Insights into the time course of evidentiality processing in Turkish heritage speakers using a self-paced reading task

Tokaç-Scheffer, Suzan D.; Arslan, Seçkin; Nickels, Lyndsey

Published in:
Frontiers in Communication

DOI:
10.3389/fcomm.2023.1070510

Publication date:
2023

frontiers | Frontiers in Communication

TYPE Original Research
PUBLISHED 22 September 2023
DOI 10.3389/fcomm.2023.1070510

OPEN ACCESS

EDITED BY
Fatih Bayram,
UiT The Arctic University of Norway, Norway

REVIEWED BY
Anna Mikhaylova,
The University of Queensland, Australia
Ayhan Aksu-Koç,
Boğaziçi University, Türkiye

*CORRESPONDENCE
Suzan D. Tokaç-Scheffer
✉ sdtokac@gmail.com

RECEIVED 14 October 2022
ACCEPTED 26 July 2023
PUBLISHED 22 September 2023

CITATION
Tokaç-Scheffer SD, Arslan S and Nickels L (2023) Insights into the time course of evidentiality processing in Turkish heritage speakers using a self-paced reading task. *Front. Commun.* 8:1070510. doi: 10.3389/fcomm.2023.1070510

Insights into the time course of evidentiality processing in Turkish heritage speakers using a self-paced reading task

Suzan D. Tokaç-Scheffer[1,2,3]*, Seçkin Arslan[4] and Lyndsey Nickels[3]

[1]International Doctorate for Experimental Approaches to Language and Brain (IDEALAB), University of Groningen, Netherlands/University of Newcastle, United Kingdom/University of Potsdam, Germany and Macquarie University, Sydney, NSW, Australia, [2]Center for Language and Cognition Groningen (CLCG), University of Groningen, Groningen, Netherlands, [3]School of Psychological Sciences, Macquarie University, Sydney, NSW, Australia, [4]Université Côte d'Azur & CNRS, Nice, France

Introduction: Studies with heritage language speakers (HLS) have often used offline measurements, investigating the post-interpretive effects which emerge after processing has been completed. Relatively few studies have investigated heritage language processing using time-sensitive methods that allow the collection of evidence regarding real-time language processing rather than post-interpretive judgments. Using a self-paced-reading paradigm, we aimed to expand our understanding of HLS language processing by investigating evidentiality-the linguistic marking of information source, which is grammatically expressed in Turkish, but not in English.

Method: Participants were 54 bilingual speakers of Turkish and English: 24 HLS (English onset: 0-5 yrs) and 30 emigrant Turkish speakers (ES) who grew up in Turkey before emigrating to Australia (English onset = 6-17 yrs). Participants read sentences with evidential-marked verb forms that either matched or mismatched to the information source context. Word-by-word reading times and end-of-sentence acceptability judgment speed and accuracy were measured.

Results: The results showed that although the HLS' responses were slower and less accurate than the ES in both reading times and end-of-sentence acceptability judgments, they showed similarities in online processing patterns. Both groups were faster at reading the mismatching sentences compared to the matching sentences; however, this pattern emerged during the time course of reading first for the indirect condition for the ES, and only later for the direct condition and for the HLS for both evidential conditions. Only HLS read faster in the target region with the direct evidential that is shown to be acquired earlier in childhood, than they did for the indirect evidential which is mastered later. In contrast, the end-of-sentence judgment data showed that while the ES group responded faster to matching direct sentences than matching indirect, this effect was missing for the HLS. Nevertheless, there were similar patterns for accuracy across evidential conditions: both groups were more accurate with the direct evidential.

Discussion: Overall, the use of the self-paced-reading paradigm allowed insights into HLS' evidentiality processing above and beyond their generally slower and less accurate processing compared to the reference group. This study provides further evidence for differences in the patterns observed using online vs. post interpretive measures in HLS, reinforcing the importance of combining these methodologies for further understanding of HLS competence and performance.

KEYWORDS

heritage language speakers, evidentiality, Turkish, sentence comprehension, Turkish-English bilinguals, self-paced reading (SPR)

Introduction

Most studies conducted with heritage language speakers report results from offline tasks including paradigms tapping sentence comprehension (sentence-picture matching, Montrul et al., 2008), oral production (structured elicitation, Bayram et al., 2017; picture naming, Hulsen et al., 2002; storey-telling (retelling), Montrul, 2002, 2004; Polinsky and Kagan, 2007; Polinsky, 2008; Montrul and Sánchez-Walker, 2013), written production (Montrul, 2002; Montrul et al., 2008), sentence judgement tasks (aural acceptability judgement task, Fuchs et al., 2015; sentence conjunction judgement task, Montrul, 2009; written acceptability judgement task, Montrul and Bowles, 2009; context/sentence-matching task, Rothman, 2007). These tasks measure the competence of heritage speakers in terms of whether they are aware of the grammatical rules in the language. In contrast, online methods (e.g., self-paced reading, eye-tracking, EEG) allow the measurement of real-time processing, which is not possible to investigate with discreet per-item responses (Felser et al., 2009; Clackson et al., 2011; Felser and Cunnings, 2012; Lago et al., 2018). According to Keating and Jagerski, the real-time component of online methods helps "tap participants' implicit knowledge of language" (2015, p. 2) since they measure processing and knowledge as it happens rather than giving speakers a chance to evaluate, deduct and make a decision based on their learned knowledge. In bilingualism research, online measures have been shown to have an advantage for disentangling sentence processing mechanisms (Keating and Jegerski, 2015).

Here we focus on self-paced reading. During a self-paced reading task, language users are presented with a single segment consisting of a word or a phrase, which disappears on command (usually a button press on the keyboard) allowing a new segment to appear. The time passing between each command/button press gives an indication of the processing load or cost of the segment on the language user. This includes enabling identification of the point in the sentence that participants may encounter processing difficulties and how long these difficulties persist. These indicators may reflect increased cognitive load, or mental effort required to process the sentence (Just and Carpenter, 1980). Consequently, as an online method, self-paced reading enables recording of segment-by-segment reading times when a reader is presented with each word or group of words in a sentence, and how they process and react to them (Chen et al., 2005; Keating and Jegerski, 2015).

There is extensive research using self-paced reading tasks with monolingual speakers (e.g., Carminati, 2002; Filiaci, 2011; Xu et al., 2018; Lee and Fraundorf, 2022) and second language learners (for review see Nicklin and Plonsky, 2020) examining the processing of various linguistic phenomena. Self-paced reading experiments with heritage language speakers are, on the other hand, relatively scarce. However, those studies that have been carried out found that although heritage language speakers were slower and/or less accurate than reference groups (monolinguals and/or late bilinguals and second language speakers), their processing patterns had qualitative similarities and showed differences to those measured with offline tasks (Keating et al., 2016; Jegerski, 2018a,b; Mikhaylova, 2018; Di Pisa et al., 2022). For example, Keating et al. (2016) investigated monolingual and heritage language speakers' antecedent choice for ambiguity resolution between null vs. overt pronouns in Spanish. They used an online self-paced reading task with sentences like *Cuando la diva visitó a la directora, Øella ofreció cantar un aria en italiano* (translation: When the diva$_F$ visited the director$_F$, ø/she offered to sing an aria in Italian). In Spanish it is more accurate to assign the null pronoun (ø) to the word in the subject position (i.e., *diva*); whereas the overt pronoun *she* is preferentially attributed to the *director* which is the object of the preceding clause (Keating et al., 2016; Supplementary material). An earlier offline study found heritage language speakers showed no such processing preferences for null vs. overt pronouns (Keating et al., 2011). In contrast, in Keating et al. (2016) self-paced reading study, heritage language speakers did show a preference for attributing null pronouns to antecedents in the subject position that is shown by monolingual native speakers. However, they did not show a preference for overt pronouns. Nevertheless, the key point here is that they showed dissimilar processing of null vs. overt pronouns during this online task. In contrast, in the responses to end-of-sentence comprehension questions (Keating et al., 2016), heritage language speakers showed no preference between null and overt pronouns, just as they had not in the earlier offline study (Keating et al., 2011). Although the heritage language speakers' processing was not completely parallel to that of monolinguals, this study clearly illustrates that tasks tapping online and offline processing can provide different insights and thereby underlines the importance of contrasting experimental methods.

Jegerski (2018b) also reports a study that supports the utility of self-paced reading tasks, for identifying which linguistic phenomena are challenging for all speakers of that language and which are only challenging for heritage language speakers. They tested heritage language speakers' *Differential Object Marking* (DOM) in Spanish compared to a group of late Spanish-English bilinguals using an online self-paced reading task interspersed with an offline end-of-sentence acceptability judgement task. In the offline, end-of-sentence acceptability judgements, heritage language speakers were less accurate and slower than the late bilinguals and did not show any differences between the conditions. This result was similar to that of Montrul and Bowles (2009) finding, also from an offline acceptability judgement task, which showed that Spanish heritage language speakers could not distinguish sentences that were ungrammatical for DOM from grammatical sentences. However, Jegerski (2018b) found that during online self-paced reading, both groups showed slower RTs for the ungrammatical DOM of the inanimate direct object, but no sensitivity to the ungrammatical omission of DOM for animate direct objects (the only condition where a direct object can be marked with "*a*" in Spanish) (Jegerski, 2018b). Indeed, they report similar online sensitivity in both heritage language speakers and late bilinguals that highlighted the fact that the variability in DOM processing could not be attributed to incomplete attainment of DOM markers. This study particularly highlights the fact that differences between heritage language speakers and a reference group of bilingual speakers were no longer apparent during online processing. Consequently, these studies demonstrate how the self-paced reading task can provide additional information to

facilitate the unravelling of heritage language speakers' processing. It is important to note, that, in these studies, a phrase-by-phrase presentation was employed. However, it has been suggested that a word-by-word presentation would allow a more refined analysis in terms of time-course of processing as it does not collapse across several words, and therefore gives smaller time windows (Keating et al., 2016).

The studies of self-paced reading cited above found heritage language speakers to show some qualitative similarity to reference groups and/or provided more insights into heritage language speakers' processing. Consequently, we were interested in whether the same would hold for heritage language speakers' processing of evidentiality in Turkish, especially given that in the primarily offline measures used in earlier studies, heritage language speakers showed slower reaction times and lower accuracy compared to monolingual (Arslan et al., 2017) and bilingual (Schmid and Karayayla, 2019; Tokaç-Scheffer et al., to appear) reference groups.

Evidentiality is the specification of how a speaker received the information in their utterance (Aikhenvald, 2004). Information can be received through different sources (i.e., visually, aurally, from a third person, etc.) and the evidential markers available in a particular language are used to specify this source. In some languages, such as Indo-European languages like English, evidential meanings are conveyed by means of lexical elements, such as *I saw* for direct visual evidence, or *I have been told or I assume* for inferred or reported knowledge (e.g., I saw John ate the apple yesterday vs. I was told/I assume that John ate the apple yesterday). These forms are optional, yet they indicate the amount of evidence for a speaker's assertion (De Haan, 1999). However, in a quarter of the world's languages, evidentiality is a grammatical unit and specification of the evidence type is obligatory in one's utterance (Aikhenvald, 2004). Turkish, the heritage language under investigation in this paper, is one such language and, in Turkish, it is obligatory to use evidential markers when referring to the past. Evidentiality in Turkish is marked as a verb inflexion that indicates the source of a past event: the evidential marker specifies whether the speaker witnessed and/or personally carried out the action firsthand or received that information non-firsthand, as in hearsay or inference (Aikhenvald, 2004; De Haan, 2005). In the case of firsthand information, the direct evidential marker -DI[1] is used. For example, in "*Bahçivan çiçekleri sula**dı**,*" (I know/saw that) the gardener watered the plants. Information that is non-firsthand is marked with the indirect evidential marker -mIş: "*Bahçivan çiçekleri sula**mış**,*" (I infer it from the wet plants or someone else told me that) the gardener watered the plants.

