Cognitive Psychology 139 (2022) 101519

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Demonstrative systems: From linguistic typology to social cognition

Paula Rubio-Fernandez

Department of Philosophy, University of Oslo, Norway

ARTICLE INFO

Keywords:
Demonstratives
Joint attention
Visual perspective taking
Peripersonal space
Social cognition

ABSTRACT

This study explores the connection between language and social cognition by empirically testing different typological analyses of various demonstrative systems. Linguistic typology classifies demonstrative systems as *distance-oriented* or *person-oriented*, depending on whether they indicate the location of a referent relative only to the speaker, or to both the speaker and the listener. From the perspective of social cognition, speakers of languages with person-oriented systems must monitor their listener's spatial location in order to accurately use their demonstratives, while speakers of languages with distance-oriented systems can use demonstratives from their own, egocentric perspective. Resolving an ongoing controversy around the nature of the Spanish demonstrative system, the results of Experiment 1 confirmed that this demonstrative system is person oriented, while the English system is distance oriented. Experiment 2 revealed that not all three-way demonstrative systems are person oriented, with Japanese speakers showing sensitivity to the listener's spatial location, while Turkish speakers did not show such an effect in their demonstrative choice. In Experiment 3, Catalan-Spanish bilinguals showed sensitivity to listener position in their choice of the Spanish distal form, but not in their choice of the medial form. These results were interpreted as a transfer effect from Catalan, which revealed analogous results to English. Experiment 4 investigated the use of demonstratives to redirect a listener's attention to the intended referent, which is a universal function of demonstratives that also hinges on social cognition. Japanese and Spanish speakers chose between their proximal and distal demonstratives flexibly, depending on whether the listener was looking closer or further from the referent, whereas Turkish speakers chose their medial form for attention correction. In conclusion, the results of this study support the view that investigating how speakers of different languages jointly use language and social cognition in communication has the potential to unravel the deep connection between these two fundamentally human capacities.

1. Demonstratives and the positive feedback loop hypothesis

Demonstratives – words like 'this' and 'that' in English, are also known as *directives* because they are used to orient the listener's attention towards an element in the speech situation, either physical (e.g., 'I prefer this one') or discoursive ('That was a good year'). Diessel (1999a, 2003, 2012a, 2012b) has shown that exophoric demonstratives serve two closely related functions: they indicate the spatial location of a referent relative to the *deictic center* (e.g., the speaker's position in English), and they coordinate the interlocutors' joint focus of attention. Diessel argues that coordinating speaker-listener joint attention is one of the most basic functions of language,

E-mail address: paula.rubio-fernandez@ifikk.uio.no.

https://doi.org/10.1016/j.cogpsych.2022.101519
Received 12 March 2022; Received in revised form 11 September 2022; Accepted 13 September 2022
Available online 24 November 2022

which also links demonstratives to social cognition. Building on extensive typological work by Diessel (1999a, 2006) and Evans and colleagues (2018a, 2018b; Bergqvist & Knuchel, 2019), I have recently argued that demonstratives are a lynchpin for the development of social cognition, training speakers of all languages in visual perspective taking (Rubio-Fernandez, 2020).

According to the *positive feedback loop hypothesis,* language and social cognition co-evolved in diachrony and co-develop in ontogeny through the acquisition and mature use of reference systems, including demonstratives (Rubio-Fernandez, 2020). Four features of demonstratives are key to this proposal (for discussion, see Diessel, 1999a, 2003, 2012a, 2012b):

1. Demonstratives are universal: they occur in all of the World's languages (Levinson, 2018).
2. Demonstratives are often accompanied by a pointing gesture, which is also a universal communicative device to establish joint attention (Kita, 2003).
3. Demonstratives emerge very early in language acquisition (Clark, 1978).
4. Demonstratives are so old, etymologically, that their roots cannot be traced back to other types of expressions. This suggests that demonstratives emerged very early in the evolution of language, probably because of their basic communicative function to co-ordinate the interlocutors' joint attention (Diessel, 2003).[1]

While demonstratives are universal, their meaning varies cross-linguistically. Thus, depending on the language, demonstratives may indicate not only the distance, but also the altitude, familiarity, position, reachability or visibility of a referent, from the perspective of the speaker, the listener, or both (Levinson, 2018). In line with the positive feedback loop hypothesis, the present study tested different typological analyses of various demonstrate systems as a way to investigate the connection between language and social cognition; or more specifically, how different demonstrative systems might make different demands on social cognition. Observing cross-linguistic differences in the recruitment of social cognition for referential communication has implications for our understanding of human cognition, which will be explored in the General Discussion.

2. Typological analyses of demonstrative systems

The present study aims to investigate the connection between language and social cognition by testing typological analyses of various demonstrative systems with adults. For example, English has a two-way demonstrative system distinguishing proximal and distal forms (e.g., 'this' vs 'that' or 'here' vs 'there'), while the Japanese system distinguishes three forms: proximal ('kore'), medial ('sore') and distal ('are') demonstratives. In linguistic typology, demonstrative systems are classified as *distance-oriented* or *person-oriented*, depending on whether they indicate the location of a referent relative only to the speaker, or to both the speaker and the listener. The English demonstrative system is distance-oriented, so the deictic center is always the speaker. The Japanese system, on the other hand, is person oriented, such that the proximal form indicates proximity to the speaker, the medial form indicates distance from the speaker but proximity to the listener, and the distal form indicates distance from both speaker and listener (Diessel, 2005, 2012a; Evans et al., 2018a).

Given the sensitivity of the Japanese demonstrative system to the spatial location of the listener, Japanese speakers must monitor their interlocutor's position in order to accurately use demonstratives. For example, the choice between the medial 'sore' and the distal 'are' hinges on whether the listener is close or far away from the referent. English speakers, on the other hand, use demonstratives from their own, egocentric perspective and need not consider the position of their interlocutor in order to select the appropriate demonstrative. These cross-linguistic differences suggest that while all demonstrative systems are used to direct an interlocutor's attention to a referent, only person-oriented systems require monitoring the listener's spatial location, potentially leading to the automatization of these processes (Goody, 1995). Empirically testing typological distinctions such as distance-oriented vs person-oriented demonstrative systems can therefore help us understand the demands that language poses on social cognition.

It must be noted, however, that the distinction between distance-oriented and person-oriented systems is not a complete characterization of demonstratives across languages, neither does it exhaust the connection between demonstrative use and social cognition. For example, several typological analyses of Turkish demonstratives have argued that their medial term 'şu' is used for referents not yet in joint attention (Hayasi, 1985, 1988, 1989 reviewed in Balpinar, 2019a, 2019b; Özyürek, 1998; Küntay & Özyürek, 2006; Hayasi & Özsoy, 2015), highlighting the connection between demonstrative use and attention monitoring. Likewise, a growing number of pragmatic studies have confirmed that interactive factors affect demonstrative use across languages, above and beyond the distance between the interlocutors and the intended referent (e.g., Piwek et al., 2008; Peeters et al., 2014, 2015; Rocca et al., 2019a, 2019b; Reile et al., 2020; Shin & Morford, 2020; Shin et al., 2020; Stukenbrock, 2020; Skilton & Peeters, 2021). For example, Rocca et al. (2018) observed a right-lateralized bias in the use of proximal demonstratives in Danish, with participants using the proximal form more frequently when the referent object was closer to their right hand. This bias suggests that proximal demonstratives are more likely to be used for referents affording easier manual manipulation. Relatedly, Rocca et al. (2018) also observed that Danish speakers shifted their peripersonal space towards their shared space with the listener when they were actively collaborating on a task, but not when the other person was merely present.

In light of these and other results, it is important to understand that speakers of languages with distance-oriented demonstrative systems (such as Catalan, English and Turkish in the present study) do not necessarily use their demonstratives "egocentrically." The

[1] See Heine et al. (2020) for a recent investigation of the origins of demonstratives showing that some demonstrative forms evolved from motion verbs (cf. Diessel & Coventry, 2020). Importantly, motion verbs also express spatial deixis, so on either account, it is deixis all the way down.

only diagnosis that we can make from confirming that they have a distance-oriented system is that listener position does not feature in the semantics of these demonstratives (e.g., other things being equal, the choice between 'aquest' vs 'aquell' or 'this one' vs 'that one' is not contingent on the listener's distance to the referent). While this is markedly different from what should be observed in languages with person-oriented systems (such as Japanese and Spanish in the present study), that does not mean that speakers of languages with distance-oriented systems are altogether insensitive to their social interaction with the listener in their use of demonstratives (see, e.g., Piwek et al., 2008; Peeters et al., 2014, 2015; Rocca et al., 2018, 2019a, 2019b; Shin et al., 2020). On the contrary, the view of demonstratives investigated here is social and interactive in nature (Rubio-Fernandez, 2020), and typological analyses will only be tested as a window into the socio-cognitive demands of different demonstrative systems.

3. A cross-disciplinary debate on the nature of demonstratives

Peeters and Özyürek (2016; see also Peeters et al., 2014, 2015) have argued against what they call the *egocentric view of demonstratives*, according to which demonstratives indicate distance relative to the speaker (e.g., Diessel, 1999b, 2014). Challenging traditional theories based on typological analyses, Peeters and Özyürek defend that demonstratives indicate the *psychological proximity* of the referent to both the speaker and the listener. Peeters et al. (2015:80-82) have further questioned the validity of typological analyses (such as the person-oriented vs distance-oriented distinction) because of typologists' reliance on reference grammars and linguistic intuitions, rather than on language use.

In line with Peeters and Özyürek (2016), the view of demonstratives investigated here is social and interactive in nature and will be tested using experimentally controlled tasks. However, unlike Peeters et al. (2015), the present study will rely on typological analyses of demonstrative systems, rather than challenging their predictions as contrary to the social view. In fact, traditional typological analyses will be tested empirically as a way to investigate some of the demands that different demonstrative systems pose on social cognition. The experiments reported in this study focused on two socio-cognitive abilities that are recruited in interactive demonstrative use: speakers' sensitivity to the listener's position (which distinguishes speaker-oriented vs distance-oriented systems; Diessel, 2005, 2012a) and speakers' sensitivity to the listener's focus of attention (which I have argued should be a universal feature of all demonstratives; Rubio-Fernandez, 2020). Once again, it is important to bear in mind that these two socio-cognitive skills are not intended to fully explain demonstrative use across contexts and languages, but rather provide experimental hypotheses that are theoretically motivated and could help us better understand how demonstrative use recruits social cognition during face-to-face interaction.

The two socio-cognitive skills investigated here are interesting because they differ in their cross-linguistic scope, with speakers' sensitivity to the listener's position being language specific (to the extent that only the semantics of person-oriented systems require monitoring the listener's position for accurate demonstrative use), whereas speakers' sensitivity to the listener's focus of attention should be universal. These abilities are therefore a test case for the theoretical claim that both universals and cross-linguistic differences in reference systems should offer insights into the connection between language and social cognition (Rubio-Fernandez, 2020). Also importantly, monitoring the listener's position may be characterized as a social ability insofar as it requires monitoring an agent's spatial location. However, monitoring the listener's attention focus is not simply a social skill involving the tracking of other agents in space, but a *mentalistic ability* that requires representing another's mental states.[2] Therefore, speakers' monitoring of the listener's attention focus is a fundamental Theory of Mind ability that may be trained through the acquisition and regular use of demonstratives across languages and cultures (Rubio-Fernandez, 2020).

Crucial to our understanding of the semantic and pragmatic factors underlying demonstrative use, recent experimental studies have investigated demonstrative use in different languages using a variety of paradigms (see, e.g., Piwek et al., 2008; Peeters et al., 2014, 2015; Rocca et al., 2018, 2019a, 2019b; for a review, see Rubio-Fernandez, 2020). These studies have clearly shown that both spatial and interactive factors affect demonstrative use, highlighting the social nature of demonstratives (Peeters & Özyürek, 2016; Shin et al., 2020). However, the experimental designs used in these studies do not serve as a test of traditional typological analyses of demonstrative systems. For example, Coventry et al. (2008) observed that Spanish-speaking participants were sensitive to the listener's position when using the proximal and medial demonstratives (i.e. 'este' vs 'ese'), whereas their English-speaking participants did not show sensitivity to listener position in their use of 'this' vs 'that'. However, Coventry et al.'s 'memory game' cannot elucidate whether the Spanish demonstrative system is a person- or distance-oriented system because the relative position of the referent and the listener were not fully crossed in their experimental design (i.e. the target object could be placed in four different locations along a table, whereas speaker and listener were sitting either on opposite ends of this long table or side by side at the same end). Thus, the memory game may reveal *addressee effects* in demonstrative use (Meira, 2003), but cannot determine if the Spanish medial form is reserved for referents far from the speaker but close to the listener, while the distal form is used for referents far from both interlocutors – as it would be the case if Spanish had a person-oriented demonstrative system. The experimental design used in the present study allows testing precisely those predictions.

