga*: she either failed to use the marker when she should have or used it for the wrong types of Access. In the remaining videos, the use of the marker was correct. Within each condition, participants were assigned to one of the presentation lists. The participants' task was to write 'yes' or 'no' in their response sheet to indicate whether or not they thought the character was using the marker correctly. At the end of the experiment, we asked participants to write down what they thought that the marker *ga* meant and when it was/was not used.

Results

Participants' responses were coded for accuracy. We calculated the accuracy means for each System. In addition, we averaged each participant's Production and Comprehension score yielding a Composite accuracy mean across tasks. We subsequently calculated a Composite Mean per System. The results can be seen in Figure 2.

For the Production task, a one-way ANOVA with System as a factor revealed a main effect of System ($F(2,98)= 4.771$, $p<0.05$). Pairwise comparisons using Bonferroni corrections revealed a significant advantage of the Reportative over the Visual System (p=.014) but no significant difference between either the Inferential and the Visual System (p=1.0), or the Inferential and the Reported System (p=.058).

For the Comprehension task, the same ANOVA revealed a main effect of System ($F(2,98)=6.509$, $p<0.01$). Pairwise comparisons (Bonferroni corrections) showed an advantage of the Reportative System over both the Inferential (p=0.01) and the Visual System (p=.005). However, there was no statistically significant difference between the Visual and Inferential System (p=1.0).

Lastly, a one-way ANOVA conducted on composite Production and Comprehension means revealed an effect of System ($F(2,98)=6.535$, $p<0.01$). Pairwise comparisons (Bonferroni corrections) revealed again a significant advantage of the Reportative System over both the Visual (p=.004) and the Inferential System (p=0.01).

Participants' answers about the meaning of the marker reflect the results' pattern: out of the 34 participants exposed to the Reportative System, 21 correctly associated the marker with reportative access, specifically alluding to the *speaker's* mental state by mentioning that she "was told" about the event. Of the 33 participants of the Visual system, only 12 associated the marker with speaker's direct visual experience of the event. only Similarly, only 9 out of the 34 participants exposed to the Inferential System correctly associated the marker with the character inferring the action. Across systems, participants that did not identify the correct marker meaning, associated the marker with some type of grammatical distinction (e.g., singular/plural forms, past or completed actions, articles such as *the*/*a*) or associated it with the incorrect type of access. Overall, these responses show that participants associated evidential meanings with the marker, but they did so much more consistently for the Reportative System.

Discussion

Our goal was to test the assumption that the frequency of cross-linguistic semantic patterns is related to the inherent learnability of these patterns, an assumption captured in Gentner and Bowerman's (2009) TPH. Using an Artificial Language Learning paradigm, we set out to compare the learnability of evidential semantic systems, focusing on those that encode a single type of information source (Table 1). The most typologically common evidential system within this group (and also the single most prevalent type of evidential system in general; Aikhenvald, 2018) is the Reportative system in which a marker is used only for the least direct type of access to information – namely, the cases when the speaker conveys information reported by another person. The least common system is the Visual system in which only direct visual access to an event is marked morphologically. In our study, as predicted by TPH, the Reportative system was learned more easily by our participants compared to the Visual system. Our experiment offers strong evidence for the conclusion that highly frequent semantic distinctions are more learnable than less frequent ones. Furthermore, it adds to previous studies that have studied learnability with the same methodological paradigm within the domains of syntax, phonology and morphology.

Not all aspects of our data are compatible with the predictions of TPH. Specifically, even though exclusive encoding of visual evidentials is rare, and there is a broad preference to mark non-visual/indirect over visual/direct sources cross-linguistically (Aikhenvald, 2018), the Inferential and Visual systems were equally learnable in our data. A possible explanation for this outcome lies with the fact that our Inference videos contained strong visual clues to what happened, bringing this type of information access closer to a direct perceptual experience than to an indirect inference on the speaker's part. This explanation is in line with several findings from a recent study by Ünal, Pinto, Bunger and Papafragou (2016). In that study, when English speakers had to state how they had found out about an event, they stated having seen events that they had experienced in their entirety. However, when they had only seen the beginning and aftermath of an event and had to "fill in" the event from these visual cues, their statements varied. Closer inspection suggested that, when the visual cues were indeterminate, participants consistently stated that they had inferred the event; but when the visual cues were more determinate and highly constrained the inference, participants were equally likely to say that they had seen vs. inferred the event. The authors proposed that there are several varieties of inference, and that stronger, more constrained (and thus more secure) inferences from visual cues might be difficult to distinguish from purely perceptual experience. These varieties of inference had implications for evidential language: Ünal et al. (2016) found that these different types of inference impacted the use of evidential morphology by speakers of Turkish, a language that grammaticalizes evidentiality. Furthermore, inference types had effects on memory: building on classic studies showing that people often have a false memory of having actually experienced events that they have only inferred (Johnson, Hashtroudi, & Lindsay, 1993; Hannigan & Reinitz, 2001; cf. Strickland & Keil, 2011), Ünal et al. (2016) found that, across English and Turkish speakers, such misattributions to perception were more common when inferences were strongly constrained by visual cues and thus harder to distinguish from pure perception. This line of reasoning leads to the prediction that replacing Inference scenarios in our paradigm with less direct cases of inference from visual cues (e.g., footsteps on snow) should allow the learnability difference between the Visual and Inferential systems to emerge.