Acquisition studies conducted with monolingual Turkish children have shown that children start producing evidentiality in their utterances very early (e.g., 1,5 years in Aksu-Koç et al., 2009); and that the acquisition of the indirect evidential marker follows the direct evidential marker (Aksu-Koç, 1988; Aksu-Koç et al., 2009). It has also been shown that it may take up to the age of seven for children to fully master the distinctions between the evidentiality markers (Öztürk and Papafragou, 2008; Ünal and Papafragou, 2016). This makes evidentiality a "late-mastered" language component and late-mastered linguistic items have been shown to be challenging for heritage language speakers (e.g., Montrul, 2002; Montrul et al., 2008; Polinsky, 2008).

Tokaç-Scheffer et al. (to appear) conducted an offline study that is of particular relevance to the study reported here. In this offline study, we compared the evidentiality processing of two groups of bilingual Turkish speakers, namely heritage language speakers of Turkish and bilingual Turkish-English speakers who had migrated from Turkey using an (offline) auditory sentence verification task (see also Arslan et al., 2017). The task was to listen to sentences presented in the heritage language and to respond as fast as possible whenever a word was detected that made the sentence unacceptable and/or semantically inappropriate. Tokaç-Scheffer et al. (to appear) found that the heritage language speakers were slower and less accurate in detecting these unacceptable sentences compared to emigrant speakers. Further analyses revealed that neither of the groups showed better processing for either evidential condition in their reaction times. However, emigrant speakers performed better in the indirect evidential condition. They rejected the sentences with the violation of a direct information source by the use of indirect evidential marker more accurately compared to when the indirect information source was followed by a mismatching use of direct evidential. This pattern was not shown by the heritage language speakers.

Karaca (2018), in an unpublished Master's thesis, also presented a relevant study using a self-paced listening task to compare three groups of Turkish speakers: a bilingual group of heritage language speakers of Turkish born in, or moved to, Canada before the age of 5, a bilingual group of first-generation migrants from Turkey to Canada and, a monolingual group of Turkish speakers residing in Turkey. Karaca administered a self-paced listening task, where participants listened to dialogues including congruent and incongruent evidentiality sentences and answered yes/no comprehension questions (for a quarter of the stimuli). While listening to the critical segment that included the verb marked with indirect evidential, monolinguals showed differences between incongruent and congruent conditions (and sustained this effect in the following segment): they were slower in the incongruent conditions for both evidentials. In contrast, the heritage language speakers showed no processing differences at any point. However, the migrant speakers mirrored the effect that the monolinguals showed for the indirect evidential sentences but only on the third segment—a delayed effect (Karaca, 2018), suggesting reduced processing speed for bilingual individuals who have experienced attrition in their first acquired language. In the third segment monolingual speakers also showed a difference between congruent and incongruent sentences for the direct evidential; such an effect was missing both for heritage language speakers and migrant speakers. However, Karaca argued that without the presentation of a fourth segment, we cannot exclude the possibility of heritage language speakers and/or migrant speakers showing a late effect. Karaca did not present the results for the comprehension questions. In sum, during this online study monolingual speakers showed processing differences between

1 The capitalization of the letters indicates that the sounds represented follow harmonization rules in Turkish, i.e., vowel harmony and consonant assimilation rules. As an agglutinative language, in Turkish sounds may be modified when appended through suffixation.

congruent and incongruent evidential conditions and rejected particularly quickly the incongruent indirect evidential marker use early in processing. The migrant speakers showed some similarities to the monolingual speakers while heritage language speakers showed no evidence of evidentiality processing preferences.

In the present study, we aimed to further explore the online and temporal processing of evidentiality of heritage language speakers relative to a reference group of migrant bilingual speakers by using a self-paced reading task with longer stimuli presented with a word-by-word moving window paradigm.

Specifically, we addressed the following research question:

(1) Does the online processing of Turkish heritage language speakers and Turkish speakers who are also late second-language speakers of English (emigrant Turkish speakers) differ during the time course of reading sentences with evidentiality marking?

Previous studies using self-paced reading have shown heritage language speakers can pattern with late bilinguals in their sensitivity during sentence reading, even when they differ in offline end-of-sentence judgement responses (Jegerski, 2018b). Therefore, although, in other studies with heritage language speakers of Turkish (e.g., Arslan et al., 2015, 2017; Karayayla and Schmid, 2019; Tokaç-Scheffer et al., to appear) heritage language speakers showed different patterns of evidentiality processing to bilingual emigrant Turkish speakers, we hypothesised that using an online task may reveal qualitative similarities between these speakers and thereby gain more information on the time course of evidentiality processing by heritage language speakers. We would nevertheless expect heritage language speakers to process the sentences more slowly than the emigrant Turkish speakers, given the differences in their Turkish exposure.

(2) Does the online processing of evidentiality by Turkish heritage language speakers and emigrant speakers of Turkish differ depending on the evidentiality distinction -direct vs. indirect? Do the two groups differ in this regard?

Based on previous research showing that the direct evidentiality condition is mastered earlier in children's acquisition and that the indirect evidential has more semantic connotations, and is therefore cognitively more complex, we expected both groups to show better processing of direct evidential which would be manifested in shorter reading times for the verbs marked with the direct evidential compared to the indirect evidential (Aksu-Koç, 1988; Öztürk and Papafragou, 2008; Aksu-Koç et al., 2009; Ünal and Papafragou, 2016).

(3) Are there differences between online and offline processing of evidentiality for Turkish heritage language speakers and emigrant speakers of Turkish? Do the two groups differ in this regard?

Similar to the previous studies described above (e.g., Keating et al., 2016), we expect online and offline tasks to show differences in the processing of evidentiality, with more similarities between heritage language speakers and emigrant speakers of Turkish during online processing.

Materials and methods

Participants

A total of 53 bilingual speakers of Turkish and English (M_{AGE} = 34.4 years; SD = 9.1; 28 Female; four left-handed) all of whom resided in Sydney, Australia participated in this study. All participants performed this self-paced reading task first. The participants were recruited via student organisations, Turkish cultural and language centres, the Turkish Consulate Sydney, schools that deliver Turkish education, and advertisements posted on social media and in neighbourhoods with large Turkish communities. Inclusion criteria were that participants were 20–54 years of age, had to have started acquiring Turkish from birth, used English actively in daily life, and had no previous psychological, neurological, or communication disorders. The participants were given a bilingual language background questionnaire constructed based on the Language and Social Background Questionnaire (LSBQ; Anderson et al., 2018) with adaptations tailored to our research questions. These adaptations included editing sub-sections such as the community language use behaviour section (now language use behaviour) and amending the scaling method for life stages to use average percentages of exposure/use. In addition, reading was expanded into a whole section (reading habits) to collect more detailed input on participants' reading behaviours given that we planned to use a reading task. We excluded questions from the language use section that were more detailed than needed for our study (i.e., language use for social, religious, extracurricular activities; shopping/restaurant/other commercial services; health care services/government/public offices/banks) and instead included only languages preferred at home, work, social life and in general). The final questionnaire consisted of four sections: social background (education, occupation, parents' language history, countries they had lived, etc.), language background, language use behaviour, and reading habits. Table 1 provides a summary of the outcomes of the questionnaire.

The participants were divided in two groups according to their heritage language situation and characteristics: heritage language speakers of Turkish and emigrant speakers.

Heritage language speakers (n = 23)

Twenty-three heritage language speakers of Turkish, speaking both Turkish and English from early childhood, were recruited. Their age ranged between 20 and 45 years. While the age onset of Turkish was always from birth in this group, the onset for English ranged between birth and 5 years of age, and these participants were either born in Australia or migrated there at a very young age (i.e. at or before the age of 5). Although they spoke Turkish as their "home" language and it was their first learned language, most of them learned to read and write in Turkish after they had acquired these skills in English. They began acquiring English in kindergarten in Australia and received all their education in English.[2] They were

2 Nevertheless, many children from Turkish backgrounds have access to some Turkish schooling in Australia. Most of the participants reported here either received Turkish instruction at school or attended "Saturday Schools," at which they carried out activities in Turkish once a week for a couple of

TABLE 1 Summary of critical outcomes from bilingualism language background questionnaire for both groups; heritage language speakers and emigrant speakers.

	HLS		ES			
	Mean	SD	Mean	SD	*Welch t-test* (*p*-value)	95% CIs
Age (years)	30.17	8.73	37.70	8.16	3.20 (0.002)*	[2.79, 12.26]
Education (years)	16.17	1.75	16.67	2.90	0.76 (0.447)	[−0.80, 1.79]
Years of residence in Australia	28.04	9.31	10.46	8.59	−7.04 (<0.001)*	[−22.61, −12.56]
Age of bilingualism Onset	1.00	1.98	11.53	1.96	19.29 (<0.001)*	[9.43, 11.63]
Turkish proficiency (self-rated)	7.91	1.52	9.90	0.28	6.22 (<0.001)*	[1.33, 2.66]
English proficiency (self-rated)	9.88	0.41	8.24	1.60	−5.39 (<0.001)*	[−2.26, −1.02]
Turkish exposure and use (current %)	30.02	18.20	33.00	19.00	0.54 (0.590)	[−7.55, 13.11]
English exposure and use (current %)	69.78	18.18	67.00	19.00	−0.54 (0.590)	[−13.11, 7.55]
Turkish material—audio and video (hr/day)	1.72	1.26	1.75	2.07	0.07 (0.947)	[−0.89, 0.96]
English material—audio and video (hr/day)	4.26	5.87	2.92	1.66	−1.06 (0.296)	[−1.24, 3.74]
Turkish material—written (current %)	19.70	17.24	33.33	22.45	2.50 (0.016)*	[2.69, 24.58]
English material—written (current %)	80.30	17.24	66.66	22.45	−2.50 (0.016)*	[−24.58, −2.69]
Turkish material—written (hr/week)	3.00	2.62	10.98	9.09	4.57 (<0.001)*	[4.43, 11.53]
English material—written (hr/week)	22.73	19.07	25.30	19.40	0.48 (0.632)	[−8.15, 13.27]

SD, Standard Deviation.
Values of significant effects ($p < 0.05$) are printed in bold and asterisked.

exposed to Turkish at home and within the Turkish community they lived in and in social environments, but their use of English became dominant over time.

Emigrant speakers (n = 30)

The emigrant speakers (age range 23–54 years) comprised participants who were raised in Turkey during their childhood and emigrated to Australia due to professional or educational circumstances at or after adolescence. They were non-native speakers of English who had started learning this language between the ages of 6 and 17 (mean = 11.50; SD = 2.0), at school in Turkey as a second/foreign language. The duration of their residence in Australia was from 1 month up to 31 years (see Table 1 for details). As skilled migrants, most of them had an upper intermediate level of English and reported using English more than Turkish since they moved to Australia.

Stimuli

The stimuli comprised 134 sentences. There were 104 target evidentiality sentences (26 in each of the four conditions) and 30 filler sentences. Twenty-six unique verbs referring to different actions were selected and each verb was inserted into a sentence frame which was then adapted for each of the four conditions of the evidentiality manipulation as described below.

Evidentiality sentences were 12 words long. The first three words constituted the contextual support which included a statement of the information source, specifically whether it was firsthand or nonfirsthand. If the information source was *firsthand*, it indicated the event was witnessed by the speaker her/himself (using *ben gördüğüme eminim*; "I am sure that I saw"). A *nonfirsthand* information source, on the other hand, specified the event was witnessed by others and that they transferred this knowledge, meaning that the speaker had heard about the event from others (*başkaları gördüğünü söylüyor*; "others say they saw it"). See Table 2 for example sentences.