4. Controversy around the Spanish demonstrative system

Spanish has a three-way demonstrative system distinguishing proximal ('este'), medial ('ese') and distal ('aquel') forms. The

[2] Thanks to Julián Jara-Ettinger for pointing out the difference between social and mentalistic abilities, which is a fundamental distinction in Theory of Mind research.

classification of this system has been contested, with some analyses characterising it as distance-oriented (e.g., Hottenroth, 1982; Kemmerer, 1999; Levinson, 2004), while others situate it among person-oriented systems (e.g., Alonso, 1968; Cifuentes- Honrubia, 1989; Eguren, 1999), or a combination of both (Jungbluth, 2003). One of the aims of the present study was to test these analyses of the Spanish demonstrative system experimentally.

Jungbluth's (2003) analysis deviates from traditional distance- vs person-oriented accounts and has received considerable attention in recent years (e.g., Peeters et al., 2015; Peeters & Özyürek, 2016; Levinson, 2018; Diessel & Coventry, 2020; Rocca & Wallentin, 2020; Shin et al., 2020; Peeters et al., 2021; Skilton & Peeters, 2021). Jungbluth argues that when Spanish speakers engage in face-to-face conversation, they 'treat their shared conversational space as uniform. Everything inside the conversational dyad is treated as proximal without any further differentiation' (Jungbluth, 2003:19). More specifically, she argues that the proximal demonstrative 'este' is generally used for referents at any location inside the face-to-face dyad. Peeters et al. (2021) recently noted that Jungbluth's proposal is in stark contrast with the results of the psycholinguistics study by Coventry et al. (2008), which showed that Spanish demonstratives express the relative distance of the referent from the speaker's position, but are also sensitive to the listener's location.

In a recent experimental study, Shin et al. (2020) used an interactive puzzle-building task in which Spanish-speaking participants had to ask an experimenter for a series of puzzle pieces. The participant was seated face-to-face with the experimenter and the space between them was shared, visible and accessible to both participants. This set up allowed Shin and colleagues to test Jungbluth's (2003) account of the Spanish proximal demonstrative as the default form inside the face-to-face dyad. Shin et al. (2020) observed that both spatial and interactive factors affected the use of the proximal and medial demonstratives in Spanish, but their results did not support Jungbluth's analysis: "The fact that there was variation on both sides, albeit very little on the proximal side, suggests that in this task the origo is not consistently situated in the shared space, contrary to the account of Jungbluth (2003). The variability in demonstrative selection in the distal context indicates that the further the referent is from the speaker, the more negotiation is needed for successful referent identification" (p. 16).

Peeters et al. (2021) describe Jungbluth's (2003) proposal as an 'in depth analysis of the Spanish demonstrative system.' Diessel and Coventry (2020) also offer an extensive discussion of Jungbluth's (2003) *face to face*, *face to back* and *side by side* analysis of demonstrative use in Spanish. However, for all its nuance in distinguishing different configurations of the conversational dyad, Jungbluth's work suffers from a fundamental shortcoming: her analysis of face-to-face, face-to-back and side-by-side conversation seems to be based on a single naturalistic example of each situation (at least no other naturalistic examples or statistical analyses of any corpus are provided). In addition, the analysis of naturalistic face-to-face conversation, which runs counter to the results of Coventry et al. (2008) and Shin et al. (2020), is potentially based on a mis-analysis of the example in question.

The exchange in Table 1 between a seller and a buyer at a market stall is the naturalistic example used by Jungbluth (2003) to support her view that Spanish speakers treat their conversational space as uniform, using proximal demonstratives to refer to any object inside the conversational dyad, without space considerations. In Spanish, as in other languages such as Japanese, Korean and Mandarin Chinese, demonstrative forms have been grammaticalized as *fillers* (also known as *hesitation markers*, *planners* or *filled pauses*), which are equivalent to etymologically opaque fillers such 'uh' and 'um' in English (see Kouteva et al., 2019). More specifically, the proximal form 'este' is the most common hesitation marker in Spanish (Clark & Fox Tree, 2002; Soler Arechalde, 2008; Graham, 2013, 2018; Stoesslein, 2014). Three discourse features suggest that the above example is an instance of the so-called 'tip of the tongue' effect:

- Disfluency: the speaker takes 2 s to produce the second word in the utterance, and then 4 s to produce the third word.
- Difficulty with lexical retrieval: the speaker wonders out loud what is the name of the vegetable they have in mind.
- Noun phrase choice: the speaker starts the utterance with a definite article ('las') and ends up producing a definite description ('las lechugas'/ 'the lettuces'), not using the demonstrative 'estas' in formulating the final noun phrase.

Despite these clear hesitation cues, Jungbluth (2003) analyses the use of 'estas' as a "true" proximal demonstrative (see the annotations in Table 1), concluding that Spanish speakers do not make spatial distinctions within the conversational dyad. It is worth noting that even if one wanted to trust in Jungbluth's criterion (assuming perhaps that her analysis is based on an extensive corpus, whose analysis was simply not reported in the paper), the fact that she selected this example as a canonical use of a proximal demonstrative in Spanish, without even acknowledging the ambiguity inherent in these forms (between deictic and filler uses), strongly suggests that she did not distinguish the two meanings of proximal demonstrative forms in Spanish.

In order to reconcile the seemingly incompatible results of Jungbluth (2003) and Coventry et al. (2008) (see also Shin et al., 2020), Peeters et al. (2021) argue that a new analysis of demonstrative use is in order. Here I propose an alternative explanation of these alleged discrepancies: Jungbluth (2003) mis-analyzed fillers as proximal demonstratives in Spanish, which led her to wrongly conclude

Table 1
Example 2 from Jungbluth (2003:20).

Speaker	Utterance and English translation	Jungbluth's annotations
Seller	***¿Qué más quiere?*** [What else do you want?]	The market woman is surrounded by lots of crates with different kinds of vegetable and fruit.
Buyer	***Las (2 sec.), estas (4 sec.). ¿Cómo se llaman? Las lechugas.*** [The (2 *sec.*), these (4 *sec.*). What's their name? The lettuces.]	The customer refers to the lettuces with 'estas' although they are nearer to the woman and they belong to her.

that Spanish speakers do not make spatial distinctions within the conversational dyad. One of the aims of this study was therefore to empirically test Jungbluth's conclusion that Spanish speakers treat the shared conversational space as uniform. This is an important test since Jungbluth's (2003) analysis is becoming the received view in the literature (see Peeters et al., 2015; Peeters & Özyürek, 2016; Levinson, 2018; Diessel & Coventry, 2020; Rocca & Wallentin, 2020; Peeters et al., 2021; Skilton & Peeters, 2021).

5. Aims and scope of the study

The overall aim of this study was to empirically test typological analyses of several demonstrative systems in order to explore some of the socio-cognitive demands that they pose on speakers of different languages (Rubio-Fernandez, 2020). Experiment 1 tried to replicate and extend the results of Coventry et al. (2008) by comparing demonstrative use in English and Spanish. The aim of the first experiment was to more accurately investigate the nature of the Spanish demonstrative system as distance oriented vs person oriented (Diessel, 2005) by varying referent and listener position parametrically. Thus, unlike in Coventry et al.'s memory game, the location of the target and the listener were fully crossed along four positions on a table (see Fig. 1).

Experiment 2 compared demonstrative choice in two languages with three-way demonstrative systems; namely, Japanese and Turkish. Since English and Spanish have two- and three-way demonstrative systems, respectively, the cross-linguistic differences observed by Coventry et al. (2008) could boil down to a more basic distinction between two- vs three-way systems. In other words, all three-way demonstrative systems may require that speakers monitor their listener's position in order to select the appropriate demonstrative form, whereas two-way systems do not require such sensitivity. Alternatively, three-way demonstrative systems may not always encode the listener's position in their medial and distal forms, revealing cross-linguistic differences that are relevant to the recruitment of social cognition in face-to-face communication. Experiment 2 investigated this question by comparing demonstrative choice in Japanese and Turkish.

Experiment 3 compared demonstrative choice in bilingual Catalan-Spanish speakers, who were administered the same task in one of their two languages. While Spanish has a three-way demonstrative system, the Catalan spoken in the region of Catalonia has undergone a paradigm reduction from three to two terms (Badia i Margarit, 1981; Perez Saldanya, 2015; Todisco et al., 2021a), such that their speakers only distinguish 'aquest' vs 'aquell' ('this' vs 'that'). The third experiment therefore investigated the degree to which Catalan-Spanish bilinguals are sensitive to their listener's position when using demonstratives in Spanish. Since Catalan has a two-way system that indicates relative distance from the speaker, bilinguals' use of Spanish demonstratives may reveal less sensitivity to the listener's position than monolinguals', especially for the medial demonstrative, which is no longer used in Catalonia.

Finally, Experiment 4 tested the hypothesis that demonstrative use does not only depend on the spatial location of the referent, but also on where the listener is looking. That is, demonstratives can be used to *correct the listener's focus of attention* when the speaker and listener perspectives are misaligned. Some typological analyses of Turkish argue that the medial demonstrative 'şu' is used for referents not yet in joint attention (Hayasi, 1985, 1988, 1989 reviewed in Balpinar, 2019a, 2019b; Özyürek, 1998; Küntay & Özyürek, 2006; Hayasi & Özsoy, 2015). The Japanese demonstrative system has been compared to the Turkish system in this respect (Özyürek & Kita, 2000; Levinson, 2004; Küntay & Özyürek, 2006; Küntay, 2012; Küntay et al., 2014). While these analyses suggest that in Turkish and Japanese the medial demonstrative may have been lexicalized as an attention correction form, I hypothesize that speakers of other languages may use their proximal and distal demonstratives flexibly, depending on where their listener is looking. This would suggest that some languages may have a preferred demonstrative form for attention correction, whereas others would flexibly use their proximal and distal demonstratives with the same pragmatic function. This hypothesis was tested in Experiment 4 with native speakers of Japanese, Spanish and Turkish.

All the experiments in this study employed an online task with static images that parametrically varied referent and listener position along a table (see Fig. 1 and Figs. 2-17). An earlier study by Peeters et al. (2014) successfully elicited demonstratives from Dutch speakers using a series of static images of a man and a woman communicating across a table. Earlier studies have also used static images to investigate demonstrative comprehension (Stevens & Zhang, 2013; Peeters et al., 2015). Like the present study, more recent studies of demonstrative production have employed online demonstrative-choice tasks, also revealing different usage patterns without real-time interaction (see Rocca et al., 2019a; Rocca & Wallentin, 2020; Todisco et al., 2021b).

While online data collection cannot possibly replace lab-based experiments and naturalistic observations (where pointing, gaze following and real-time interaction affect the use of demonstratives; Cooperrider, 2016), picture-based tasks allow researchers to control for a number of interactive factors such as referent position, speaker and listener distance to the referent and their focus of attention (including joint attention). Given the relevance of these factors to the study of social cognition in referential communication, the present study took advantage of the accuracy of experimental manipulations in online testing, while acknowledging the limitations of this form of language elicitation relative to real-time interactive paradigms and naturalistic observation. As a first test of the reliability of the online paradigm used in the study, Experiment 1 tried to replicate and extend the results of the lab study by Coventry et al. (2008). However, future studies should try to replicate the results of this study using real-time interactive tasks.

6. Experiment 1

In distance-oriented systems, demonstrative choice is dictated by the relative distance of the referent from the speaker. Under this account, Object Position (relative to the speaker), but not Listener Position, should determine participants' responses in both English and Spanish.