On a broader level, our results raise questions about the origins of the typological generalizations in the domain of evidentiality. According to the basic observation motivating the present work, across languages, the least formally marked source of information is visual, or direct access (Aikhenvald, 2018, a.o.). Why should this be so? One possibility is that "the tendency to mark direct, or visual, or sensory evidentials less than others may reflect the primacy of vision as an information source" (Aikhenvald, 2018, p.16). Direct perceptual experience of an event is regarded as a very reliable source because it is assumed to correspond to reality (Dancy, 1985). Additionally, developmental research suggests that children draw the connection between seeing and knowing from early on (Pillow, 1989; Pratt & Bryant, 1990; Ozturk & Papafragou, 2016), which highlights the the primacy of visual perception as an information source. Relatedly, indirect sources of information such as inference or reports are deemed more peripheral and less reliable in the sense that the former may be based on incomplete premises while the latter depends on the informant's reliability (Dancy, 1985; Koring & De Mulder, 2014; Papafragou et al., 2007; Matsui & Fitneva, 2009; McCready, 2015; Aikhenvald, 2018; Wiemer, 2018;). This has been found to be true even for languages that do express information access through *perception verbs* and not through obligatory grammatical morphemes (Lesage, Ramlakhan, Toivonen & Wildman, 2015). According to some researchers (Sperber, Clement, Heintz, Mascaro, Mercier, Origgi & Wilson, 2010), human cognition uses epistemic vigilance as a mechanism to avoid unreliable sources and the risk of being misinformed. However, exercising epistemic vigilance could entail an additional processing cost: listeners would have to give up the assumption that the communicative exchange they are engaged in offers truthful, informative contributions and would need to evaluate not only the actual information they receive but also their interlocutor's reliability and intentions. Thus, pragmatic pressures to mark sources of information would affect indirect and probably unreliable sources more than direct/perceptual, and hence more reliable, experience.

If this perspective is on the right track, our results would support a more nuanced version of TPH. Recall that, on Gentner and Bowerman's original proposal, the roots of TPH lie in the cognitive naturalness of the semantic classes that the learner acquires. Here we have proposed a broadened notion of naturalness that also includes pragmatic (and not only conceptual) factors. In our studies, adult learners acquired semantic systems of varying cross-linguistic frequency but both the frequency patterns and the learnability outcomes were pragmatically (not conceptually) motivated.

If the frequency patterns for linguistic evidentiality systems reflect the pragmatic need for information source marking, as we have suggested, a further prediction follows: it might be possible to obtain similar learnability patterns even if we used a non-linguistic marker to encode information source (e.g., a pictorial symbol). We are currently pursuing this possibility in ongoing work.

Acknowledgements

This work was supported by a Leventis Foundation Graduate Educational Grant (Dionysia Saratsli) and NSF Grant #1632849 (Anna Papafragou). We would like to thank the undergraduate research assistant Casey Corallo for her help in collecting and coding data for this project.

References

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

Aikhenvald, A. (2018). Evidentiality: The Framework. In A.Y. Aikhenvald (Ed.), *The Oxford Handbook of Evidentiality*, 1-43. Oxford: Oxford University Press.

Clark, E. (1976). Universal categories: On the semantics of classifiers and children's early word meanings. In A. Juilland (Ed.), *Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth birthday*, Volume 3: Syntax, 449–462. Saratoga, CA: Anna Libri.

Culbertson, J. (2012). Typological Universals as Reflections of Biased Learning: Evidence from Artificial Language Learning. *Linguistics and Language Compass*, 6(5), 310–329. https://doi.org/10.1002/lnc3.338.

Culbertson, J., & Newport, E. (2015). Harmonic biases in child learners: In support of language universals. *Cognition*, 139, 71–82.

Dancy, J. (1985). *An introduction to contemporary epistemology.* Oxford: Blackwell.

de Haan, F. (2013a). Coding of evidentiality. In M.,Dryer and M.,Haspelmath, (Eds.) *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.

de Haan, F. (2013b). Semantic distinctions of evidentiality. In Dryer, M. and Haspelmath, M. (Eds.), *The World Atlas of Language Structures Online.* Max Planck Institute for Evolutionary Anthropology, Leipzig.

Fedzechkina, M., & Newport, E. (2016). Miniature artificial language learning as a complement to typological data. *The Usage-Based Study of Language Learning and Multilingualism*, 1–22. Retrieved from https://books.google.com/books?hl=en&lr=&id=88gvDAAAQBAJ&oi=fnd&pg=PA211&ots=I6oc-Bwvl-&sig=SORG26sQudO-ZohoKyOHqtzAUo0.