The statement of the information source (contextual clause) was followed by the critical clause, which included the target verb region (R-TV)—the verb inflected with the evidentiality marker. The evidential marker on the verb was either *direct* or *indirect* (condition) and either matched or did not match the information source (firsthand or nonfirsthand) in the contextual clause. In the *match* sentences, the evidentiality marker appended on the main verb of the critical clause matched the preceding information source: the direct evidential marker (-DI) was used in firsthand information source sentences (*firsthand—direct;* see Table 2) and the indirect evidential marker (-mIş) was used in the nonfirsthand information source conditions (*nonfirsthand—indirect*). In the *mismatch* sentences the evidential marker on the verb did not match with the information source: Following a nonfirsthand information source the direct evidential marker (-DI) was given (*nonfirsthand-direct**) and following a firsthand information source the indirect evidential marker (-mIş) was provided (*firsthand-indirect**).

The last part of the sentence was the padding phrase (last five words) which included the spillover region (divided into two

hours. One participant reported having been home schooled in Turkish by her/his parents, and another participant had received Turkish as a second language lessons during her regular schooling.

TABLE 2 Example of evidentiality sentences used in the self-paced reading task.

	Condition	Contextual support	Critical clause	Padding phrase
Direct	Firsthand-direct	Ben gördüğüme eminim,	Mehmet ceketinin düğmesini kopardı	bu sebepten terziye gitmesi gerekecek.
		I see.$_{\text{DIRECT EVID.1SG}}$ sure.$_{\text{1SG}}$	Mehmet jacket.$_{\text{POSS.GEN}}$ button.$_{\text{POSS.DEF}}$ pull off.$_{\text{DIRECT EVID.3SG}}$	this reason.$_{\text{ABL}}$ tailor.$_{\text{DAT}}$ go.$_{\text{MOD}}$ require.$_{\text{FUTURE}}$
		'I am sure *I saw* Mehmet pull off (witnessed) the button of his jacket, that's why he will need to go to the tailor.'		
	Nonfirsthand-direct*	Başkaları gördüğünü söylüyor,	Mehmet ceketinin düğmesini *kopardı	bu sebepten terziye gitmesi gerekecek.
		Others see.$_{\text{DIRECT EVID.1SG}}$ say.$_{\text{3PL}}$	Mehmet jacket.$_{\text{POSS.GEN}}$ button.$_{\text{POSS.DEF}}$ *pull off.$_{\text{DIRECT EVID.3SG}}$	this reason.$_{\text{ABL}}$ tailor.$_{\text{DAT}}$ go.$_{\text{MOD}}$ require.$_{\text{FUTURE}}$
		"Others say *they saw* Mehmet *pull off (witnessed) the button of his jacket, that's why he will need to go to the tailor."		
Indirect	Nonfirsthand-indirect	Başkaları gördüğünü söylüyor,	Mehmet ceketinin düğmesini koparmış	bu sebepten terziye gitmesi gerekecek.
		Others see.$_{\text{DIRECT EVID.1SG}}$ say.$_{\text{3PL}}$	Mehmet jacket.$_{\text{POSS.GEN}}$ button.$_{\text{POSS.DEF}}$ pull off.$_{\text{DIRECT EVID.3SG}}$	this reason.$_{\text{ABL}}$ tailor.$_{\text{DAT}}$ go.$_{\text{MOD}}$ require.$_{\text{FUTURE}}$
		"Others say *they saw* Mehmet pull off (reportedly) the button of his jacket, that's why he will need to go to the tailor."		
	Firsthand-indirect*	Ben gördüğüme eminim,	Mehmet ceketinin düğmesini *koparmış	bu sebepten terziye gitmesi gerekecek.
		I see.$_{\text{DIRECT EVID.1SG}}$ sure.$_{\text{1SG}}$	Mehmet jacket.$_{\text{POSS.GEN}}$ button.$_{\text{POSS.DEF}}$ *pull off.$_{\text{DIRECT EVID.3SG}}$	this reason.$_{\text{ABL}}$ tailor.$_{\text{DAT}}$ go.$_{\text{MOD}}$ require.$_{\text{FUTURE}}$
		"I am sure *I saw* Mehmet *pull off (reportedly) the button of his jacket, that's why he will need to go to the tailor."		

Examples 1 and 3 are Match conditions; Examples 2 and 4 are Mismatch conditions (indicated with *).

spillover regions R-SO1 and R-SO2, two words in each region) and the final word region (R-FW). This padding phrase (last five words) included extra information regarding the event to enhance the clarity and comprehension of the overall meaning conveyed in the text. This phrase also allowed us to observe potential delays in processing since in self-paced-reading, effects may carry over to next segments. These phrases were presented in the present continuous or simple future tense to avoid any confusion with the time of the event (past).

The 30 filler sentences which also included morphosyntactic mismatches were sentences without specification of an information source. Half of the sentences were ungrammatical, created either by person/number disagreements or semantically incorrect verb choice (for the full list of stimuli see Supplementary material). The number of fillers (30) was determined based on the number of unique verbs (26) selected for the evidentiality sentences. These 26 verbs were then manipulated across four conditions, resulting in a total of 104 experimental sentences. Given the relatively large number of experimental sentences, to prevent fatigue, we decided not to add additional filler sentences. Furthermore, this self-paced reading experiment was only one part of a longer study, which in total took 2 h for participants to complete.

The 26 critical action verbs used to construct the sentence stimuli described above were chosen from a larger set of stimuli, that were normed for surface frequency, cloze probability of the evidential verb (see Tokaç-Scheffer, 2023 for further details).

Procedure

The sentence materials were programmed in a non-cumulative self-paced reading design with end-of-sentence acceptability judgement (Just et al., 1982) using the web platform Ibex Farm (Drummond, 2013). The sentences were presented in black font (96px) on a white background. The stimulus sentence advanced segment-by-segment with each press of the SPACE button in a *moving-window paradigm*. The first segment always contained the contextual clause where the information source was presented (consistently including 3 words; see Table 2) followed by the critical sentence material presented per word *per se*gment. *The uninformative mask* technique was employed to the sentences with the word boundaries shown on the screen. With the first press, the information source for each sentence was presented as a single chunk of three words (e.g. *başkaları gördüğünü söylüyor*; "others say they saw it") at the beginning of the sentence. Following a space bar press, this first segment disappeared from view, and the next word appeared to the right, such that only one segment (comprising a single word) was visible at any one time. At the end of each sentence participants were required to judge whether the sentence was "grammatically coherent." After the answer, the next sentence appeared automatically.

The experiment started with an explanation of the task and what was expected. Participants were shown the first practise item and instructions in Turkish were given with a demonstration of the first practise item "This is a sentence. You will read each word by pressing the space key on the keyboard. After each press, a word will appear and with the next press that word will disappear and the next one will appear. You will see each word and consequently each sentence only once and will not be able to go back. This long line here [pointing to a line that indicates where the information source phrase will appear] gives you [relevant] information and then the rest of the sentence will follow. At the end of each sentence, you will decide if this sentence was grammatically coherent or not by choosing the smiley face emoji for *yes* and sad face emoji for *no* that will appear on the screen which corresponds to the *f* and *j* keys, respectively, on the keyboard." They were not given any instructions regarding the speed of reading or responding (to the judgement questions). Testing started after four example

trials. Each participant was presented with sentences in all four conditions together with the filler items, as described in Table 2. The presentation order of the sentences was randomised for each participant. Participants were given the opportunity to have breaks when needed.

Data pre-processing and analyses

All data pre-processing and analyses were conducted in R studio version 1.2.5 (R Core-Team, 2012). First, sub-datasets were created for the analysis of each region of interest and acceptability question answers. Accordingly, five separate datasets were created: four for the analyses of reading times included the following regions: Target Verb (R-TV; target verb inflected with the evidentiality marker), SpillOver 1 (R-SO1; comprising the individual reading times to the two words following the target verb), SpillOver 2 (R-SO2; comprising the individual reading times to the next two words following the R-SO1), and Final Word (R-FW; the final word of the sentence seen before the presentation of the sentence acceptability judgement). A separate dataset was created for the reaction time analyses for the Sentence Judgements (SJ-RT; the acceptability question after the presentation of each sentence), and for the analyses of the sentence judgement accuracy (SJ-Acc; accuracy for the sentence judgements). Each dataset consisted of 5512 data points initially (26 sentences * 4 conditions * 53 participants). For data cleaning procedures, we followed Nicklin and Plonsky (2020), which presents a comprehensive overview of data pre-processing practises in bilingual studies and adopted those that fitted our population and research objectives. Specifically, we started with a visual inspection of the data and looked at histograms, boxplots, and Q-Q plots (see Supplementary material) to determine appropriate cut-offs to exclude outliers: Trials in the reading time data sets that were faster than 100 ms (e.g., Luce, 1991; Jegerski, 2016; Litcofsky and Van Hell, 2017; Kim et al., 2018) and slower than 8,000 ms (R-TV and R-SO) and 15,000 ms (R-FW) were excluded.[3] Eight trials at R-TV, 12 at R-SO1 (6^{th} word − 8 trials; 7^{th} word − 3 trials), 14 at R-SO2 (8^{th} word − 8 trials; 9^{th} word − 6 trials), and 40 at R-FW regions were excluded. Lower and upper boundaries for SJ reaction times were 100 and 15,000 ms respectively, which resulted in the exclusion of 46 data points. This pre-processing resulted in the exclusion of <1% of the data for each dataset. The accuracy analyses were computed on the same data set that was created for the SJ-RT analysis. We analysed response times for both accurately and inaccurately judged sentences as these responses have been shown to be informative (Jegerski, 2015).

Statistical analyses of the reading/reaction times were performed using mixed-effects models computed with the "lme4" package in R (Bates et al., 2015). We started by constructing a maximal model including random intercepts and slopes for participants and items (Barr et al., 2013) and also used an optimizer in the models analysing the reading time data (optimizer = 'bobyqa; Powell, 2009). As the maximal models failed to converge, we simplified them by removing the random slopes. Each final model included both participant and item as random intercepts. The two-level factorial interaction variables, evidentiality (direct, indirect), grammaticality (match, mismatch), and group (heritage speakers, emigrant speakers) were sumcoded.[4] All reading times were log-transformed in the models to reduce the positive skew. As the participant groups differed significantly in age and verbal working memory, we controlled for these variables by adding them as fixed effects. Since it is known to affect reading speed, region length, that is the number of letters composing each word, was also included as a control variable (Jegerski, 2014). To explore the accuracy differences between groups, similar models to those outlined above were built and the scores were analysed using logit generalised mixed-effects models. *Post hoc* pairwise comparisons to explore the nature of the interactions were conducted using the "emmeans" package (Lenth, 2019) and adjusted using Holm correction for multiple comparisons.

Results

Offline processing: sentence judgement results

Table 3 shows the mixed-effect model output for the response time data and the generalised mixed-effects model output for accuracy to sentence judgement questions.

Sentence judgement accuracy

The generalised mixed effect model for sentence judgement accuracy revealed a three-way interaction between group, grammaticality and evidential (see Figure 1A). *Post-hoc* analyses showed that, in the *direct* condition, both groups were significantly more accurate in their judgements of the match sentences compared to mismatch sentences (HLS: $\beta = 1.22$, $SE = 0.184$, $z = 6.61$, $p < 0.001^*$; ES: $\beta = 1.55$, $SE = 0.183$, $z = 8.51$, $p < 0.001^*$). However, while the same held for the *indirect* condition for heritage language speakers ($\beta = 0.64$, $SE = 0.180$, $z = 3.57$, $p = 0.002^*$) it did not for the emigrant speakers who showed no significant difference between the match and mismatch sentences in this condition ($\beta = 0.23$, $SE = 0.173$, $z = 1.33$, $p = 0.369$). In terms of evidentiality, for match sentences, both groups were significantly more accurate in the *direct* compared to *indirect* condition (HLS $\beta = 0.58$, $SE = 0.186$, $z = 3.12$, $p = 0.007^*$; ES $\beta = 1.01$, $SE = 0.18$, $z = 5.46$, $p < 0.001^*$), but showed no such difference for the mismatch sentences (HLS $\beta < 0.001$, $SE = 0.177$, $z = 0.04$, $p = 0.970$; ES $\beta = -0.32$, $SE = 0.169$, $z = -1.87$, $p = 0.185$).