According to the person-oriented account, Spanish proximal demonstratives (e.g., 'este') indicate proximity to the speaker, and would be unaffected by Listener Position. Thus, for Object Position 0 (i.e. when the referent is closest to the speaker) both theoretical

accounts make the same predictions in Spanish. However, under the person-oriented account, Spanish medial demonstratives (e.g., 'ese') indicate distance from the speaker but proximity to the listener. Thus, in Object Positions 1 and 2, participants' choice of 'ese' should be affected by Listener Position. Finally, distal demonstratives (e.g., 'aquel') indicate distance from both speaker and listener. Therefore, responses in Position 3 should also be sensitive to Listener Position (with participants choosing between 'ese' and 'aquel' depending on the listener's proximity to the target object).

Finally, according to Jungbluth (2003), in face-to-face conversation, Spanish speakers treat their shared conversational space as uniform, referring to everything inside the conversational dyad as proximal without any further differentiation. The speaker and listener depicted in this task were standing face to face (with the target object standing in between them) in three trials: when both the target and the listener were in Position 0, when the target was in Position 1 and the listener in Position 2, and when the target was in Position 2 and the listener was in Position 3 (see Fig. 6). To test Jungbluth's account experimentally, participants' choice of the proximal demonstrative 'este' in face-to-face trials was analyzed separately to test for a potential effect of Object Position. Such an effect would challenge Jungbluth's conclusion that Spanish speakers do not establish spatial distinctions within the conversational dyad.

6.1. Methods

6.1.1. Participants

104 participants were recruited through Prolific, a crowdsourcing platform. N was originally set to 50 in each language group, but four participants timed out (one in the Spanish-speaking group and three in the English-speaking group), so extra participants were automatically recruited and their data was retained for analysis. All participants were compensated for their time.

In order to recruit a comparable sample to Coventry et al. (2008), for the Spanish-speaking group, the Demographics in Prolific were set to monolingual speakers of Spanish, who were Spanish nationals currently living in Spain. For the English-speaking group, demographics were set to monolingual speakers of English, who were British nationals currently living in England.

6.1.2. Materials and procedure

A series of 16 displays were created showing a boy and a girl on opposite sides of a table (for a sample display, see Fig. 1). The speaker in each display represented the participant (labelled 'You') and the listener represented the participant's friend (labelled 'Her'). The cover story for the task was that the participant was moving houses and asked a friend for help packing his things in boxes. In each trial, there were four different objects on the table, and the participant had to ask his friend to pass him one of the objects.

The target was the object that the speaker was looking at in each trial, as determined by both his body orientation and two intermittent white lines representing his line of gaze (i.e. connecting his eyes with the target object). The speaker position was fixed across trials (upper right corner of the table, or Position 0), whereas the target object position varied along the table (Positions 0–3, moving left-wise). The friend's position also varied across trials (Positions 0–3), resulting in a fully-crossed 4x4 design (Object Position × Listener Position).

Participants were asked to adopt the role of the speaker and complete the request in the speech bubble ('Now I need…') by clicking one of three radio buttons in Spanish: 'este' (this one), 'ese' (that one) or 'aquel' (yonder). In English, participants were given a choice of two radio buttons: 'this one' or 'that one'. The instructions highlighted that participants should treat the scene as an interactive scenario, and imagine what their natural choice would be in each trial.

To avoid that participants would develop strategies ahead of the task, the sample display shown in the instructions included 3 nouns to choose from ('The camera'/ 'The plane'/ 'That duckie'), rather than the demonstrative pronouns they would be offered during the actual task. The task was very short (16 trials / approx. 3 min.) and the trials were randomized individually to reduce the chance

Fig. 1. Spanish version of a sample display from the online task used in Experiments 1–3. Participants were asked to adopt the role of the speaker (figured labelled 'You' in Spanish), who is asking a friend to pass him one of the objects on the table. The speaker is always in Position 0, whereas the positions of the target object and the friend (figure labelled 'Her' in Spanish) varied across trials. In this instance, the friend is in Position 1, while the target is in Position 2. The speech bubble says 'Now I need…' in Spanish. To complete the speaker's request, participants were given a choice of two or three demonstrative pronouns, depending on the language.

that participants would develop strategies during the task.

The task was set up using the online survey builder Qualtrics.

6.2. Results

Descriptive statistics are plotted in Figs. 2-6 (see also the Appendix). Following Coventry et al. (2008), English and Spanish data were analyzed separately.

6.2.1. Spanish data

Because Spanish has a three-way demonstrative system, analyses distinguished a *Proximal zone* (corresponding with Object Position 0), a *Medial zone* (corresponding with Object Positions 1 and 2) and a *Distal zone* (corresponding with Object Position 3). The proximal demonstrative 'este' was the most frequent choice in the Proximal zone (93 % on average), while the medial demonstrative 'ese' was the preferred option in the Medial zone (68 % on average) and the distal demonstrative 'aquel' was the most frequent choice in the Distal zone (80 % on average).

6.2.1.1. Proximal zone (Object Position 0). Using logistic mixed effects regression, the binary outcome variable of Proximal Demonstrative choice ('ESTE' = 1, 'ESE' and 'AQUEL' = 0) was modelled for Object Position 0 (or Proximal zone), with Listener Position (0–3) as a continuous predictor variable, with the Speaker's fixed position as the zero point. All analyses in the study were run using R statistical software (R Core Team, 2019) and all models were fit with the maximal random effect structure for Participants and Items (Barr et al., 2013), according to the experimental design of the tasks. The first model included random intercepts for Participants and Items and random by-participant slopes for Listener Position. In this and all other models in the study, no random by-item slopes for Listener Position or Object Position were included because there was only one trial for each Object Position × Listener Position combination. However, the objects on the table changed across trials, potentially introducing by-item variability that was accounted for with a random intercept for Items. Because of model convergence issues, in this first model the random intercept and random slope effects for Participants were uncorrelated.

As predicted by both theoretical accounts, Listener Position did not have a significant effect on the choice of the proximal demonstrative 'este' in Object Position 0 ($\beta = 1.345$; SE = 1.292; $p = 0.2979$; see Fig. 2).

6.2.1.2. Medial zone (Object Positions 1 and 2). Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('ESE' = 1, 'ESTE' and 'AQUEL' = 0) was modelled for Object Positions 1 and 2 (or Medial zone), with Listener Position (0–3) and Object Position (1–2) as continuous predictor variables, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position and Object Position plus their interaction.

As predicted by both the distance-oriented and person-oriented accounts, Object Position had an effect on the choice of the medial demonstrative 'ese', albeit only marginally significant ($\beta = -1.2746$; SE = 0.6542; $p < 0.0514$). Only supporting the predictions of the

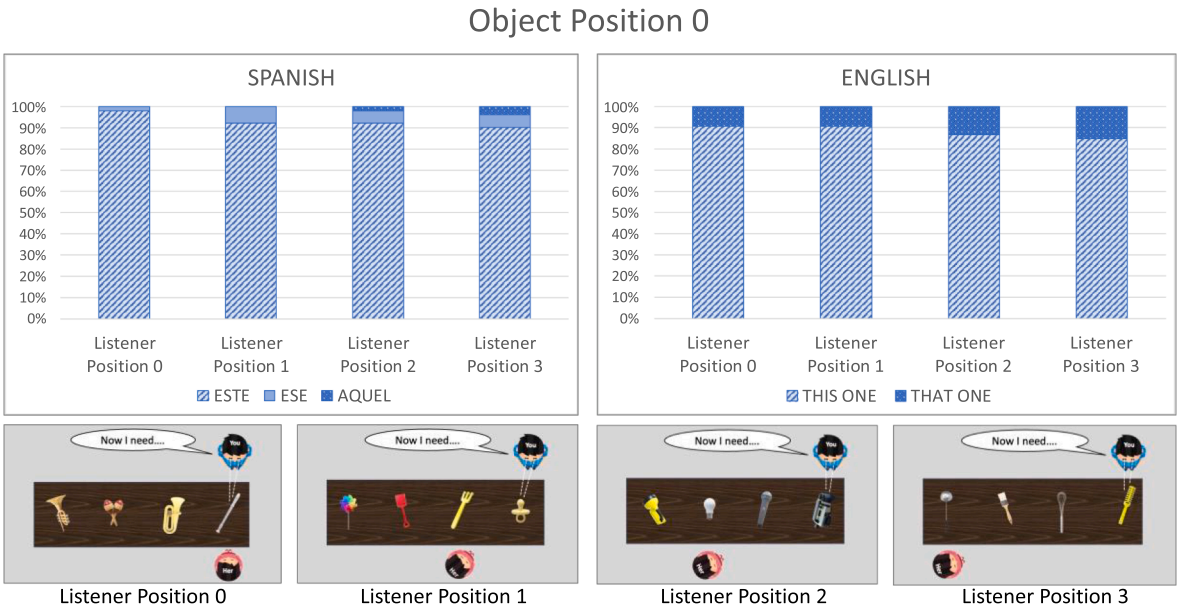

Fig. 2. Average percentage of demonstrative choice in Spanish and English (Experiment 1) in each listener position when the target object was in Position 0. Displays were shown in the language of test.

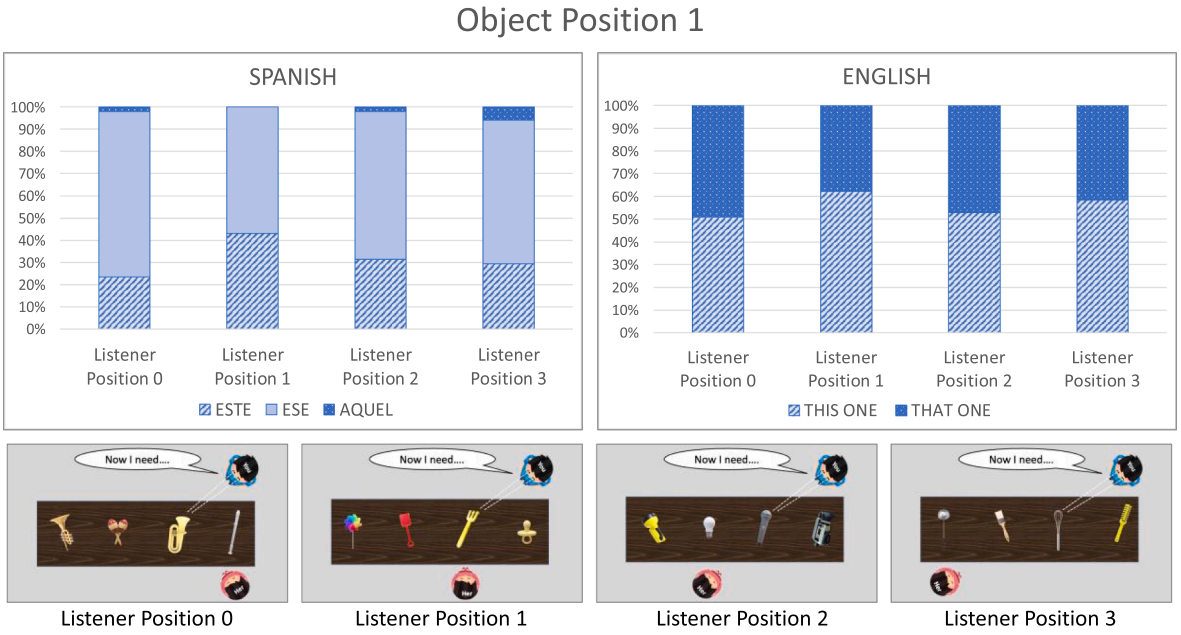

Fig. 3. Average percentage of demonstrative choice in Spanish and English (Experiment 1) in each listener position when the target object was in Position 1. Displays were shown in the language of test.