Folia, V., Uddén, J., de Vries, M. H., Forkstam, C., & Petersson, K. M. (2010). Artificial language learning in adults and children. *Language Learning*, 60, 188–220.

Gentner, D., & Bowerman, M. (2009). Why some spatial semantic categories are harder to learn than others: The typological prevalence hypothesis. In J. Guo et al., (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin*, 465–480. Hillsdale, NJ: Erlbaum.

Hannigan, S. L., & Reinitz, M. T. (2001). A demonstration and comparison of two types of inference-based memory errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 931–940. http://dx.doi.org/10.1037//0278-7393.27.4.931.

Jakobson, R. (1971). *Studies on child language and aphasia*. The Hague/Paris: Mouton.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3–28. http://dx.doi.org/10.1037/0033- 2909.114.1.3.

Koring, L., & De Mulder, H. (2014). Understanding different sources of information: the acquisition of evidentiality. *Journal of Child Language*, 947–968. https://doi.org/10.1017/S030500091400052X.

Lesage, Claire & Ramlakhan, Nalini & Toivonen, Ida & Wildman, Chris. (2015). The reliability of testimony and perception: connecting epistemology and linguistic evidentiality. *Proceedings from the 37th Annual Meeting of the Cognitive Science Society*.

Matsui, T. & Fitneva, S. A. (2009). Knowing how we know: evidentiality and cognitive development. In S. A. Fitneva & T. Matsui (eds), *Evidentiality: a window into language and cognitive development* (New Directions for Child and Adolescent Development 125) (pp. 1–11). San Francisco: Jossey-Bass.

McCready, E. (2015). *Reliability in pragmatics*. Oxford: Oxford University Press.

Merkx, M., Rastle, K., & Davis, M. H. (2011). The acquisition of morphological knowledge investigated through artificial language learning. *Quarterly Journal of Experimental Psychology*, 64(6), 1200–1220. https://doi.org/10.1080/17470218.2010.538211.

Newport, E., & Aslin, R. (2004). Learning at a distance i. statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48(2), 127–162.

Ozturk, O., & Papafragou, A. (2016). The Acquisition of Evidentiality and Source Monitoring. *Language Learning and Development*, *12*(2), 199–230. https://doi.org/10.1080/15475441.2015.1024834.

Papafragou, A., Li, P., Choi, Y., & Han, C. (2007). Evidentiality in language and cognition. *Cognition*, 103(2), 253–299. https://doi.org/10.1016/j.cognition.2006.04.001.

Pillow, B. (1989). Early understanding of perception as a source of knowledge. Journal of *Experimental Child Psychology*, *47*, 116–129.

Pinker, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Pratt, C., & Bryant, P. (1990). Young Children Understand That Looking Leads to Knowing (So Long as They Are Looking into a Single Barrel). *Child Development*, *61*(4), 973-982. doi:10.2307/1130869. Rosch, E. H. (1972). Universals in color naming and memory. *Journal of Experimental Psychology*, 93(1), 10–20.

Rosch, E. H. (1972). Universals in color naming and memory. *Journal of experimental psychology*, *93*(1):10–20.

Sperber, D., Clement, F., Heintz, C., Mascaro, O., Mercier, H. Origgi G. & Wilson, D. (2010). Epistemic vigilance, *Mind & Language*, 25(4), 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x.

Strickland, B., & Keil, F. (2011). Event completion: Event based inferences distort memory in a matter of seconds.

Cognition, 121, 409–415. http://dx.doi.org/10.1016/j.cognition.2011.04.007.

Tabullo, Á., Arismendi, M., Wainselboim, A., Primero, G., Vernis, S., Segura, E., & Yorio, A. (2012). On the learnability of frequent and infrequent word orders: An artificial language learning study. *Quarterly Journal of Experimental Psychology*,65(9),1848–1863. https://doi.org/10.1080/17470218.2012.677848.

Ünal, E., Pinto, A., Bunger, A., & Papafragou, A. (2016). Monitoring sources of event memories: A cross-linguistic investigation. *Journal of Memory and Language*, *87*, 157–176. https://doi.org/10.1016/j.jml.2015.10.009

Ünal, E., & Papafragou, A. (2018). Relations between language and cognition: Evidentiality and sources of knowledge. Special issue of *Topics in Cognitive Science* on Lexical Learning, 1-21. https://doi.org/10.1111/tops.12355.

Wiemer, B. (2018). Evidentials and epistemic modality. In A. Aikhenvald (Ed.), *The Oxford Handbook of Evidentiality*. Oxford: Oxford University Press.

Willett, T. (1988). A cross-linguistic survey of the grammaticization of evidentiality. *Studies in language*, *12*(1):51–97.

Wonnacott, E., Brown, H., & Nation, K. (2017). Skewing the evidence: The effect of input structure on child and adult learning of lexically based patterns in an artificial language. *Journal of Memory and Language*, 95, 36–48. https://doi.org/10.1016/j.jml.2017.01.005.