3 A variety of upper boundaries (1,000, 2,000, 2,500, 5,000 ms) have been mentioned in self-paced reading studies (e.g., Hofmeister, 2011; Vasishth and Drenhaus, 2011; Nicklin and Plonsky, 2020). Given the population studied in this study and their varying language competencies we aimed to not exclude any critical data and selected an upper boundary of 8,000 ms, taking into account the visual inspection of the data.

4 Although we labelled this variable "grammaticality" the mismatching sentences are not ungrammatical *per se*. It is possible in some contexts for these sentences to be plausible. Nevertheless, without the presentation of a full context (as is the case in this experiment) these mismatching sentences are not acceptable, as they present opposite sources of information.

10.3389/fcomm.2023.1070510

TABLE 3 Mixed-effects estimates of accuracy and response times for the sentence judgement questions.

	Sentence judgement questions accuracy				Sentence judgement questions response times			
	ß	SE	*t*	*p*	ß	SE	*z*	*p*
(Intercept)	0.76	0.11	7.00	**<0.001***	7.13	0.05	133.03	**<0.001***
Age	>-0.01	0.01	−0.30	0.764	<0.001	0.01	0.28	0.779
Verbal Working Memory	0.11	0.06	1.90	0.057	0.04	0.03	1.08	0.285
Group	0.19	0.11	1.61	0.108	−0.08	0.06	−1.21	0.232
Evidential	−0.16	0.05	−2.91	**0.004***	0.01	0.01	1.04	0.299
Grammaticality	−0.45	0.05	−8.25	**<0.001***	>-0.01	0.01	−0.35	0.728
Evidential—Grammaticality	0.24	0.05	4.32	**<0.001***	−0.01	0.01	−1.15	0.252
Evidential—Group	−0.01	0.03	−0.41	0.679	0.01	0.01	0.70	0.486
Grammaticality—Group	<0.001	0.03	0.29	0.769	−0.01	0.01	−0.88	0.379
Evidential—Grammaticality—Group	0.09	0.03	2.96	**0.003***	−0.02	0.01	−2.62	**0.009***
Observations	<5468				<5468			
Marginal R^2/Conditional R^2	0.092/0.245				0.011/0.254			
	glmer (Accuracy ~ Evidential * Grammaticality * Group + c. (VerbalWM) + c. (Age) + (1 |ParticipantCode) + (1 |Item), data=dataAR, family = binomial (link = "logit"))				lmer (log (ResponseRT) ~ Evidential * Grammaticality * Group + c. (RegionLength) + c. (VerbalWM) + c. (Age) + (1 |Participant) + (1 |Item), data = dataSJ, REML = FALSE, control = lmerControl (optimizer = "bobyqa"))			

Values of significant effects ($p < 0.05$) are printed in bold and asterisked.

Sentence judgement response time

The model output for the response times to the sentence judgement questions showed a three-way interaction between group, grammaticality and evidential (see Figure 1B). The emmeans analyses showed that in the *indirect* condition emigrant speakers were marginally, but not significantly, faster in their responses to the mismatch sentences ($\beta = 0.09$, $SE = 0.036$, $t = 2.60$, $p = 0.067$) compared to the match sentences. They were marginally faster when firsthand information was (incorrectly) followed by the indirect evidential marker compared to when nonfirsthand information was (correctly) followed by the indirect evidential marker. In the match sentences, emigrant speakers' responses were significantly *faster* in the direct condition, that is *firsthand-direct* sentences, compared to the indirect, *nonfirsthand-indirect* sentences ($\beta = -0.10$, $SE = 0.036$, $t = -2.90$, $p = 0.031^*$). The heritage language speakers showed no significant differences in sentence judgement response times across the sentence types.

Online processing: reading time results

The reading time analysis of each critical region revealed, as expected, slower reading times for heritage language speakers compared to the emigrant speakers at every time point (HLS overall mean RT = 907.1, SD = 834.0; ES overall mean RT = 683.6, SD = 690.9). Figure 2 shows by-region reading time averages.

Outputs of the mixed-effects models, computed at each region, are presented in Table 4. At all four regions, as expected, there was a significant effect of the Speaker Group on reading times: heritage language speakers were slower in their reading times at the target verb compared to emigrant speakers. We will discuss the remaining results for each region in turn.

Target verb

Reading times at the target verb showed a significant two-way interaction between evidential and group (see Figure 3A). *Post hoc* analysis indicated that the source of this interaction was that, while the emigrant speaker group showed no significant difference between the direct and indirect evidentiality conditions ($\beta < 0.01$, $SE = 0.018$, $t = 0.23$, $p = 0.816$), the heritage language speaker group did show a significant difference ($\beta = -0.05$, $SE = 0.020$, $t = -2.60$, $p = 0.040^*$). Irrespective of grammaticality, the heritage language speakers read verbs marked with *direct* evidentiality faster than those with *indirect* evidentiality (see Supplementary material for the full pairwise comparisons of the emmeans analyses).

Spillover region 1

In the first spillover region, comprising the reading times of the first two words following the evidentiality-marked verb, there was a significant three-way interaction between group, grammaticality and evidential (see Figure 3B). This reflected that in the *indirect* condition (when the verb is marked with the indirect evidential) the emigrant speakers were faster when the evidentiality marker did not match with the information source than when it did ($\beta = 0.07$, $SE = 0.016$ $t = 4.47$, $p < 0.001^*$), but this was not the case in the *direct* condition ($\beta = 0.01$, $SE = 0.016$, $t = 0.60$, $p = 1.000$). In addition, emigrant speakers were faster in the *indirect* condition

10.3389/fcomm.2023.1070510

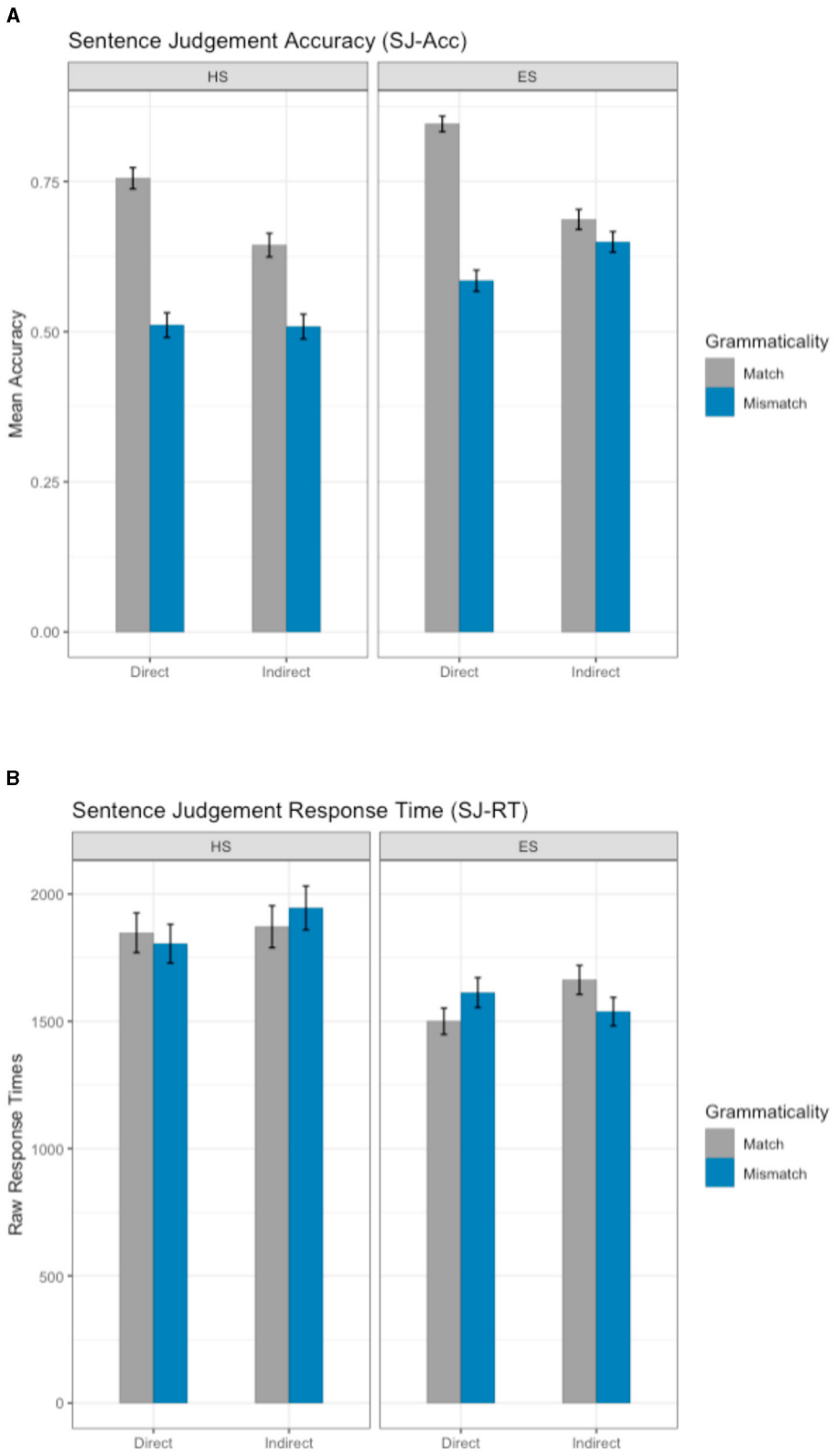

FIGURE 1
Sentence judgement accuracy **(A)** and response times **(B)**. Error bars represent the standard error of the mean. The means and standart deviations are provided in Table A1.

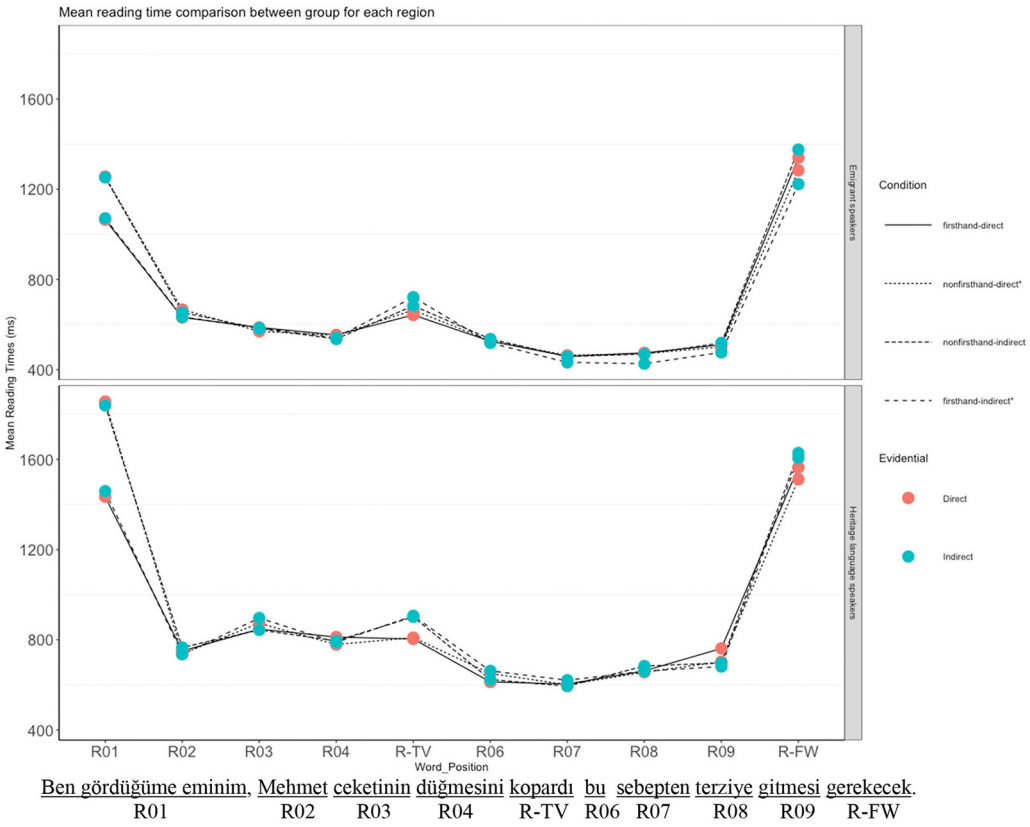

FIGURE 2
Word reading time comparisons across the sentence, for each evidential condition for emigrant speakers and heritage language speakers.

compared to the *direct* condition for mismatch sentences ($\beta = 0.06$, $SE = 0.016$, $t = 3.70$, $p < 0.001^*$) but in the match sentences there was no significant condition difference ($\beta = -<0.01$, $SE = 0.016$, $t = -0.17$, $p = 1.000$). However, for the heritage language speakers, there were no significant differences in any of the comparisons.