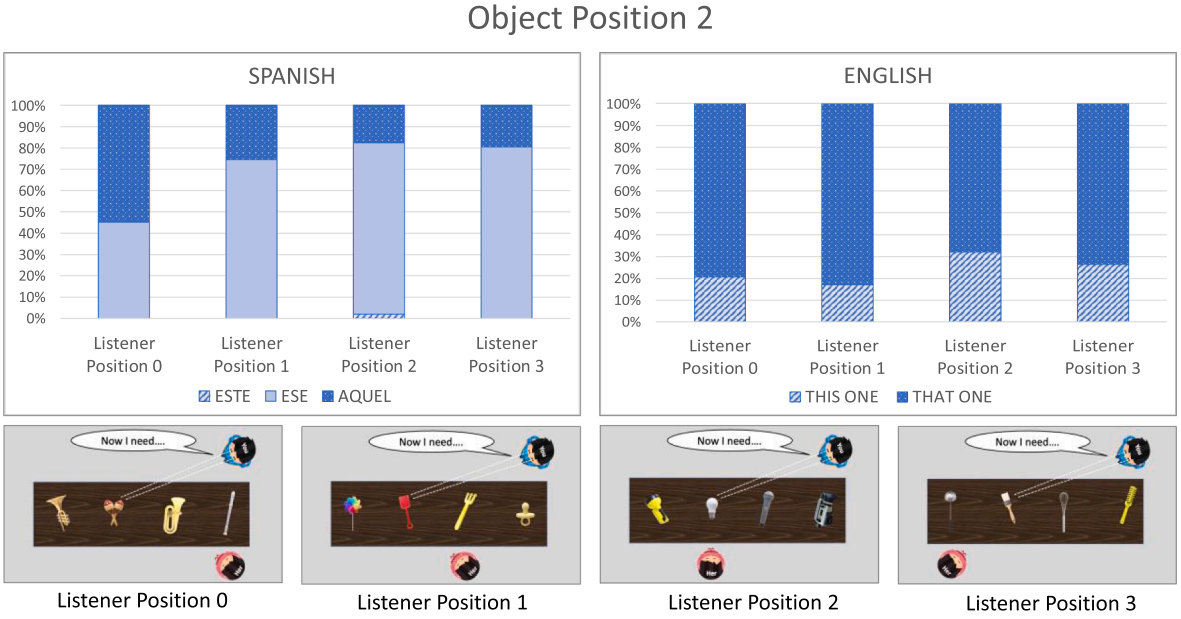

Fig. 4. Average percentage of demonstrative choice in Spanish and English (Experiment 1) in each listener position when the target object was in Position 2. Displays were shown in the language of test.

person-oriented account, Listener Position had a significant effect ($\beta = -1.7612$; SE = 0.5664; $p < 0.0019$). Finally, there was a significant Object Position × Listener Position interaction ($\beta = 1.5716$; SE = 0.4540; $p < 0.0006$).

To follow up on the interaction, Medial Demonstrative choice was modelled separately in Object Positions 1 and 2, with Listener Position (0–3) as a continuous predictor variable, with the Speaker's fixed position as the zero point (see Figs. 3 and 4). These two simpler models included random intercepts for Participants and Items and random by-participant slopes for Listener Position.

The effect of Listener Position was not significant in Object Position 1 ($\beta = -0.1929$; SE = 0.2139; $p = 0.3672$), but was significant in Object Position 2 ($\beta = 1.3548$; SE = 0.3949; $p < 0.0007$), with the choice of the medial demonstrative 'ese' increasing as the listener was closer to the referent.

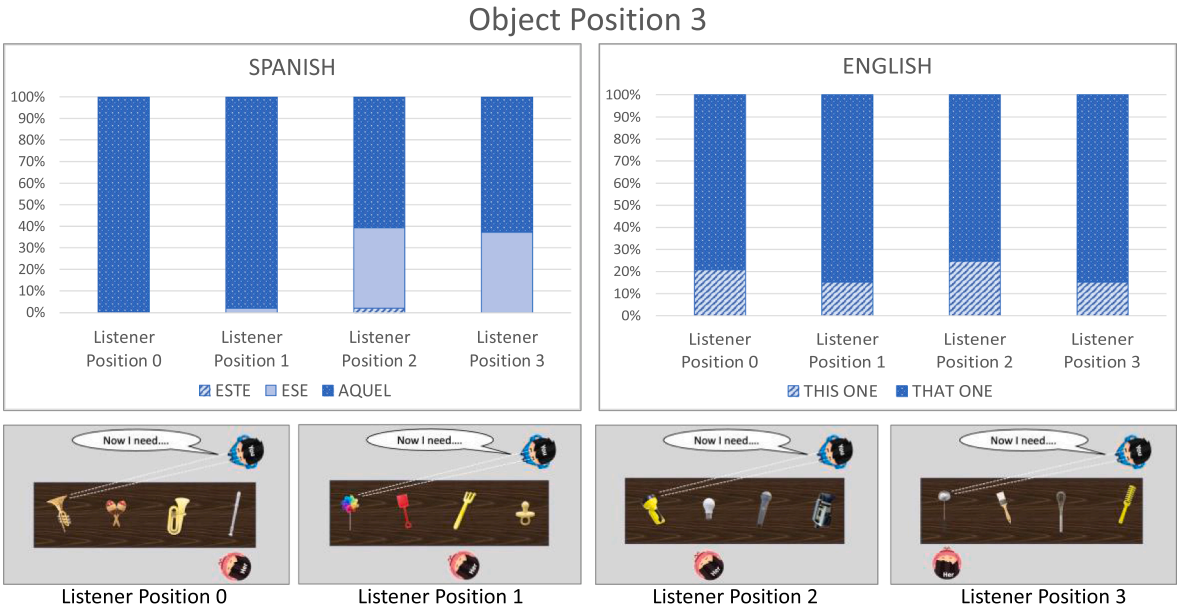

Fig. 5. Average percentage of demonstrative choice in Spanish and English (Experiment 1) in each listener position when the target object was in Position 3. Displays were shown in the language of test.

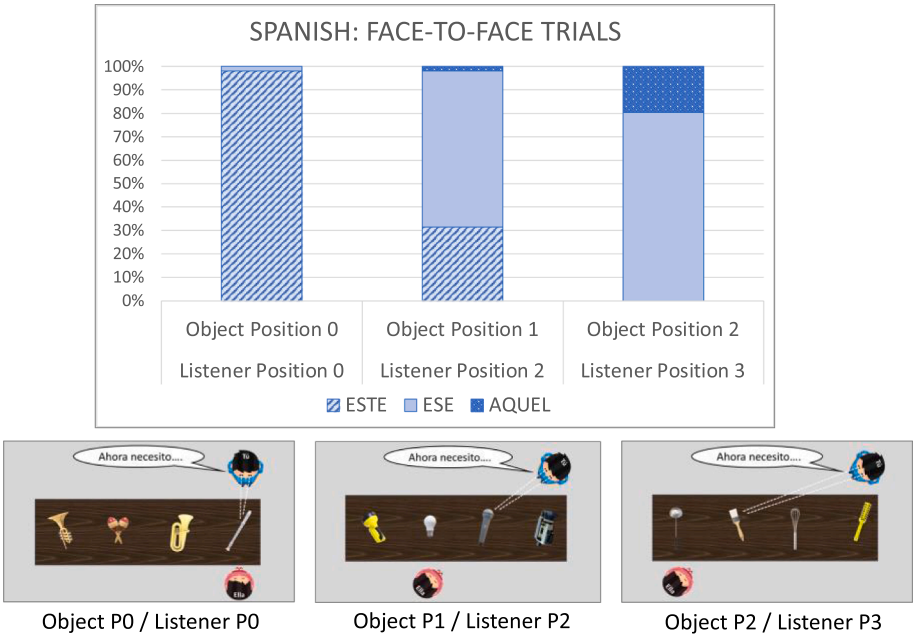

Fig. 6. Average percentage of Spanish demonstrative choice in face-to-face trials (Experiment 1), which were used to test Jungbluth's (2003) account of Spanish demonstrative use inside the conversational dyad.

6.2.1.3. Distal zone (Object Position 3). Using logistic mixed effects regression, the binary outcome variable of Distal Demonstrative choice ('AQUEL' = 1, 'ESTE' and 'ESE' = 0) was modelled for Object Position 3 (or Distal zone), with Listener Position (0–3) as a continuous predictor variable, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position.

Supporting the person-oriented account, Listener Position had a significant effect on participants' use of 'aquel' ($\beta = -1.8277$; $SE = 0.6723$; $p < 0.0066$; see Fig. 5).

6.2.1.4. Analysis of face-to-face trials. Using logistic mixed effects regression, the binary outcome variable of Proximal Demonstrative

choice ('ESTE' = 1, 'ESE' and 'AQUEL' = 0) was modelled for face-to-face trials (Object Position 0 / Listener Position 0, Object Position 1 / Listener Position 2, and Object Position 2 / Listener Position 3), with Object Position as a continuous predictor variable, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Object Position.

Contrary to Jungbluth's (2003) account of Spanish demonstratives, Object Position had a significant effect on the choice of the proximal demonstrative 'este' inside the conversational dyad ($\beta = -13.694$; SE = 6.780; $p < 0.0435$; see Fig. 6).

6.2.2. English data

Since the English demonstrative system only has two terms, analyses did not distinguish between Proximal, Medial and Distal zones. The choice of the proximal demonstrative 'this one' decreased across Object Positions 0–3 (88 % − 56 % − 24 % − 19 %, on average).

Using logistic mixed effects regression, the binary outcome variable of Proximal Demonstrative choice ('THIS ONE' = 1, 'THAT ONE' = 0) was modelled with Object Position (0–3) and Listener Position (0–3) as continuous predictor variables, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position and Object Position plus their interaction.

Replicating the results of Coventry et al. (2008), Object Position had a significant effect on the choice of the proximal demonstrative 'this one' ($\beta = -2.6912$; SE = 0.5328; $p < 0.0001$), which decreased as the object was further away from the speaker. The effect of Listener Position was not significant ($\beta = 0.0489$; SE = 0.2957; $p = 0.869$), and neither was the Object Position × Listener Position interaction ($\beta = 0.0123$; SE = 0.2344; $p = 0.958$).

6.3. Discussion

The results of Experiment 1 replicated and extended the findings of Coventry et al. (2008) using an online demonstrative-choice task: English speakers were sensitive to the location of the target object relative their own position, whereas Spanish speakers were also sensitive to the listener's position. More specifically, the choice of 'ese' in the Medial zone revealed an interaction between Object and Listener Position, such that the proximity of the listener to the target object affected demonstrative choice more strongly in Object Position 2 than in Object Position 1. That is, Listener Position had a stronger effect when participants were choosing between the medial and distal forms ('ese' vs 'aquel') than between the proximal and medial forms ('este' vs 'ese'). Listener Position also had a significant effect in Object Position 3, with the distal form 'aquel' being selected more frequently when the target object was far from both speaker and listener.

In addition to confirming the person-oriented account of Spanish demonstratives (Alonso, 1968; Cifuentes- Honrubia, 1989; Eguren, 1999), the results of the first experiment disconfirmed Jungbluth's (2003) conclusion that Spanish speakers treat everything inside the conversational dyad as proximal, without any further differentiation. Contrary to this prevalent account, the analysis of face-to-face trials revealed a significant effect of Object Position, with the choice of 'este' decreasing with increasing distance of the target object from the speaker's position.

7. Experiment 2

Since English and Spanish have two- and three-way demonstrative systems, respectively, their different sensitivity to the listener's location may be generalizable to similar demonstrative systems. To investigate whether medial and distal demonstratives are always selected in relation to the listener's location (as was observed in Spanish), the same online task was administered to native speakers of another two languages with three-way demonstrative systems: Japanese and Turkish.

Japanese demonstratives form a person-oriented system, with 'kore' indicating proximity to the speaker, 'sore' indicating distance from the speaker but proximity to the listener, and 'are' indicating distance from both speaker and listener (Diessel, 2005; Evans et al., 2018a). Traditional grammars of Turkish have characterized the demonstrative system as distance oriented (e.g., Swift, 1963; Lewis, 1967; Kornfilt, 1997), although other scholars have described it as person oriented, with the proximal form 'bu' referring to entities close to the speaker, the medial form 'şu' referring to entities far from the speaker but close to the listener, and the distal form 'o' referring to entities away from both speaker and listener (Bazin, 1968/1987; Lyons, 1977; see also references in Balpinar, 2019a). As in the case of Spanish, work in linguistic typology has therefore contested whether the speaker is the deictic center in all Turkish demonstratives, or the medial and distal forms also indicate distance from the listener. Experiment 2 employed the same online demonstrative-choice task to empirically test these two accounts of the Turkish demonstrative system, and the person-oriented classification of the Japanese system.

7.1. Methods

7.1.1. Participants

105 participants were recruited through Prolific. N was originally set to 50 in each language group, but five participants timed out (four in the Japanese-speaking group and one in the Turkish-speaking group), so extra participants were automatically recruited and their data was retained for analysis. All participants were compensated for their time.

