The Problem – This paper develops a fine-grained semantics of the Turkish indirect evidential marker that predicts unexpected discrepancies in its distribution. Turkish, as well as a wide swath of languages spoken in the Balkans, the Caucasus, and Central Asia, exhibits a variety of evidentiality in which information acquired “indirectly” licenses an evidential marker – in the case of Turkish, the vowel-harmonic suffix -mİş (Slobin & Aksu 1982, Johanson & Utas 2000). In descriptive and theoretical accounts alike, “indirect evidence” is used as a cover term for information acquired via hearsay (reportative evidence) or inference (Izvorski 1997, Coşkun 2010, Şener 2011). The central claim of this paper is that the notion of “indirect evidence” must be honed and defined in order to adequately account for the full range of data in Turkish.

The preference for -mİş in (1a) is expected, as direct observation of a birth that took place before one’s own is clearly impossible. Yet -mİş is dispreferred in (1b). While one might explain away this infelicity by positing an exemption for historical facts, such an exemption fails to hold for lesser-known historical information, as in (1c), which favors the indirect evidential marker.

(1) Context: During a conversation with a taxi driver, a speaker might utter the following. (Turkish speakers find these judgments canonical, for most readily imagined contexts.)

(a) Dede-m Türkiye- ’de doğ- {#du / muş}.
   grandfather-1SG.Poss Turkey-LOC be.born-{#PAST / EVID}
   ‘My grandfather was born in Turkey.’
(b) Atatürk Selanik- ’te doğ- {#du / #muş}.
   Atatürk Thessaloniki-LOC be.born-{PAST / #EVID}
   ‘Ataturk (Turkey’s first president) was born in Thessaloniki.’
(c) İsmet İnonü İzmir- ’de doğ- {#du / muş}.
   İsmet Inonu Izmir-LOC be.born-{#PAST / EVID}
   ‘Ismet Inonu (Turkey’s second president) was born in Izmir.’

The contrasts in (1) reveal that the use of -mİş is sensitive to subtle distinctions within the category of “indirect evidence.” Importantly, -mİş does not merely convey doubt on the part of the speaker: a speaker can utter (1a) even when perfectly confident about her grandfather’s place of birth. I claim that the generalization behind (1), among other similarly delicate contrasts, is as follows.

(2) The use of -mİş signals that a speaker’s evidence for the proposition expressed is AT BEST THE SECOND-BEST POSSIBLE given the context and general knowledge about the world.

The Analysis – Izvorski’s (1997) seminal work cast the indirect evidential in Bulgarian and Turkish as an epistemic modal, an approach that captures formal commonalities between evidentials and the present perfect. Yet this approach and others that have built upon it (Faller 2002, Matthewson et al. 2007, Şener 2011) provide no unified account of reportative and inferential evidentials and, furthermore, fail to predict the preferences in (1). Other accounts have made headway characterizing the not-at-issue nature of evidentials, their scope interactions, or their function in discourse, but again the concept of “indirect evidence” is in large part left to be defined by the reader (Davis et al. 2007, Murray 2010). To explain the data in (1), I close in on this definition.

I assume a possible worlds semantics, where \( W \) is the set of all possible worlds and \( f_s(w, p) \)
is a constrained epistemic modal base: the function that assigns to every possible world \( w \) the set of propositions that the speaker \( s \) knows in \( w \), and which serve as evidence for \( p \). This starting point differs from those of previous accounts in that it does not bake an “indirect” flavor into the modal base – i.e., \( f_s(w, p) \) contains not just what qualifies as indirect evidence for \( p \), as in Izvorski (1997), or as perceived or reportative evidence for \( p \), as do the separately defined modal bases of Faller (2011). Next, I define the set \( \mathcal{G}(w) \) of all general world knowledge in \( w \), containing propositions like ‘people cannot witness the births of those older than them,’ ‘there generally exist records one has not examined regarding the birthplaces of relatives,’ and ‘historical figures’ birth records have typically been examined in exhaustive detail.’ \( \mathcal{G}(w) \) is then intersected with \( p \). The result is \( \mathcal{I}(w, p) \), intuitively, the set of all inferences that might arise from the utterance of \( p \) in \( w \). If \( p \) is the proposition expressed by ‘my grandfather was born in Turkey,’ then \( \mathcal{I}(w, p) \) contains the inference ‘the speaker did not witness his grandfather’s birth,’ among others.

(3)  
(a) \( \mathcal{G}(w) = \{g_1, g_2, g_3, \ldots, g_n\} \)  
(b) \( \mathcal{I}(w, p) = \{q : u \in \bigcap_{i=1}^{n} (g_i \cap p) \rightarrow u \in q\} \)

To more rigorously define the sort of evidence that licenses -\( mLs \), I introduce \( f_0(w, p) \), the partially ordered set of all possible evidence for \( p \) in \( w \), independent of speaker: e.g., ‘\( s \) directly observed \( p \)’ outranks ‘\( s \) found previously unexamined records documenting \( p \),’ which might outrank ‘\( s \) was told by reliable sources over many years that \( p \).’ Intersecting the set of all possible evidence \( f_0 \) with \( \mathcal{I} \) eliminates all evidence not consistent with world knowledge in light of \( p \).

(4)  
(a) \( f_0(w, p) = \{q_{best}, q_{best-1}, q_{best-2}, \ldots\}, \quad q_{best} \leq q_{best-1} \leq q_{best-2}, \ldots \)  
(b) \( f_c(w, p) = f_0(w, p) \cap \mathcal{I}(w, p) \)

For the case of a grandparent’s birth, for example, the set of world-compatible evidence \( f_c \) lacks ‘\( s \) directly observed \( p \),’ because it is inconsistent with the general world knowledge that this is impossible. Although \( f_c \) has shed some propositions, those that remain do retain the partial ordering imposed on \( f_0 \). The indirect evidential -\( mLs \), then, contributes the following denotation.

(5)  
\[
[-\text{\( mLs \)}]^{s,w} = \lambda p. [\forall r \in f_s(w, p) \exists q \in f_c(w, p) \text{ s.t. } q \leq r]
\]

This denotation formalizes the generalization that when a proposition \( p \) is marked with -\( mLs \), the evidence possessed by a speaker \( s \) in support of \( p \) is worse than the best evidence possible given general knowledge about the world. It unifies inferential and reportative evidence as special cases of ‘at best second-best’ evidence without stipulating an ambiguity between them into the semantics, and also predicts the observed optionality of speaker doubt in \( p \): depending on what propositions of \( f_0 \) have been ruled out in \( f_c \), one’s evidence for \( p \) might at best be entirely reliable, or quite unreliable. In the paper, I show that this account explains the use of -\( mLs \) in mirative contexts, where -\( mLs \) marks information counter to a speaker’s expectations despite the availability of first-hand evidence, and implement (5) within a commitment-based discourse model (Gunlogson 2001, Farkas & Bruce 2010) to derive the behavior of -\( mLs \) in questions: so-called “interrogative flip.”

**Selected References:**  