Spillover region 2

In the second spillover region, comprising the responses to the third and fourth words following the evidentiality-marked verb, there was also a significant three-way interaction between group, grammaticality and evidential (see Figure 3C). In the *indirect* condition, similar to R-SO1, the emigrant speakers were faster when the evidentiality marker did not match with the information source than when it did ($\beta = 0.11$, $SE = 0.018$ $t = 6.26$, $p < 0.001^*$), and also showed a significant but much smaller difference between match and mismatch sentences for the *direct* evidential (ES $\beta = 0.04$, $SE = 0.018$, $t = 2.67$, $p = 0.032^*$).

The heritage language speakers also showed significantly faster reading times for mismatch than match sentences for both the *indirect* ($\beta = 0.06$, $SE = 0.020$ $t = 3.17$, $p = 0.008^*$), and *direct* conditions ($\beta = 0.08$, $SE = 0.020$, $t = 3.98$, $p < 0.001^*$).

When comparing the direct and indirect conditions, the patterns were the same as for R-SO1, in the mismatch conditions emigrant speakers were, once again faster in the *indirect* condition compared to the *direct* condition ($\beta = 0.06$, $SE = 0.018$, $t = 3.61$, $p = 0.002^*$), but not in the match condition ($\beta < 0.001$, $SE = 0.018$, $t = 0.02$, $p = 1.000$). Similar to R-SO1, the heritage language speakers did not show any significant differences between the direct and indirect conditions in R-SO2 (match: $\beta = 0.01$, $SE = 0.020$, $t = 0.51$, $p = 1.000$; mismatch: $\beta = 0.02$, $SE = 0.020$, $t = 1.32$, $p = 0.563$).

Final word

The model for word reading time at the Final Word region of interest revealed main effects only for grammaticality, but no significant interactions (see Table 4; Figure 3D). Reading times for *mismatch* sentences were significantly shorter than for the *match* sentences in this region.

Summary of results

The patterns of response across the analyses are summarised in Table 5.

Reading times during self-paced reading (i.e., online results) showed that the Turkish heritage language speakers were slower in their reading times overall compared to the other bilingual group, i.e., Turkish emigrant speakers. At the target verb, targets marked with direct evidential markers (-DI) were processed

TABLE 4 Mixed-effects estimates for the reading times at the regions of interest.

	Region target verb				Region spillover 1				Region spillover 2				Region final word			
	ß	SE	*t*	*p*	ß	SE	*t*	*p*	ß	SE	*t*	*p*	ß	SE	*t*	*p*
(Intercept)	6.47	0.04	171.59	<0.0**01***	6.19	0.03	223.01	**<0.001***	6.21	0.03	182.91	**<0.001***	6.95	0.06	109.68	**<0.001***
Region length	0.05	<0.001	9.80	**<0.001***	0.01	<0.001	3.07	**0.002***	0.03	<0.001	13.60	**<0.001***	0.02	<0.001	4.33	**<0.001***
Age	>-0.01	<0.001	−0.70	0.483	0.01	<0.001	1.69	0.097	0.01	<0.001	1.64	0.108	0.01	<0.001	1.74	0.088
Verbal working memory	−0.03	0.02	−1.40	0.167	−0.04	0.02	−2.42	**0.019***	−0.05	0.02	−2.27	**0.027***	−0.07	0.04	−1.84	0.072
Speaker group	−0.09	0.04	−2.09	**0.042***	−0.11	0.03	−3.53	**<0.001***	−0.16	0.04	−4.06	**<0.001***	−0.13	0.07	−1.73	0.089
Evidential	0.01	0.01	1.58	0.116	> −0.01	<0.001	−1.15	0.252	−0.01	0.01	−2.36	**0.020***	>-0.01	0.01	−0.32	0.748
Grammaticality	0.01	0.01	1.06	0.291	−0.01	<0.001	−2.50	**0.016***	−0.04	0.01	−7.00	**<0.001***	−0.05	0.01	−4.40	**<0.001***
Evidential—grammaticality	<0.001	0.01	0.23	0.817	−0.01	<0.001	−1.30	0.194	−0.01	0.01	−1.13	0.260	−0.01	0.01	−1.19	0.237
Evidential—group	−0.01	0.01	−2.40	**0.016***	−0.01	<0.001	−2.29	**0.022***	>-0.01	<0.001	−0.93	0.354	−0.01	<0.001	−1.48	0.138
Grammaticality—group	<0.001	0.01	0.24	0.811	−0.01	<0.001	−2.34	**0.019***	>-0.01	<0.001	−0.62	0.535	−0.02	<0.001	−1.68	0.092
Evidential—grammaticality—group	<0.001	0.01	0.23	0.821	−0.01	<0.001	−2.46	**0.014***	−0.01	<0.001	−2.55	0.**010***	>-0.01	<0.001	−0.98	0.327
Observations	<5504				<11013				11010				<5472			
Marginal R^2/conditional R^2	0.084/0.334				<0.097/0.267				0.154/0.368				<065/0.323			

lmer (log(RT) ~ Evidential * Grammaticality * Group + c. (RegionLength) + c. (VerbalWorkingMemory) + c. (Age) + (1 |Participant) + (1 |Item), data = data, REML = FALSE, control = lmerControl (optimizer = "bobyqa")). Note: Values of significant effects ($p < 0.05$) are printed in bold and asterisked.

10.3389/fcomm.2023.1070510

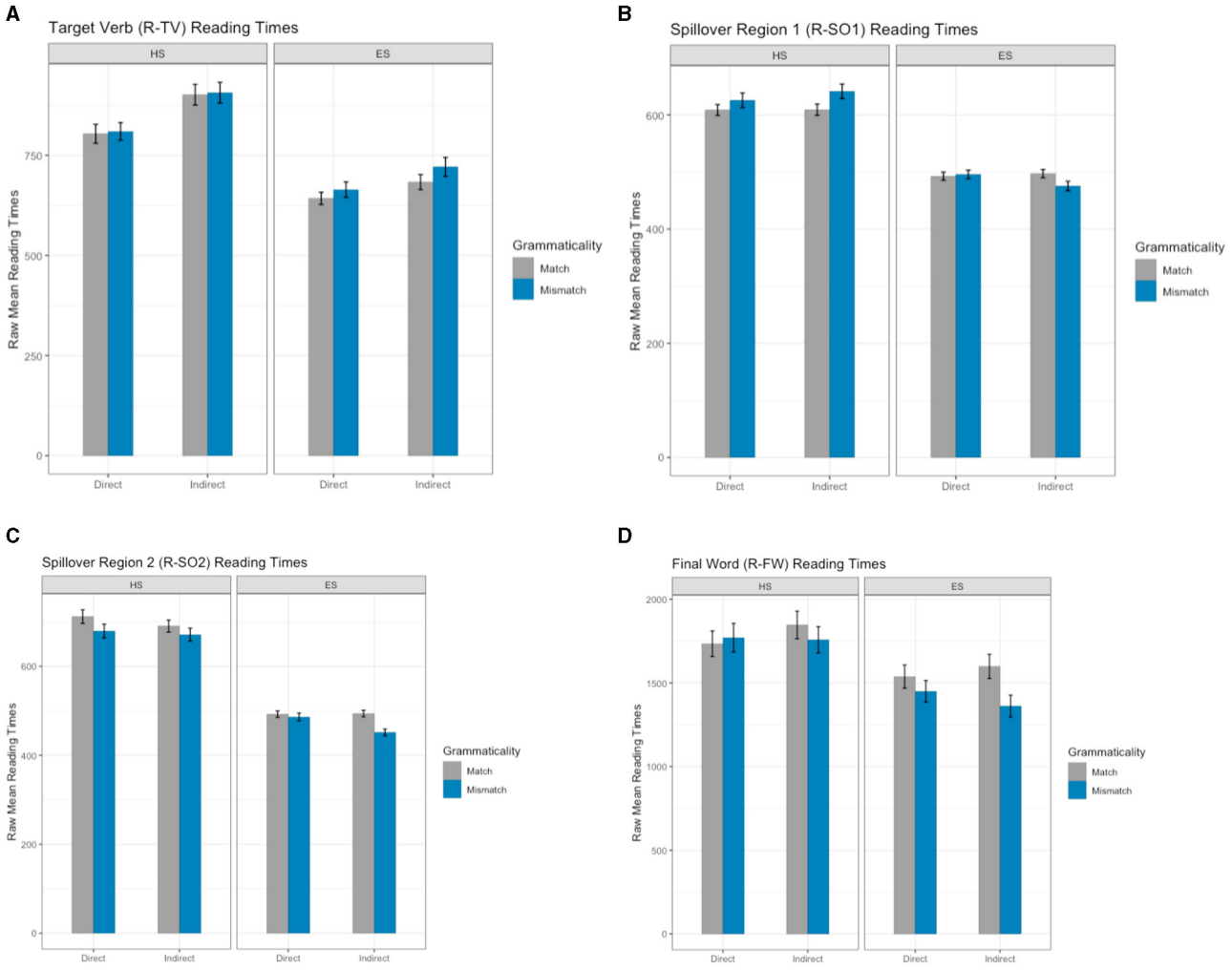

FIGURE 3
Raw reading times for the Target Verb region **(A)**, Spillover Region 1 **(B)** and 2 **(C)**, and the Final word region **(D)**. Error bars represent the standard error of the mean. The means and standart deviations are provided in Table A1.

faster than those marked with the indirect marker (-mIş) by the heritage language speakers whereas emigrant speakers showed no processing differences.

Emigrant speakers showed significantly faster reading times for mismatch than match sentences with indirect evidential markers at both Spillover Region 1 and Spillover Region 2. However, for the heritage language speakers, this pattern appeared later in the sentence, with a significant difference only at Spillover Region 2. Moreover, in Spillover Region 2, the same effect was apparent for sentences where the verb was inflected with the direct evidential: Both groups were faster at processing mismatched direct conditions compared to the direct match conditions.

In sum, in the second spillover region, both groups showed similar processing: they processed mismatching sentences faster. This faster processing of mismatching sentences was sustained in the final word region without the condition effect for both groups.

The speed of the-end-of-sentence acceptability judgements (i.e., offline results) revealed no effect of evidentiality condition or sentence grammaticality for the heritage language speakers. But the emigrant speakers, showed a similar pattern to their online responses in the spillover regions, responding marginally faster in mismatching sentences when the verb was inflected with the indirect evidential marker following a firsthand information source compared to sentences when the indirect evidential marker matched with a nonfirsthand information source. In the acceptable, match, sentences, their responses were significantly faster when the sentences included the direct evidential marker than the indirect evidential marker.