In the Demographics section of Prolific, first language was set to Japanese and Turkish, in each language group. In addition to this initial screening, two qualifying questions were posed at the start of the task:

(i) The first language you learned, was it Japanese/ Turkish?
   a. Yes
   b. No
(ii) What is your current level of Japanese/ Turkish?
   a. Native, I speak Japanese/ Turkish daily.
   b. I don't speak Japanese/ Turkish well anymore, I don't use it daily.

Only those participants who responded (a) to both questions qualified for the task.

7.1.2. Materials and procedure

The materials and procedure employed in Experiment 1 were used again in Experiment 2. The instructions and materials were translated from English to Turkish and Japanese with the help of native speakers of these languages. Participants in both language groups were given three demonstrative pronouns to choose from ('kore'/ 'sore'/ 'are' in Japanese, and 'buna'/ 'şuna'/ 'ona' in Turkish).

7.2. Results

Descriptive statistics are plotted in Figs. 7-10 (see also the Appendix). As in Experiment 1, data from the two languages were analyzed separately. This is also in line with the overall goal of the study, which is to test typological analyses of different demonstrative systems, rather than directly comparing task performance across languages.

7.2.1. Japanese data

Because Japanese has a three-way demonstrative system, analyses distinguished Proximal, Medial and Distal zones (corresponding with Object Positions 0, 1–2 and 3, respectively). The proximal demonstrative 'kore' was the most frequent choice in the Proximal zone (95 % on average), while the medial demonstrative 'sore' was the preferred option in the Medial zone (68 % on average) and the distal demonstrative 'are' was the most frequent choice in the Distal zone (75 % on average).

7.2.1.1. Proximal zone (Object Position 0). Using logistic mixed effects regression, the binary outcome variable of Proximal Demonstrative choice ('KORE' = 1, 'SORE' and 'ARE' = 0) was modelled for the Proximal zone, with Listener Position (0–3) as a continuous predictor variable, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position.

Listener Position did not have a significant effect on the choice of the proximal demonstrative 'kore' in Object Position 0 ($\beta = 0.6490$; SE = 0.9593; $p = 0.4987$; see Fig. 7).

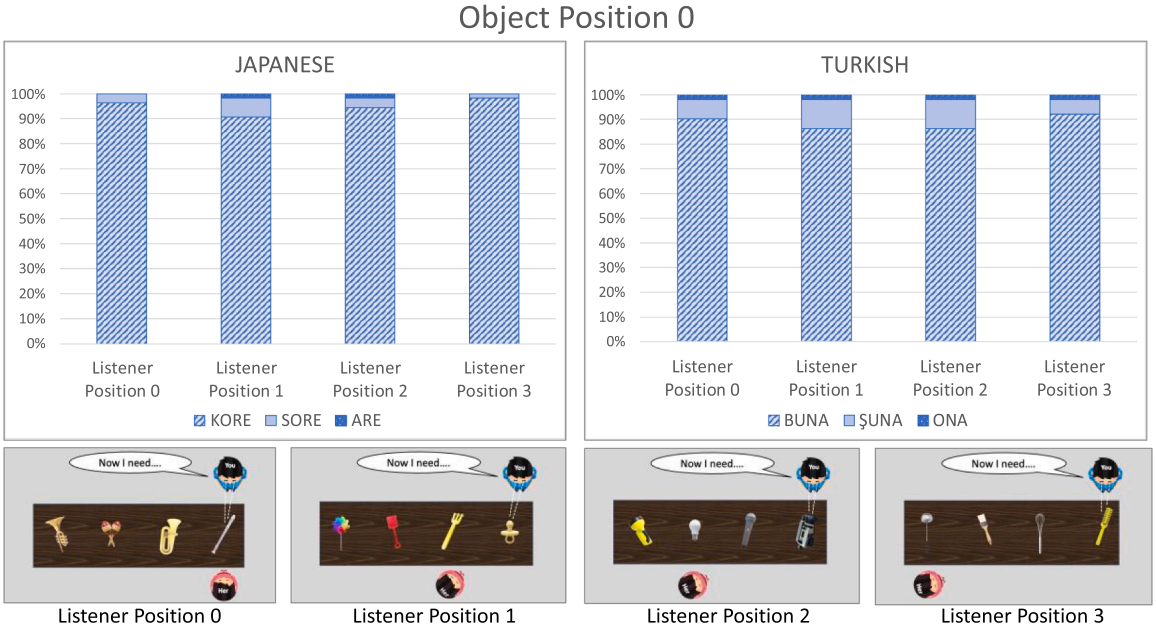

Fig. 7. Average percentage of demonstrative choice in Japanese and Turkish (Experiment 2) in each listener position when the target object was in Position 0. Displays were shown in the language of test.

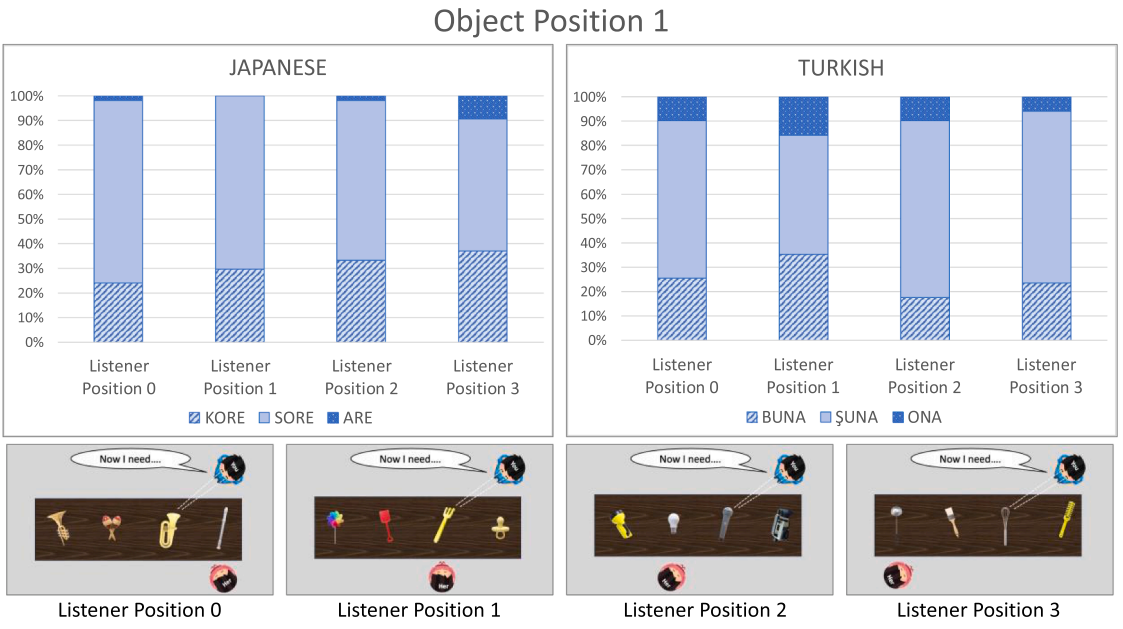

Fig. 8. Average percentage of demonstrative choice in Japanese and Turkish (Experiment 2) in each listener position when the target object was in Position 1. Displays were shown in the language of test.

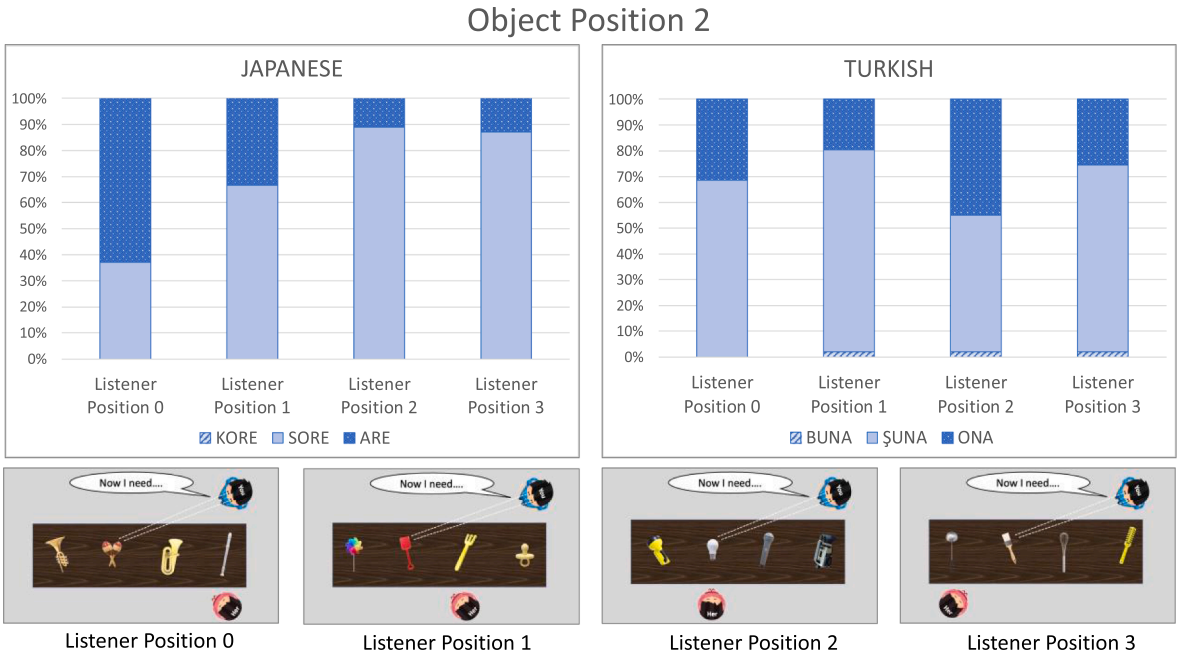

Fig. 9. Average percentage of demonstrative choice in Japanese and Turkish (Experiment 2) in each listener position when the target object was in Position 2. Displays were shown in the language of test.

7.2.1.2. Medial zone (Object Positions 1 and 2). Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('SORE' = 1, 'KORE' and 'ARE' = 0) was modelled in the Medial zone, with Listener Position (0–3) and Object Position (1–2) as continuous predictor variables, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position and Object Position plus their interaction.

Object Position had a significant effect on the choice of the medial demonstrative 'kore' ($\beta = -2.4937$; SE = 0.6975; $p < 0.0004$). Listener Position also had a significant effect ($\beta = -2.9739$; SE = 0.6052; $p < 0.0001$). Finally, there was a significant Object

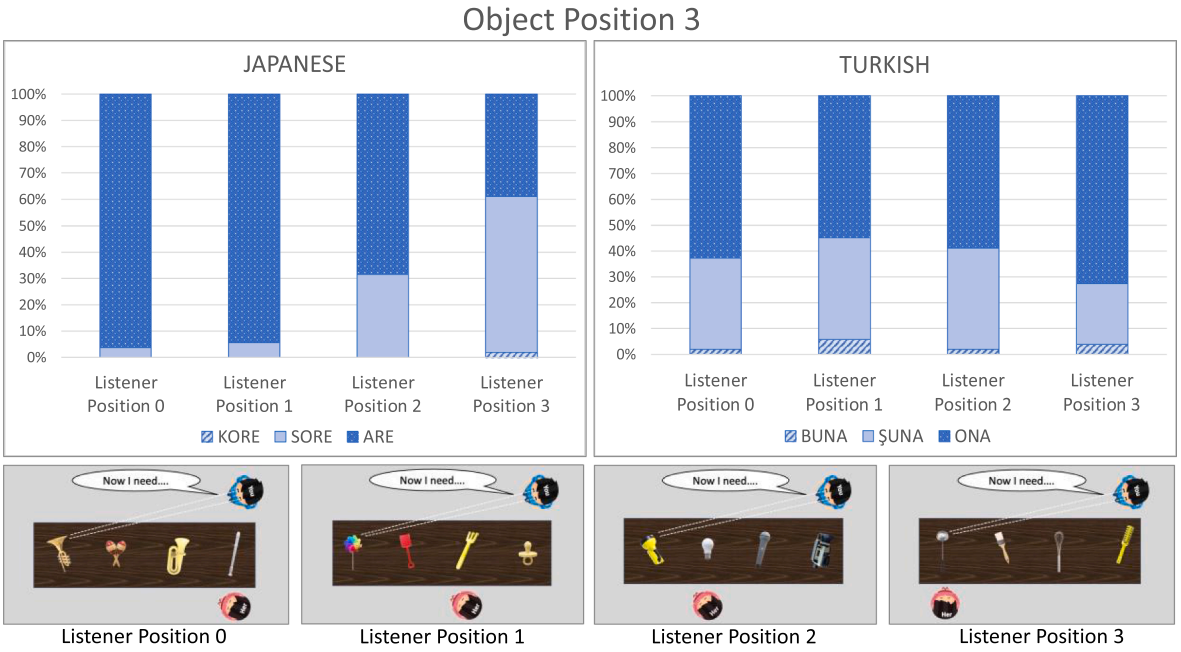

Fig. 10. Average percentage of demonstrative choice in Japanese and Turkish (Experiment 2) in each listener position when the target object was in Position 3. Displays were shown in the language of test.