In terms of accuracy of sentence acceptability judgements, both heritage language speakers and emigrant speakers were more accurate in their judgements of the match vs. mismatch sentences in the direct evidential condition. In the indirect evidential condition, only heritage language speakers showed a significant difference, with match sentences being more accurately responded to than mismatch sentences. Both groups were more accurate in the direct than the indirect condition for the match sentences.

10.3389/fcomm.2023.1070510

TABLE 5 Summary of analyses examining patterns across reading time (online) and end of sentence judgement (offline) measures by condition.

		Online reading times								Offline end-of-sentence acceptability judgements			
		Target verb region		Spillover region 1		Spillover region 2		Final word region		Response Time		Accuracy	
		HLS	ES	HLS	ES	HLS	ES	HLS	ES	HLS	ES	HLS	ES
Match vs. mismatch													
Indirect	Match vs. mismatch				Mis < Mat	Mis < Mat	Mis < Mat	Mis<Mat	Mis < Mat		Mis = < Mat	Mat<Mis	
Direct	Match vs. mismatch					Mis < Mat	Mis < Mat	Mis < Mat	Mis < Mat			Mat<Mis	Mat<Mis
Direct vs. indirect		Dir < Ind											
Match	Direct vs. Indirect										Dir < Ind	Dir < Ind	Dir < Ind
Mismatch	Direct vs. indirect				Ind < Dir		Ind < Dir						

A < B, A is significantly faster, or less error prone (more accurate), than B; A ≤ B, A is marginally significantly faster than B ($p = 0.06$); HLS, Heritage Language Speakers; ES, Emigrant Speakers; Dir, Direct; Ind, Indirect; Mat, Match; Mis, Mismatch.

Discussion

This study was motivated by two issues in the study of the language processing of heritage language speakers: (1) increasing interest in the use of, and additional awareness of the importance of, online methods; (2) the importance of including a reference group that is comparable to the heritage language speakers. We, therefore, aimed to provide new insights into the processing of evidentiality in Turkish heritage language speakers using a task, self-paced reading, that provides online and offline measures, and by comparing their performance to that of a reference group of late bilingual emigrant speakers of Turkish also living in Australia. During the self-paced reading task, participants were presented with sentences that started with an indication of either a firsthand or a nonfirsthand information source followed by matching or mismatching evidentiality markers. We will first discuss the offline, behavioural data, from end of sentence judgements before moving to the online data and a comparison of the two.

In sentence acceptability judgements, both groups were more accurate in their judgements of the matching vs. mismatching sentences in the direct evidential condition. However, in the indirect evidential condition, the ES group were equally accurate for matching and mismatching sentences, whereas the heritage language speakers showed poor response accuracy for mismatching sentences relative to matching sentences for both evidential contexts. The heritage language speaker group showed poor response accuracy for mismatching sentences relative to matching sentences for both evidential contexts. Simply put, the heritage language speaker group showed a tendency to judge mismatching sentences to be acceptable in about 50% of all the trials: they lacked sensitivity in judging the acceptability of both the evidential forms in reference to given information source contexts. Such a finding is completely consistent with the pattern found in a listening task for the same heritage language speaker group (see Tokaç-Scheffer et al., to appear) and for a separate group of heritage language speakers residing in the Netherlands reported in Arslan et al. (2017).

As mentioned above, we did not observe a significant difference between direct and indirect evidentials in either groups' accuracy to mismatching sentences, suggesting that they considered mismatches of both information sources to evidential forms equally unacceptable to an extent. Following Arslan et al. (2017) data from Turkish monolingual speakers, we may have expected to find an asymmetry here with higher response accuracy in detecting violations when the firsthand information source mismatches to the indirect evidential than when the nonfirsthand information source mismatches to the direct evidential. Recall that such an evidential context, mismatching use of indirect evidential marker following firsthand information source, is what Aikhenvald (2004, p. 217) refers to as counter-intuitive. Note that in the previous literature this behavioural asymmetry was found for monolingual speakers (Arslan et al., 2017; Karaca, 2018; Schmid and Karayayla, 2019), and it is not surprising that this counter-intuitiveness was not reflected in our bilingual groups' responses who may have reduced sensitivity to evidentiality, as even the emigrant speaker group were living in conditions where Turkish was not the language of the society (i.e., what could be considered heritage language

conditions). It is important to note, however, as we discuss below, that we found such an asymmetry towards faster reading times in firsthand information sources mismatched to indirect evidential forms at the post-critical regions during word-by-word processing for the emigrant speakers. This suggests that living under heritage language conditions and being exposed to a majority language which lacks grammatical evidentiality limits and/or influences the representation of evidentiality in our groups of bilinguals but in online processing we still find a clear response to counter-intuitive evidential contexts by the late bilingual group who grew up in Türkiye.

We turn now to online processing in more detail. Our first research question asked whether the online processing of the Turkish heritage language speakers and the Turkish emigrant speakers differed during the time course of reading these sentences with evidentiality marking. As expected, the heritage language speakers were slower in their reading times across the critical segments compared to the emigrant speakers. This is not a surprising outcome when heritage language research is considered (e.g., Sekerina and Trueswell, 2011; Montrul, 2016). For example, in Hulsen et al. (2002), second-generation heritage language speakers of Dutch born in New Zealand, an experimental group similar to ours, were slower in word retrieval compared to first-generation adult migrant speakers of Dutch. However, as we discuss below, when we examined the pattern of responses of our participants in more detail, we found similarities as well as differences in the performance of the two groups.

Our second research question asked whether the speakers' online processing of evidentiality differed depending on whether the evidentiality marker on the verb was direct or indirect, and whether the two groups differed in this regard. We examined the reading times in four segments of interest including the segment with the target evidential-marked verb, a first spillover region comprised of the first two words following the target verb, a second spillover region with the next two words following the first spillover region, and the final word of the sentence. At the target verb region, the heritage language speaker group was faster in their reading times for the direct than the indirect evidential condition (irrespective of the grammaticality of the sentence). This condition difference seems to be consistent with previous heritage language research conducted with children (e.g., Aarssen, 2001; Karakoç, 2007) and adults (Schmid and Karayayla, 2019; Arslan, 2020; Arslan et al., 2020), all of which reported heritage language speakers' better processing of direct evidential forms. Note that, however, direct evidential is the default form in the Turkish evidentiality paradigm as opposed to the indirect evidential which is conceptually rather complex and semantically more "marked" as it refers to an assortment of contexts including reportative and inferential. Outcomes from language acquisition studies often mirror this asymmetry in monolingual Turkish-speaking children's acquisition trajectories in that the direct evidential marker emerges earlier in children's language, both in comprehension and production, and is also fully mastered prior to the indirect evidential marker (Aksu-Koç, 1988; Öztürk and Papafragou, 2008; Aksu-Koç et al., 2009; Ünal and Papafragou, 2016). Borrowing insights from these acquisition studies, we suggest that heritage language speakers growing up in homes where Turkish is spoken as a heritage language are likely to better acquire the direct evidential marker than the indirect evidential marker. A possibility here is that, as the majority language (English) which lacks grammatical evidentiality gains dominance after early childhood, the Turkish evidentiality paradigm may have been impacted by negative language transfer effects, and as a consequence, Turkish heritage speakers are tending to take direct evidentiality as a default past tense form (see also Arslan et al., 2015). The slower reading times of our heritage language speakers for the indirect than the direct evidential conditions, in fact, reflect the erosion of the indirect evidential in Turkish heritage grammar. The lack of such condition differences between the direct and indirect evidential forms in the emigrant speakers' reading times suggests that the individual bilingualism profiles of our bilingual participants (including dominant language setting during childhood, language of formal education) influence the way the evidentiality system erodes in Turkish heritage grammar.

The pattern of results in the spillover regions demonstrated that the emigrant speakers differed from the heritage language speakers in terms of timing, which suggested delayed processing for heritage language speakers compared to emigrant speakers. In the first spillover region, the emigrant speaker group was faster at reading mismatching sentences in the indirect condition, than matching sentences. In the mismatching sentences, a phrase indicating a firsthand information source was followed by an indirect evidential marker. This pairing is not plausible in any circumstance and is therefore easy for native Turkish speakers to reject. The heritage language speakers on the other hand showed a similar pattern but not until the second spillover region (where this pattern was still evident for the emigrant speakers). Hence, while the heritage language speakers were sensitive to the mismatch of firsthand information, they were slower to do so. A similar latency effect in evidentiality processing was shown by Karaca (2018): In comparison to monolinguals, bilingual first-generation migrants were slow to show sensitivity to the mismatch of firsthand information, relative to monolinguals. The significantly slower listening that monolinguals showed for the incongruent indirect evidential compared to congruent in the second segment did not appear until the third segment for first-generation immigrants. As pointed out by Jegerski (2014), any effects on latency during self-paced reading should be interpreted cautiously and that it is not simply that heritage language speakers are merely slower overall. Importantly, we showed that, despite their overall slower latencies, heritage language speakers showed qualitative similarities to the reference group with both groups showing similar sensitivity to the same evidential condition (direct vs. indirect) and grammaticality (match vs. mismatch) combination: when mismatch indirect sentences are compared to match sentences.

With regard to our results from the analysis of the second spillover region, the two groups were similarly faster at reading the mismatch sentences compared to match sentences in the direct condition. In contrast, in both the first and second spillover regions, only emigrant speakers exhibited a difference between the direct and indirect evidential conditions, and this was only in the mismatching sentences. That is, when a firsthand information source was violated by the use of the indirect evidential marker (mismatching indirect condition), they were faster than when

a nonfirsthand information source was violated by the direct evidential marker (mismatching direct condition). This was similar to the pattern shown by monolinguals in previous studies (Arslan et al., 2017; Tokaç-Scheffer et al., to appear). According to Aikhenvald (2004, p. 217) upon being given firsthand information, it is "counterintuitive" to challenge this experience with the use of indirect evidential and native speakers are faster at rejecting such mismatches.

It is worth reiterate that while the heritage language speakers showed a difference between direct and indirect sentences at the target verb (for both match and mismatch sentences), they did not show it in the spillover regions. The lack of an effect of condition in later processing shows some similarity to previous (offline and online) studies: Neither of the studies using (offline) auditory verification (go/no go) tasks (Arslan et al., 2017; Tokaç-Scheffer et al., to appear) revealed any effect of evidentiality condition for heritage language speakers. Similarly, during a self-paced listening task, heritage language speakers in Karaca (2018) did not show any significant processing differences between the conditions in any of the sentence segments. This is at odds with Arslan et al. (2015) findings from the eye-movement monitoring experiment which demonstrated that both early and late bilingual Turkish speakers were slower and less accurate to respond to direct evidential than indirect evidential conditions. Arslan et al. (2015) used a visual representation of evidence while their participants listened to sentences with evidential forms, and it seems these condition differences are reflecting a somewhat different aspect of evidentiality processing. The precise nature of these condition differences requires further studies critically investigating grammaticality/acceptability judgement tasks as opposed to naturalistic tasks. We will turn to this issue below.

In the spillover regions, it is also important to underline that there was a significant difference between match and mismatch sentences in the indirect condition. This effect was observed in both post critical regions for emigrant speakers, only in the second spillover region for heritage language speakers, possibly due to delayed processing speed. Both bilingual groups rejected the indirect mismatch sentences faster both in-between and within conditions. However, the difference between match and mismatch sentences in the direct condition was only observed in the second spillover region for both groups. We argue that participants were faster at reading times when presented with a violation of firsthand direct information with a mismatching indirect evidential marker.

The final word reading times were longer than those for any other word/segment of the sentence and this was true for both groups. It is important to note that in this experiment, the last word of the sentence was indicated by the presence of segment lines (see the Method section for details). Consequently, participants knew that with the next key press they would be asked to make a decision about the acceptability of the sentence they had just read. This resulted in participants generally having longer reading times for this segment compared to the rest of the sentence (see Figure 1). This is quite a common phenomenon in self-paced reading and eye-tracking studies known as the "wrap-up effect" (Just and Carpenter, 1980). At this stage, the parser evaluates and assesses all the presented information and resolves any "inconsistencies" (Just and Carpenter, 1980, p. 345). Longer reading times at the end of sentence have also been hypothesised to be due to readers preparing for the execution of the next task (Stowe et al., 2018). It is not possible to tease apart the role of these two effects in our experiment nor was it the purpose of this study. At this time point (the final word), qualitatively, the two groups showed similar processing with no effect of evidentiality condition and faster reading times for sentences with unacceptable evidentiality marker use following a mismatching information source. The re-evaluation taking place at the final word was much faster for the violated (mismatching) sentences for both groups. The similarities between the two groups that were observed in this final segment provide support for the validity of using a self-paced reading task. As highlighted out by previous researchers, self-paced reading offers qualitative insight on the nature of processing differences between heritage language speakers and the reference group they are being compared to. This extends beyond quantitative comparison of slower processing, and, furthermore, enables seeing whether the differences between monolinguals and non-monolinguals are diminishing or becoming less apparent (Jegerski et al., 2016; Jegerski, 2018a,b).

Our third research question sought to address the benefits of including online methods in heritage language studies and asked whether there were differences between online and offline processing of evidentiality for heritage language speakers and emigrant speakers. As noted above, the examination of the offline results, end-of-sentence judgement questions, revealed that overall, heritage language speakers were slower and less accurate than the emigrant speakers. End-of-sentence acceptability judgements are thought to measure metalinguistic knowledge gained most likely through formal teaching (Bayram et al., 2021). At this point, the individuals parsing the sentences are aware that they need to make a decision, so they re-evaluate their processing, complete the missing information, and solve the "linguistic problems" to make their judgements (Keating and Jegerski, 2015, p. 3). The emigrant speakers, who had received formal education in the language under investigation, showed better accuracy. However, we cannot ignore the fact that the heritage language speakers of the current study, who were residing in Australia had also participated in learning activities through community schools and, the accuracy of these heritage language speakers was higher than that of the heritage language speakers in other studies, who had not received schooling in their heritage language (Arslan et al., 2017; discussed in Tokaç-Scheffer et al., to appear).

The detailed examination of the offline patterns in comparison to the online patterns showed some differences for both groups. The differences between direct and indirect evidentiality conditions that were captured in the target verb region for heritage language speakers were reflected in the accuracy results but only for the match sentences. Their online processing of direct evidentiality markers was faster than that of the indirect markers, and at the (offline) end-of-sentence judgements they were more accurate at judging grammatical direct evidentiality sentences (matching between a firsthand information source and a direct evidential marker) compared to the indirect grammatical sentences. Why would the processing of direct evidential markers be more accurate and quicker as compared to their indirect counterparts? Above, we mentioned the impact of potential processing asymmetries given the primacy of the direct evidential in Turkish monolingual

children. This is based on the idea that indirect evidential is semantically more "marked" as it refers to a number of indirect information sources (i.e., inference, reportative), and the ability to monitor indirect information sources develop with a delay in children's acquisition. In adults, by contrast, such an asymmetry is not necessarily reflected in behavioural responses. For example, Arslan (2020), using a similar design to ours, showed that a group of Turkish monolingual speakers judged both evidential forms with around 90% task accuracy. Therefore, explaining the presence of strong asymmetry in our adult bilingual speakers as enhanced performance in direct evidential forms over indirect forms based solely on the markedness of the indirect form seems rather unreasonable.[5] In the absence of data on the developmental trajectories of evidential forms in children acquiring Turkish as a heritage language, we also cannot speculate on whether this asymmetry emerged as result of any possible maturational constraints. The only possible explanation that we can offer at present is the lack of a grammatical indirect evidential form in English, which is the dominant societal language for all our participants. It is conceivable that cross-linguistic transfer effects were at play here and these Turkish speakers in Australia developed a greater tendency towards accepting the direct evidential as the more plausible form in a past time context. Such instances of restructuring of Turkish grammar in contact with English have been attested before. For instance, Gürel (2002) showed that Turkish speakers in North America have attuned to English-like overt subject pronouns as opposed to null pronouns.

A second difference in online and offline processing was that our participants exhibited faster reading times for mismatching sentences for both indirect and direct conditions during moment-by-moment reading, while later, at the end of the trial, their judgements were more accurate for the matching sentences. A similar finding was reported for second language learners of English by Juffs and Harrington (1996), who found that participants were more accurate in their judgements of those sentences on which they spent more reading time.

A final point that is important to note is that there might be a task effect in measuring heritage language outcomes. Grammaticality/acceptability judgement tasks in heritage speakers have been suggested to lead to biassed outcomes as opposed to naturalistic tasks with time-sensitive online measures as these speakers tend to have low metalinguistic awareness of their heritage language (see Polinsky, 2018). Therefore, it is conceivable that the heritage language speakers under examination in this study were inaccurate in their judgements of sentence acceptability due either to insensitivity in their grammar and/or weakened language awareness of their heritage language. This argument fits in well with why some studies report no condition differences in heritage speakers' evidentiality processing (Arslan et al., 2017; Tokaç-Scheffer et al., to appear) while some others report critical condition differences (Arslan et al., 2015).

Evidentiality studies with heritage language speakers are scarce, which leaves limited room for comparisons. On the other hand, there are studies comparing online and offline results with other grammatical phenomena in other languages (e.g., Keating et al., 2016; Jegerski, 2018a,b). Keating et al. (2016), who studied the differential object marking in Spanish, obtained similar results that heritage language speakers showed processing differences between conditions during online self-paced reading but not in their reaction times to the comprehension questions. The lack of difference between the conditions on the sentence judgement results prove that what is observed via offline measures may not be as informative as online measures for non-monolingual processing.

In addition to providing detailed information on heritage language speakers' evidentiality processing patterns, while showing similarities and differences between online vs. offline tasks, this study also underlines the importance of choosing an appropriate reference group to compare heritage language speakers. Although the (late bilingual) emigrant speakers were faster and more accurate in their processing compared to the heritage language speakers, they had similar moment-by-moment processing of evidentiality to heritage language speakers. The similarities in sentence processing between bilingual groups (heritage language speakers and late bilinguals) during self-paced reading that have been shown in previous studies (Jegerski, 2018b) were confirmed by our results and the differences between online vs. offline processing were clear.

The results from both online and offline measures in the current study suggest that heritage language speakers do process the grammatical details of evidentiality. Although it has been hard to measure the extent of this knowledge comprehensively with offline measures, the moment-by-moment investigation revealed that heritage language speakers can activate and integrate this knowledge during online reading. Results from this study can motivate future research comparing heritage language speakers' processing to other bilingual groups with a variety of age of onsets and language backgrounds. This can help us understand and come to grounded conclusions on the important factors affecting evidentiality processing in bilinguals.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by Macquarie University Human Research Ethics Committee (approval number: 3531). The patients/participants provided their written informed consent to participate in this study.

Author contributions

ST-S primary author and the editor of the original manuscript collected the data. ST-S and SA prepared the stimuli and the task. All authors participated in project conceptualisation and implementation, data analysis, and preparation of this manuscript.

5 This point was also suggested by an anonymous reviewer.

Funding

This research was supported by a Macquarie University International Research Excellence Scholarship (iMQRS) held by ST-S.

Acknowledgments

The authors are grateful to Turkish Welfare Association and Cultural Centre in Sydney and their president Beşir Karasu, also the teachers and parents at Ataturk School who helped greatly with participant requirement; to Roelien Bastiaanse and Srđan Popov for their priceless feedback and discussion during the construction of this study. We also would like to thank Serje Robidoux and Solène Hameau for their statistical advice and support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2023.1070510/full#supplementary-material

References

Aarssen, J. (2001). "Development of temporal relations in narratives by Turkish-Dutch bilingual children," in *Narrative Development in a Multilingual Context*, eds L. T. Verhoeven and S. Stromqvist (Amsterdam: Benjamins), 209–231.

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

Aksu-Koç, A. (1988). *The Acquisition of Aspect and Modality: The Case of Past Reference in Turkish*. Cambridge: Cambridge University Press.

Aksu-Koç, A., Ögel-Balaban, H., and Alp, I. E. (2009). "Evidentials and source knowledge in Turkish," in *Evidentiality: A Window Into Language and Cognitive Development, New Directions for Child and Adolescent Development*, eds S. A. Fitneva and T. Matsui (San Francisco: Jossey-Bass), 13–28.

Anderson, J. A., Mak, L., Keyvani Chahi, A., and Bialystok, E. (2018). The language and social background questionnaire: assessing degree of bilingualism in a diverse population. *Behav. Res. Methods* 50, 250–263. doi: 10.3758/s13428-017-0867-9

Arslan, S. (2020). When the owner of information is unsure: epistemic uncertainty influences evidentiality processing in Turkish. *Lingua* 247, 102989. doi: 10.1016/j.lingua.2020.102989

Arslan, S., Bastiaanse, R., and Bayram, F. (2020). "First language exposure predicts attrition patterns in Turkish heritage speakers' use of grammatical evidentiality," in *Studies in Turkish as a Heritage Language*, ed F. Bayram (Amsterdam: John Benjamins), 105–126.

Arslan, S., Bastiaanse, R., and Felser, C. (2015). Looking at the evidence in visual world: eye movements reveal how bilingual and monolingual Turkish speakers process grammatical evidentiality. *Front. Psychol.* 6, 1387. doi: 10.3389/fpsyg.2015.01387

Arslan, S., De Kok, D., and Bastiaanse, R. (2017). Processing grammatical evidentiality and time reference in Turkish heritage and monolingual speakers. *Bilingual. Lang. Cogn.* 20, 457–472. doi: 10.1017/S136672891500084X

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Memory Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Statistic. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bayram, F., Kubota, M., Luque, A., Pascual y Cabo, D., and Rothman, J. (2021). You can't fix what is not broken: contextualizing the imbalance of perceptions about heritage language bilingualism. *Front. Educ.* 6, 628311. doi: 10.3389/feduc.2021.628311

Bayram, F., Rothman, J., Iverson, M., Kupisch, T., Miller, D., Puig-Mayenco, E., et al. (2017). Differences in use without deficiencies in competence: passives in the Turkish and German of Turkish heritage speakers in Germany. *Int. J. Bilingu. Educ. Bilingual.* 22, 919–939. doi: 10.1080/13670050.2017.1324403

Carminati, M. N. (2002). *The Processing of Italian Subject Pronouns, PhD thesis, University of Massachusetts Amherst*.

Chen, E., Gibson, E., and Wolf, F. (2005). Online syntactic storage costs in sentence comprehension. *J. Memory Lang.* 52, 144–169. doi: 10.1016/j.jml.2004.10.001

Clackson, K., Felser, C., and Clahsen, H. (2011). Children's processing of reflexives and pronouns in English: evidence from eye-movements during listening. *J. Mem. Lang.* 65, 128–44. doi: 10.1016/j.jml.2011.04.007

De Haan, F. (1999). Evidentiality and epistemic modality: setting boundaries. *Southwest J. Linguistics* 18, 83–101.

De Haan, F. (2005). "Encoding speaker perspective: evidentials," in *Linguistic Diversity and Language Theories*, eds Z. Frajzyngier, A. Hodges, and D. S. Rood (Amsterdam: John Benjamins B. V), 379–417.

Di Pisa, G., Rothman, J., and Marinis, T. (2022). *Gender and Number Agreement in Italian as a Heritage Language: A Self-paced Reading Study* [Poster presentation]. Heritage Languages Around the World, Lisbon, Portugal. Available online at: http://cehum.ilch.uminho.pt/languages#registration (accessed October 2022).