Position × Listener Position interaction (β = 2.4543; SE = 0.4892; $p < 0.0001$).

To follow up on the interaction, Medial Demonstrative choice was modelled separately in Object Positions 1 and 2, with Listener Position (0–3) as a continuous predictor variable, with the Speaker's fixed position as the zero point (see Figs. 8 and 9). These two simpler models included random intercepts for Participants and Items and random by-participant slopes for Listener Position.

The effect of Listener Position was significant in Object Position 1 (β = -0.4975; SE = 0.2021; $p < 0.0139$), and highly significant in Object Position 2 (β = 1.8801; SE = 0.4450; $p < 0.0001$), with the choice of the medial demonstrative 'kore' increasing as the listener was closer to the referent.

7.2.1.3. Distal zone (Object Position 3). Using logistic mixed effects regression, the binary outcome variable of Distal Demonstrative choice ('ARE' = 1, 'KORE' and 'SORE' = 0) was modelled for the Distal zone, with Listener Position (0–3) as a continuous predictor variable, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position.

Supporting the person-oriented account, Listener Position had a significant effect on participants' use of 'are' (β = -2.5524; SE = 0.6643; $p < 0.0002$; see Fig. 10).

7.2.2. Turkish data

Analyses distinguished Proximal, Medial and Distal zones, in line with the Turkish demonstrative system. The proximal demonstrative 'buna' was the most frequent choice in the Proximal zone (89 % on average), while the medial form 'şuna' was the preferred option in the Medial zone (66 % on average) and the distal demonstrative 'ona' was the most frequent choice in the Distal zone (62 % on average).

7.2.2.1. Proximal zone (Object Position 0). Using logistic mixed effects regression, the binary outcome variable of Proximal Demonstrative choice ('BUNA' = 1, 'ŞUNA' and 'ONA' = 0) was modelled for the Proximal zone, with Listener Position (0–3) as a continuous predictor variable, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position.

Listener Position did not have a significant effect on the choice of the proximal demonstrative 'buna' in Object Position 0 (β = -0.3267; SE = 0.9257; $p = 0.7241$; see Fig. 7).

7.2.2.2. Medial zone (Object Positions 1 and 2). Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('ŞUNA' = 1, 'BUNA' and 'ONA' = 0) was modelled in the Medial zone, with Listener Position (0–3) and Object Position (1–2) as continuous predictor variables, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position and Object Position plus their interaction.

Object Position did not have a significant effect on the choice of the medial demonstrative 'şuna' (β = 0.7614; SE = 0.7539;

$p = 0.313$), neither did Listener Position ($\beta = 0.6624$; SE = 0.5734; $p = 0.248$). Finally, the Object Position × Listener Position interaction was also non-significant ($\beta = -0.3886$; SE = 0.3635; $p = 0.285$; see Figs. 8-9).

7.2.2.3. Distal zone (Object Position 3). Using logistic mixed effects regression, the binary outcome variable of Distal Demonstrative choice ('ONA' = 1, 'BUNA' and 'ŞUNA' = 0) was modelled for the Distal zone, with Listener Position (0–3) as a continuous predictor variable, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position.

Contrary to the person-oriented account, Listener Position did not have a significant effect on participants' use of 'ona' ($\beta = 0.1182$; SE = 0.2549; $p = 0.643$; see Fig. 10).

7.3. Discussion

The results of Experiment 2 confirmed that the Japanese demonstrative system is person oriented, with the medial and distal forms encoding distance from the speaker but proximity to the listener, and distance from both interlocutors, respectively. Interestingly, the results of the Medial zone revealed an Object Position × Listener Position interaction analogous to the one observed in Spanish, where the effect of Listener Position was stronger in Object Position 2 than in Object Position 1. Therefore, both Japanese and Spanish speakers show greater sensitivity to the listener's position when choosing between medial and distal forms in far-away space than when choosing between medial and proximal forms in closer space.

The analyses of the Turkish data confirmed that its demonstrative system is distance oriented, with the choice of the medial and distal forms not being affected by the listener position (unlike in Japanese). The results of Experiment 2 therefore support the characterization of the Turkish demonstrative system found in traditional grammars (e.g., Swift, 1963; Lewis, 1967; Kornfilt, 1997), according to which the speaker is always the deictic center in Turkish demonstratives.

8. Experiment 3

Given the clear differences in sensitivity to listener position observed in Experiments 1 and 2 between person-oriented and distance-oriented demonstrative systems, Experiment 3 investigated the effect of speaking two languages with different demonstrative systems. To this end, Catalan-Spanish bilinguals were administered the same online task in one of their two languages. While Spanish has a three-way person-oriented system (as the results of Experiment 1 confirmed), Catalan has a two-way distance-oriented system, after it lost the medial demonstrative through diachronic change (Badia i Margarit, 1981; Perez Saldanya, 2015; Todisco et al., 2021a). It follows from the hypothesis that the acquisition and mature use of demonstratives trains perspective taking across the lifespan (Rubio-Fernandez, 2020), that speaking two languages with different demonstrative systems may affect speakers' sensitivity to their listener's position. In the case of Catalan-Spanish bilinguals, their use of the Spanish medial form 'ese' may be less sensitive to the listener's position as a result of Catalan having lost the medial form 'aqueix'. This hypothesis was tested in Experiment 3.

8.1. Methods

8.1.1. Participants

102 participants were recruited through Prolific. N was originally set to 50 in each language group, but two Catalan-speaking participants timed out, so extra participants were automatically recruited and their data was retained for analysis. All participants were compensated for their time.

The task was advertised for Catalan native speakers from the regions of Catalonia and the Balearic Islands. In the Demographics section of Prolific, nationality was set to Spanish and fluent languages was set to Catalan. In addition to this initial screening, three qualifying questions were posed at the start of the task:

(i) Did you learn Catalan when you were young? (before the age of 12)
   a. Yes
   b. No

(ii) What is your current level of Catalan?
   a. Native, I speak Catalan daily.
   b. I don't speak Catalan well anymore, I don't use it daily.

(iii) In the area of Catalonia and the Balearic Islands, they use two demonstratives ('aquest' / 'aquell'), whereas in Valencia they use three ('est' / 'eixe' / 'aquell'). What variety of Catalan do you speak?
   a. I use 'aquest' / 'aquell'.
   b. I use 'est'/ 'eixe'/ 'aquell'.

Only those participants who responded (a) to the three questions qualified for the task. It was further confirmed through an open-box demographics question that all qualifying participants were originally from either Catalonia or the Balearic Islands.

8.1.2. Materials and procedure

The materials and procedure employed in Experiments 1 and 2 were used again in Experiment 3. The instructions and materials were translated to Catalan with the help of native speakers of the language. Participants in the Spanish group were given three demonstrative pronouns to choose from ('este'/ 'ese'/ 'aquel'), and in the Catalan group they were given two ('aquest'/ 'aquell').

8.2. Results

Descriptive statistics are plotted in Figs. 11-14 (see also the Appendix). As in Experiments 1 and 2, and in line with the overall goal of the study, data from the two languages were analyzed separately.

8.2.1. Spanish data (Catalan bilinguals)

Analyses distinguished Proximal, Medial and Distal zones (corresponding with Object Positions 0, 1–2 and 3, respectively). The proximal demonstrative 'este' was the most frequent choice in the Proximal zone (99.5 % on average), while the medial demonstrative 'ese' was the preferred option in the Medial zone (63 % on average). The distal demonstrative 'aquel' was the most frequent choice in the Distal zone (84 % on average).

8.2.1.1. Proximal zone (Object Position 0). Using logistic mixed effects regression, the binary outcome variable of Proximal Demonstrative choice ('ESTE' = 1, 'ESE' and 'AQUEL' = 0) was modelled for the Proximal zone, with Listener Position (0–3) as a continuous predictor variable, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position.

Listener Position did not have a significant effect on the choice of the proximal demonstrative 'este' in Object Position 0 ($\beta = -20.63$; SE = 219.52; $p = 0.925$; see Fig. 11).

8.2.1.2. Medial zone (Object Positions 1 and 2). Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('ESE' = 1, 'ESTE' and 'AQUEL' = 0) was modelled in the Medial zone, with Listener Position (0–3) and Object Position (1–2) as continuous predictor variables, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position and Object Position plus their interaction.

Object Position had a significant effect on the choice of the medial demonstrative 'ese' ($\beta = 1.8484$; SE = 0.8098; $p < 0.0225$), but Listener Position did not ($\beta = -0.1435$; SE = 0.4774; $p = 0.7638$). The Object Position × Listener Position interaction was not significant either ($\beta = 0.2147$; SE = 0.3761; $p = 0.5680$; see Figs. 12-13).

8.2.1.3. Distal zone (Object Position 3). Using logistic mixed effects regression, the binary outcome variable of Distal Demonstrative choice ('AQUEL' = 1, 'ESTE' and 'ESE' = 0) was modelled for the Distal zone, with Listener Position (0–3) as a continuous predictor variable, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and

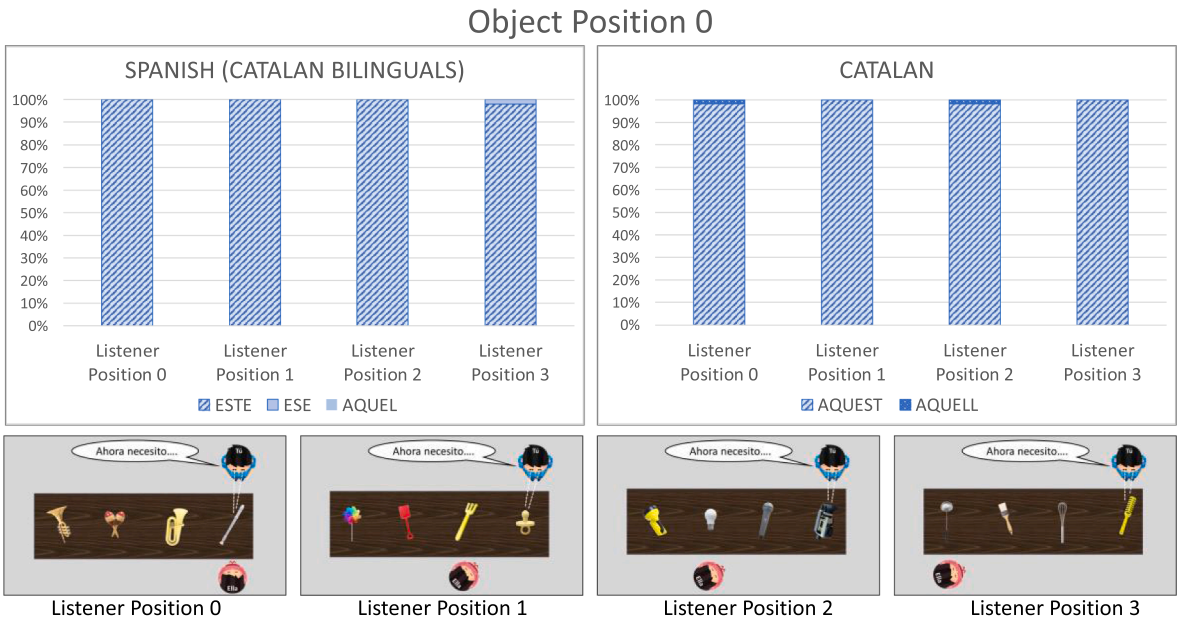

Fig. 11. Average percentage of demonstrative choice in Spanish (Catalan bilinguals sample) and Catalan (Experiment 3) in each listener position when the target object was in Position 0. Displays were shown in the language of test.

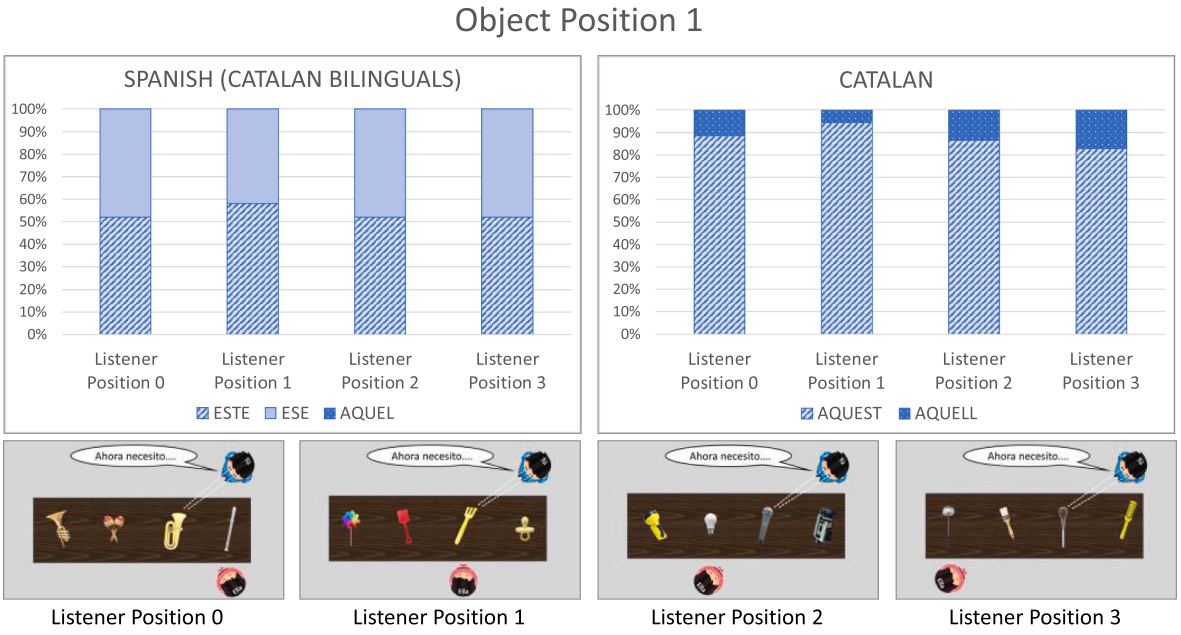

Fig. 12. Average percentage of demonstrative choice in Spanish (Catalan bilinguals sample) and Catalan (Experiment 3) in each listener position when the target object was in Position 1. Displays were shown in the language of test.

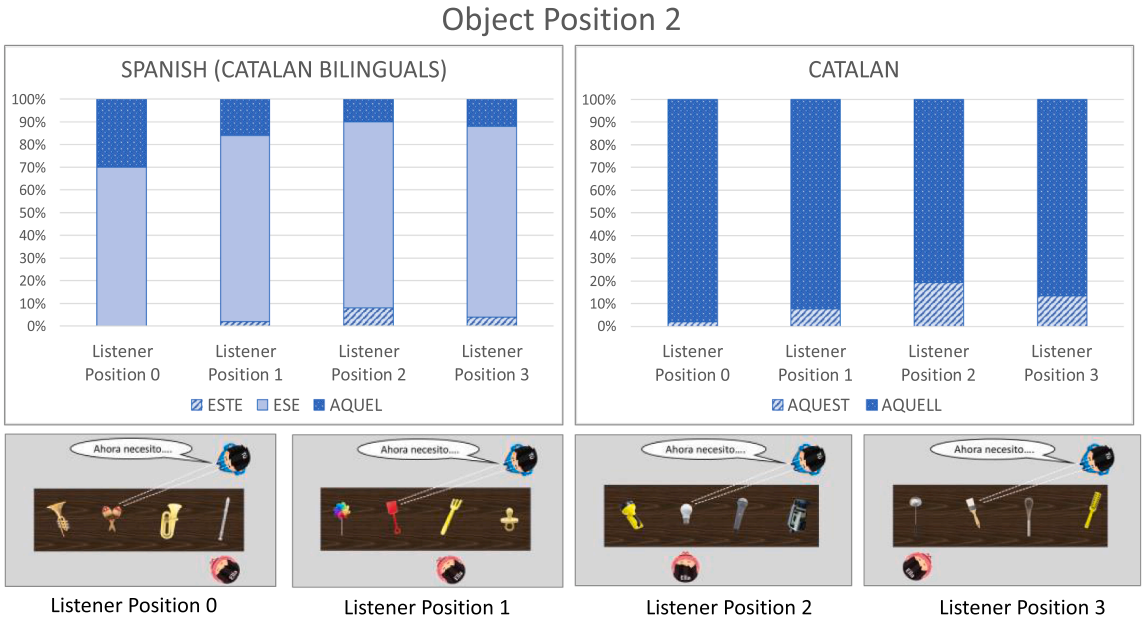

Fig. 13. Average percentage of demonstrative choice in Spanish (Catalan bilinguals sample) and Catalan (Experiment 3) in each listener position when the target object was in Position 2. Displays were shown in the language of test.

random by-participant slopes for Listener Position. Because of model convergence issues, in this model the random intercept and random slope effects for Participants were uncorrelated.

Listener Position had a significant effect on participants' use of 'aquel' ($\beta = -0.9335$; SE = 0.4689; $p < 0.0466$; see Fig. 14).

8.2.2. Catalan data

Since the Catalan demonstrative system only has two terms, analyses did not distinguish between Proximal, Medial and Distal zones. The choice of the proximal demonstrative 'aquest' decreased across Object Positions 0–3 (99 % – 88 % – 11 % – 4 %, on average).

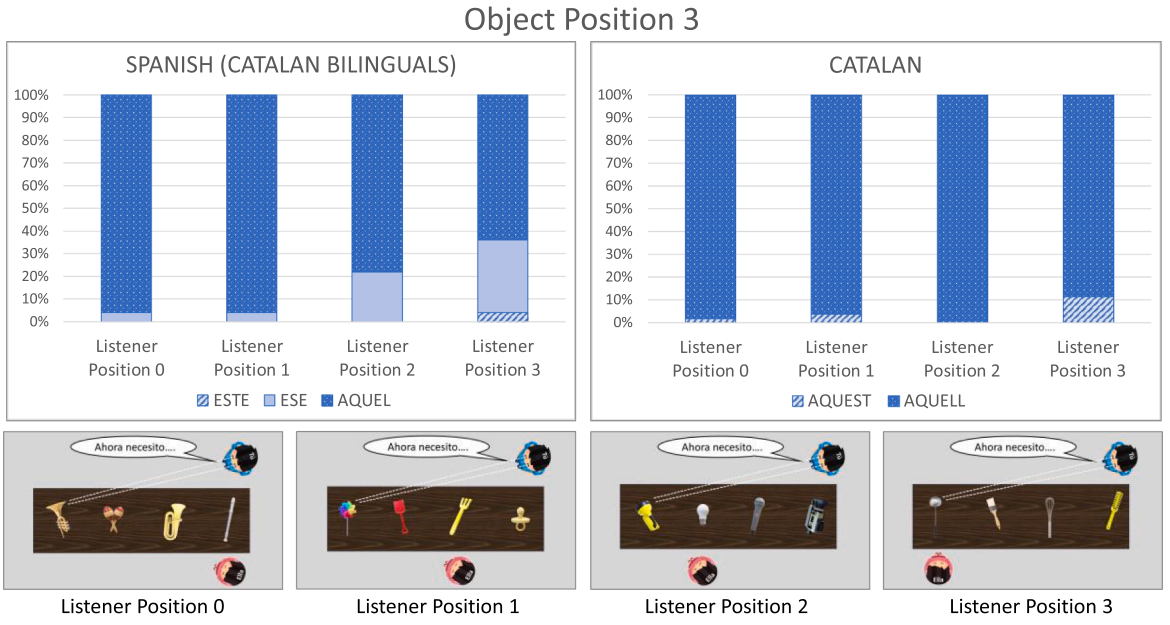

Fig. 14. Average percentage of demonstrative choice in Spanish (Catalan bilinguals sample) and Catalan (Experiment 3) in each listener position when the target object was in Position 3. Displays were shown in the language of test.

Using logistic mixed effects regression, the binary outcome variable of Proximal Demonstrative choice ('AQUEST' = 1, 'AQUELL' = 0) was modelled with Object Position (0–3) and Listener Position (0–3) as continuous predictor variables, with the Speaker's fixed position as the zero point. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Position and Object Position plus their interaction.

Object Position had a significant effect on the choice of the proximal demonstrative 'aquest' ($\beta = -8.3148$; SE = 1.9563; $p < 0.0001$), which decreased as the object was further away from the speaker. The effect of Listener Position was not significant ($\beta = -0.8046$; SE = 1.5669; $p = 0.608$), and neither was the Object Position × Listener Position interaction ($\beta = 0.6180$; SE = 0.9868; $p = 0.531$; see Figs. 11-14).

8.3. Discussion

The analyses of the Catalan data revealed a significant effect of speaker-referent distance on demonstrative choice, but no main effect of or interaction with Listener Position. This pattern of results is therefore analogous to the one observed in English in Experiment 1, confirming that the Catalan demonstrative system is distance oriented.

The analyses of the Spanish data from Catalan-Spanish bilinguals did not reveal a significant effect of Listener Position, nor a significant interaction with Object position in the choice of 'ese' in the Medial Zone. However, the effect of Listener Position was significant in Object Position 3, where participants showed a preference for 'aquel' when the referent was far away from both interlocutors. In comparison with the results observed with monolingual Spanish speakers in Experiment 1, the results of Experiment 3 were therefore mixed: while both samples revealed sensitivity to the listener's position in their choice of the distal demonstrative 'aquel', only the monolingual Spanish speakers were sensitive to their listener's position in their choice of the medial form 'ese'.

Another pattern that was different between the two Spanish-speaking groups was the choice of 'ese' in Object Position 1: while monolingual Spanish speakers selected the medial form 66 % of the time (with Japanese speakers also using 'sore' 66 % of the time in Object Position 1 in Experiment 2), bilingual Catalan-Spanish speakers selected 'ese' 46 % of the time in the same object position. This pattern of demonstrative choice may be the result of a transfer effect where both the lesser use of 'ese' and the lack of sensitivity to the listener's position in the Medial Zone result from the Catalan demonstrative system not having a medial form and being distance oriented.

9. Experiment 4

In an early grammar of Turkish, Underhill (1976) maintains that the use of the medial demonstrative 'şu' is associated with the use of a deictic gesture, while the use of 'bu' and 'o' is not. Bastuji (1976) argues that 'bu' and 'şu' refer to entities within the *space of communication*, where 'şu' has an additional emphatic component, which 'bu' lacks. Hayasi (1985, 1988, 1989, reviewed in Balpinar, 2019a, 2019b; Hayasi & Özsoy, 2015) argues that in Turkish, the proximal and distal demonstratives are used when the referent has already been introduced into the discourse, whereas 'şu' is used for new discourse referents. Based on naturalistic observations during multiparty interaction, Özyürek (1998) proposed an analysis of Turkish demonstatives where the medial form 'şu' is used when the

listener's visual attention is not on the intended referent, while the proximal form 'bu' and the distal form 'o' are used when the referent is already in joint attention.

Küntay and Özyürek (2006; Küntay, 2012; Küntay et al., 2014) refer to an unpublished study by Özyürek and Kita (2000), where they propose a parallel analysis of the Turkish medial demonstrative 'şu' and the Japanese medial demonstrative 'so' (see also Levinson, 2004). Experiment 4 tested these theoretical accounts of the Turkish and Japanese medial demonstratives, and extended the investigation to Spanish as a control.