Drummond, A. (2013). *Ibex farm*. Available online at: http://spellout.net/ibexfarm (accessed September 30, 2021).

Felser, C., and Cunnings, I. (2012). Processing reflexives in a second language: The timing of structural and discourse-level constraints. *Applied Psycholinguistics* 33, 571–603. doi: 10.1017/S0142716411000488

Felser, C., Sato, M., and Bertenshaw, N. (2009). The on-line application of Binding Principle A in English as a second language. *Bilingual. Lang. Cogn.* 12, 485–502. doi: 10.1017/S1366728909990228

Filiaci, F. (2011). *Anaphoric Preferences of Null and Overt Subjects in Italian and Spanish: a Cross-linguistic Comparison, PhD thesis, The University of Edinburg*.

Fuchs, Z., Polinsky, M., and Scontras, G. (2015). The differential representation of number and gender in Spanish. *Linguistic Rev.* 32, 703–737. doi: 10.1515/tlr-2015-0008

Gürel, A. (2002). *Linguistic characteristics of second language acquisition and first language attrition: Overt vs. null pronouns*. [doctoral dissertation], McGill University.

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Lang. Cogn. Processes* 26, 376–405. doi: 10.1080/01690965.2010.492642

Hulsen, M., Bot, K. D., and Weltens, B. (2002). Between two worlds social networks, language shift, and language processing in three generations of Dutch migrants in New Zealand. *Int. J. Sociolinguistic.* 153, 27–52. doi: 10.1515/ijsl.2002.004

Jegerski, J. (2014). "Self-paced reading," in *Research Methods in Second Language Psycholinguistics*, eds J. Jegerski and B. VanPatten (New York, NY: Routledge), 20–49.

Jegerski, J. (2015). The processing of case in near-native Spanish. *Sec. Lang. Res.* 31, 281–307. doi: 10.1177/0267658314563880

Jegerski, J. (2016). Number attraction effects in near-native Spanish sentence comprehension. *Stud. Sec. Lang. Acquisit.* 38, 5–33. doi: 10.1017/S027226311400059X

Jegerski, J. (2018a). Sentence processing in Spanish as a heritage language: a self-paced reading study of relative clause attachment. *Lang. Learn.* 68, 598–634. doi: 10.1111/lang.12289

Jegerski, J. (2018b). The processing of the object marker *a* by heritage Spanish speakers. *Int. J. Bilingual.* 22, 585–602. doi: 10.1177/1367006916681083

Jegerski, J., Keating, G. D., and VanPatten, B. (2016). On-line relative clause attachment strategy in heritage speakers of Spanish. *Int. J. Bilingual.* 20, 254–268. doi: 10.1177/1367006914552288

Juffs, A., and Harrington, M. (1996). Garden path sentences and error data in second language processing research. *Lang. Learn.* 46, 286–324. doi: 10.1111/j.1467-1770.1996.tb01237.x

Just, M. A., and Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychologic. Rev.* 87, 329–354. doi: 10.1037/0033-295X.87.4.329

Just, M. A., Carpenter, P. A., and Wooley, J. D. (1982). Paradigms and processes in reading comprehension. *J. Exp. Psychol.* 111, 228–238.

Karaca, F. (2018). *Comprehension of Evidentiality in Spoken Turkish: Comparing Monolingual and Bilingual Speakers* (Unpublished Master's thesis), University of Alberta.

Karakoç, B. (2007). "Connectivity by means of finite elements in monolingual and bilingual Turkish discourse," in *Connectivity in Grammar and Discourse*, eds J. Rehbein, L. Pietsch, and C. Hohenstein (Amsterdam: John Benjamins), 199–227.

Karayayla, T., and Schmid, M. S. (2019). First language attrition as a function of age at onset of bilingualism: first language attainment of Turkish–English bilinguals in the United Kingdom. *Lang. Learn.* 69, 106–142. doi: 10.1111/lang.12316

Keating, G. D., and Jegerski, J. (2015). Experimental designs in sentence processing research. *Stud. Sec. Lang. Acquisit.* 37, 1–32. doi: 10.1017/S0272263114000187

Keating, G. D., Jegerski, J., and VanPatten, B. (2016). Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Bilingual. Lang. Cogn.* 19, 36–49. doi: 10.1017/S1366728914000418

Keating, G. D., VanPatten, B., and Jegerski, J. (2011). Who was walking on the beach? anaphora resolution in Spanish heritage speakers and adult second language learners. *Stud. Sec. Lang. Acquisit.* 33, 193–221. doi: 10.1017/S0272263110000732

Kim, M., Crossley, S. A., and Skalicky, S. (2018). Effects of lexical features, textual properties, and individual differences on word processing times during second language reading comprehension. *Read. Writ.* 31, 1155–1180. doi: 10.1007/s11145-018-9833-x

Lago, S., Stutter, A., G., and Felser, C. (2018). The role of native and non-native grammars in the comprehension of possessive pronouns. *Sec. Lang. Res.* 35, 319–349. doi: 10.31219/osf.io/v72gu

Lee, E., and Fraundorf, S. (2022). Do L1-L2 differences in discourse processing reflect processing demands or difficulty of form-function mapping?: evidence from self-paced listening of contrastive prosody. *Stud. Sec. Lang. Acquisit.* 44, 942–966. doi: 10.1017/S0272263121000619

Lenth, R. (2019). *Emmeans: Estimated marginal means, aka least-squares means. R package version 1.3.5.1*. Available online at: https://cran.r-project.org/web/packages/emmeans/index.html

Litcofsky, K. A., and Van Hell, J. G. (2017). Switching direction affects switching costs: behavioral, ERP and time-frequency analyses of intra-sentential codeswitching. *Neuropsychologia* 97, 112–139. doi: 10.1016/j.neuropsychologia.2017.02.002

Luce, R. D. (1991). *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford: Oxford University Press.

Mikhaylova, A. (2018). Morphological bottleneck: the case of Russian heritage speakers. *J. Lang. Contact* 11, 268–303. doi: 10.1163/19552629-01102005

Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingual. Lang. Cogn.* 5, 39–68. doi: 10.1017/S1366728902000135

Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: a case of morphosyntactic convergence. *Bilingual. Lang. Cogn.* 7, 125–142. doi: 10.1017/S1366728904001464

Montrul, S. (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers. *Int. J. Bilingual.* 13, 239–269. doi: 10.1177/1367006909339816

Montrul, S. (2016). *The Acquisition of Heritage Languages*. Cambridge: Cambridge University Press.

Montrul, S., and Bowles, M. (2009). Back to basics: Differential object marking under incomplete acquisition in Spanish heritage speakers. *Bilingual. Lang. Cogn.* 12, 363–383. doi: 10.1017/S1366728909990071

Montrul, S., Foote, R., and Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: the effects of age and context of acquisition. *Lang. Learn.* 58, 503–553. doi: 10.1111/j.1467-9922.2008.00449.x

Montrul, S., and Sánchez-Walker, N. (2013). Differential object marking in child and adult spanish heritage speakers. *Lang. Acquisit.* 20, 109–132. doi: 10.1080/10489223.2013.766741

Nicklin, C., and Plonsky, L. (2020). Outliers in l2 research in applied linguistics: a synthesis and data re-analysis. *Ann. Rev. Appl. Linguistic.* 40, 26–55. doi: 10.1017/S0267190520000057

Öztürk, Ö., and Papafragou, A. (2008). The acquisition of evidentiality in Turkish. *Univ. Pennsylvan. Working Papers Linguistic.* 14, 297–309. Available online at: https://repository.upenn.edu/handle/20.500.14332/44673

Polinsky, M. (2008). Relative clauses in heritage Russian: fossilization or divergent grammar? *Form. Approach. Slavic Linguistic.* 16, 333–358. Available online at: http://nrs.harvard.edu/urn-3:HUL.InstRepos:3403062

Polinsky, M. (2018). *Heritage Languages and Their Speakers* (Vol. 159). Cambridge: Cambridge University Press.

Polinsky, M., and Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. *Lang. Linguistics Compass* 1, 368–395. doi: 10.1111/j.1749-818X.2007.00022.x

Powell, M. J. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*.

R Core-Team (2012). R*: A Language and Environment for Statistical Computing. Vienna, Austria*. Available online at: http://www.r-project.org/

Rothman, J. (2007). Heritage speaker competence differences, language change, and input type: inflected infinitives in Heritage Brazilian Portuguese. *Int. J. Bilingual.* 11, 359–389. doi: 10.1177/13670069070110040201

Schmid, M. S., and Karayayla, T. (2019). The roles of age, attitude, and use in first language development and attrition of Turkish–English bilinguals. *Lang. Learn.* 70, 54–84. doi: 10.1111/lang.12361

Sekerina, I. A., and Trueswell, J. C. (2011). Processing of contrastiveness by heritage Russian bilinguals. *Bilingual. Lang. Cogn.* 14, 280–300. doi: 10.1017/S1366728910000337

Stowe, L. A., Kaan, E., Sabourin, L., and Taylor, R. C. (2018). The sentence wrap-up dogma. *Cognition* 176, 232–247. doi: 10.1016/j.cognition.2018.03.011

Tokaç-Scheffer, D. (2023). *An investigation of heritage language speakers of Turkish: Evidentiality processing and the effects of language experience* (Doctoral dissertation), University of Groningen, Groningen, Netherlands. doi: 10.33612/diss.593422246

Tokaç-Scheffer, S. D., Nickels, L., and Arslan, S. (to appear). "One suitcase, two grammars: What can we conclude about Australian Turkish heritage speakers' divergent processing of evidentiality?" in *Cognitive mechanisms driving contact-induced language change* (LangVang).

Ünal, E., and Papafragou, A. (2016). Production–comprehension asymmetries and the acquisition of evidential morphology. *J. Memory Lang.* 89, 179–199. doi: 10.1016/j.jml.2015.12.001

Vasishth, S., and Drenhaus, H. (2011). Locality in German. *Dialog. Discourse* 2, 59–82. doi: 10.5087/dad.2011.104

Xu, X., Chen, Q., Panther, K. U., and Wu, Y. (2018). Influence of concessive and causal conjunctions on pragmatic processing: online measures from eye movements and self-paced reading. *Discourse Process.* 55, 387–409. doi: 10.1080/0163853X.2016.1272088

10.3389/fcomm.2023.1070510

Appendix

Table A1 Means (M) and standard deviations (SD) for each condition of end-of-sentence judgement accuracy (SJ-Acc), response times (SJ-RT), and reading times at each region.

		SJ-Acc		SJ-RT		R-TV		R-SO1		R-SO2		R-FW	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
HLS	Direct match	0.75	0.43	1,847.9	1,898.1	803.7	577.4	608.6	330.9	711.9	530.6	1,733.8	1,859.8
	Direct mismatch	0.51	0.50	1,804.9	1,855.3	809.6	537	625.6	439.8	679.4	528.9	1,770.7	2,082.8
	Indirect match	0.64	0.48	1,871.8	2,003.8	901.4	632.5	609	337.3	690.5	468.7	1,846.5	2,018.1
	Indirect mismatch	0.51	0.50	1,945.5	2,082.6	906.5	634	641.3	437.3	671.1	507.2	1,758	1,903.9
ES	Direct match	0.85	0.36	1,500	1,441.2	642.4	428.5	492.7	281	492.5	284.1	1,538	1,926.2
	Direct mismatch	0.58	0.49	1,612.7	1,634.8	664.2	538.6	495.7	291.2	486.1	346.7	1,449.3	1,802.1
	Indirect match	0.69	0.46	1,662.8	1,585.3	683.1	526.4	497	285.7	493.8	288.4	1,598.7	2,012.5
	Indirect mismatch	0.65	0.48	1,537.8	1,557.3	721.3	657.6	475.4	331.3	451.5	291.8	1,361.7	1,818