In addition to these typological analyses, Experiment 4 tested the hypothesis that demonstratives may be used for attention correction in all languages, even if they do not have a specific demonstrative form for this function. For example, a Spanish speaker may refer to an object in Position 1 as 'ese' (i.e. using the medial form) when their listener is looking at that object, but they may refer to the same object as 'este' (i.e. using the proximal form) if their listener is looking at Position 3 (i.e. past the location of the target; gloss: *Look over here!*). Confirming this hypothesis would support the view that attention monitoring is fundamental to demonstrative use across languages, and does not depend on having a specific demonstrative for attention correction (as it has been argued for Turkish and Japanese; for related analyses of Yucatec Mayan, see Bohnemeyer (2018); Hanks (2009)).

9.1. Methods

9.1.1. Participants

A new group of 101 native speakers of Spanish and Turkish were recruited through Prolific. N was originally set to 50 in each language group, but one Spanish-speaking participant timed out and an extra participant was automatically recruited and their data was retained for analysis. All participants were compensated for their time.

In the Demographics section of Prolific, first language was set to Spanish and Turkish, in each language group. In addition to this initial screening, two qualifying questions were posed at the start of the task:

(iii) The first language you learned, was it Spanish/ Turkish?
  a. Yes
  b. No
(iv) What is your current level of Spanish/ Turkish?
  a. Native, I speak Spanish/ Turkish daily.
  b. I don't speak Spanish/ Turkish well anymore, I don't use it daily.

Only those participants who responded (a) to both questions qualified for the task.

The Japanese sample was not recruited through Prolific because that platform has a very limited pool of Japanese native speakers, and it was not possible to run a second experiment with new participants. Instead, the link for the task was distributed amongst Japanese native speakers who were studying at Tohoku University in Sendai (Japan). N was also set to 50, but an additional student performed the task before the link was disabled.

9.1.2. Materials and procedure

The visual materials and instructions were similar to those in Experiments 1–3 ($n = 18$). There were only three differences in the displays: the position of the listener was always parallel to the target object (i.e. object and listener positions were not crossed). Object Position ranged across Positions 1–3 on the table (i.e. the target object did not appear in Position 0). In addition, both speaker and listener were looking at one of the objects on the table, as determined by both their body orientation and two intermittent white lines representing their line of gaze. In half the trials, the speaker and listener were looking at the target object (Aligned Perspectives condition), whereas in the other half, they were looking at different objects (Misaligned Perspectives condition).

Target position was counterbalanced across trials (6 trials per position). The listener and the target object did not appear in Position 0 (the speaker's position) because the proximal demonstratives 'kore', 'este' and 'buna' (equivalent to 'this one' in English) would be used in that position with both aligned and misaligned perspectives, weakening any evidence of flexible demonstrative use for attention correction. On the other hand, the use of the 3 demonstratives would be more flexible in Positions 1–3, which made for a better test of attention correction. Participants in all language groups were given three demonstrative pronouns to choose from ('kore'/ 'sore'/ 'are' in Japanese, 'este'/ 'ese'/ 'aquel' in Spanish, and 'buna'/ 'şuna'/ 'ona' in Turkish).

The instructions were the same as in previous experiments with one difference: participants were told that their friend would sometimes be looking at the wrong object. However, it was stressed that the target object was always the one the participant/speaker was looking at, and that was the object they needed to request. The instructions therefore treated the listener's focus of attention as irrelevant to the speaker's request.

9.2. Results

Descriptive statistics are plotted in Figs. 15-17 (see also the Appendix). As in previous experiments and in line with the overall goal of the study, data from the three languages were analyzed separately.

Two analyses were performed at each object position. First, to test the hypothesis that the Japanese and Turkish medial demonstratives are used when the speaker and listener perspectives are misaligned, the choice of this form was compared in the Aligned and Misaligned Perspectives conditions at each object position. Second, to test the hypothesis that proximal and distal demonstratives

may be used flexibly depending on the direction in which the listener is looking, the results of the Misaligned condition were analyzed as follows: proximal demonstrative choice was compared when the listener was looking closer vs further from the target in Object Position 1, and distal demonstrative choice was compared when the listener was looking closer vs further from the target in Object Position 2. This comparison could not be performed in Object Position 3 because in all Misaligned Perspectives trials, the listener was looking closer from the target (see Fig. 17).

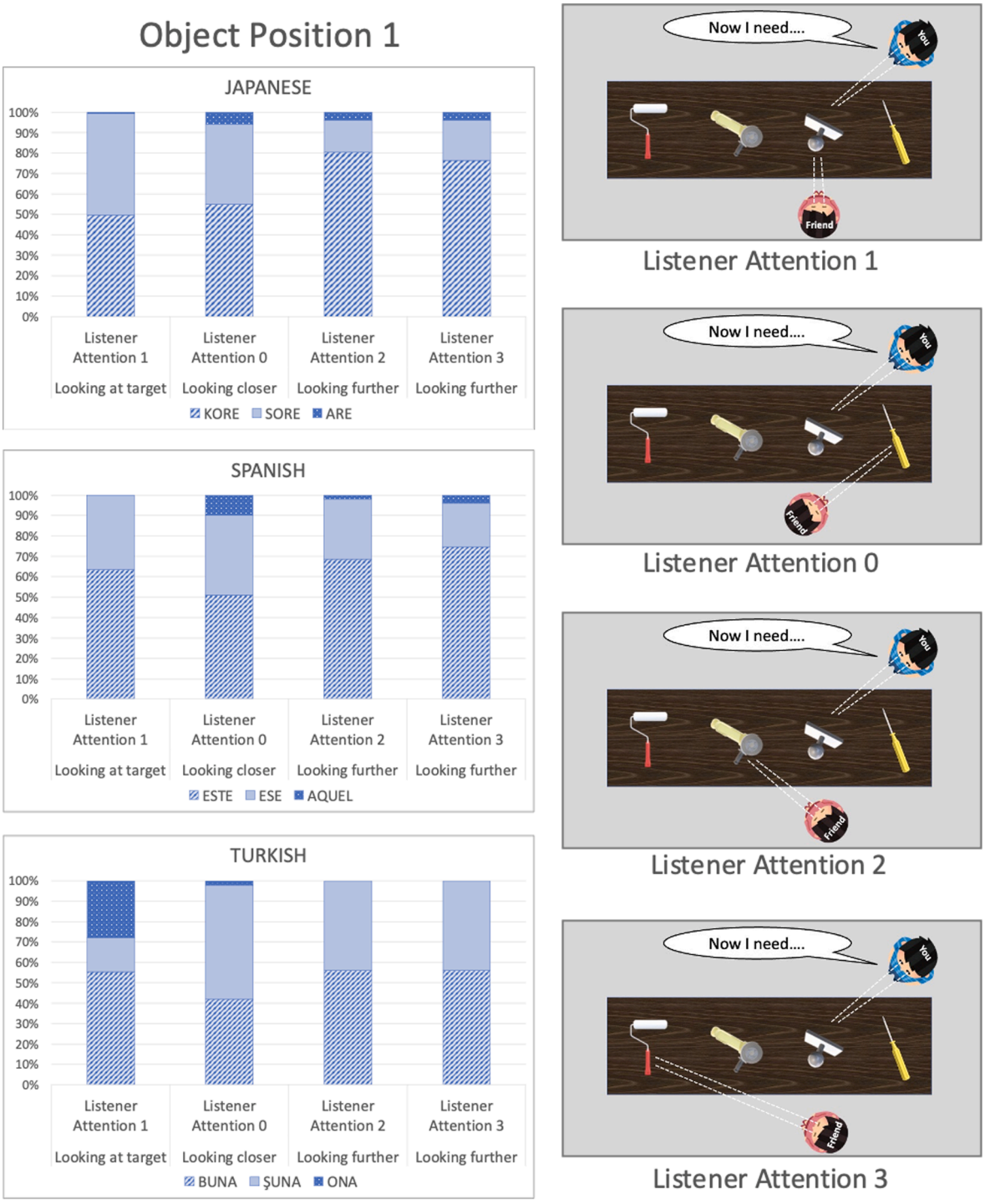

Fig. 15. Average percentage of demonstrative choice in Japanese, Spanish and Turkish (Experiment 4) in each listener-attention position when the target object was in Position 1. Displays were shown in the language of test.

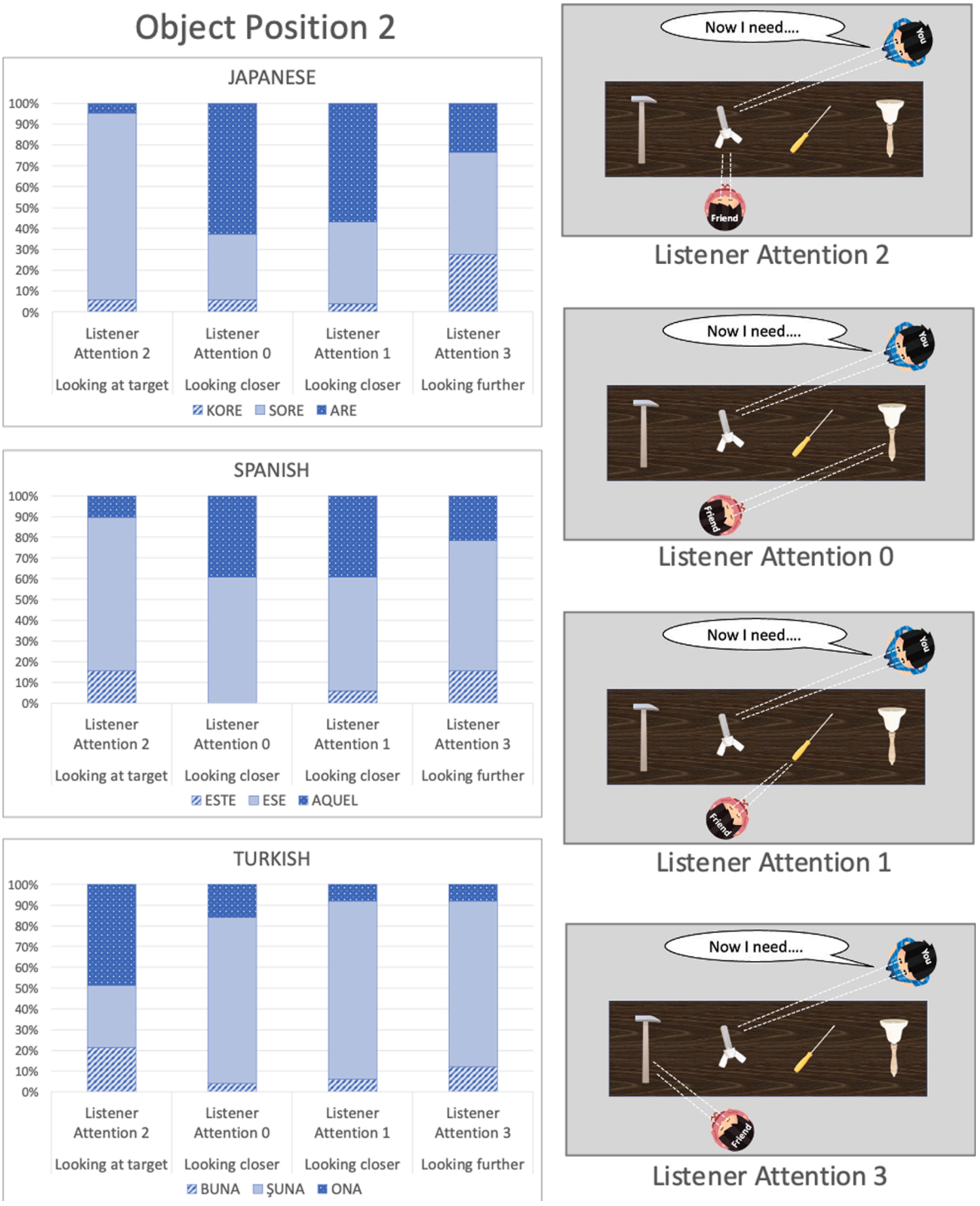

Fig. 16. Average percentage of demonstrative choice in Japanese, Spanish and Turkish (Experiment 4) in each listener-attention position when the target object was in Position 2. Displays were shown in the language of test.

9.2.1. Japanese data

The proximal demonstrative 'kore' was the most frequent choice in Object Position 1 (65 % on average), while the medial demonstrative 'sore' was the preferred option in Object Position 2 (52 % on average) and the distal demonstrative 'are' was the most frequent choice in Object Position 3 (63 % on average).

Regarding the use of the medial demonstrative to correct the listener's attention, 'sore' was selected 72 % of the time in the Aligned

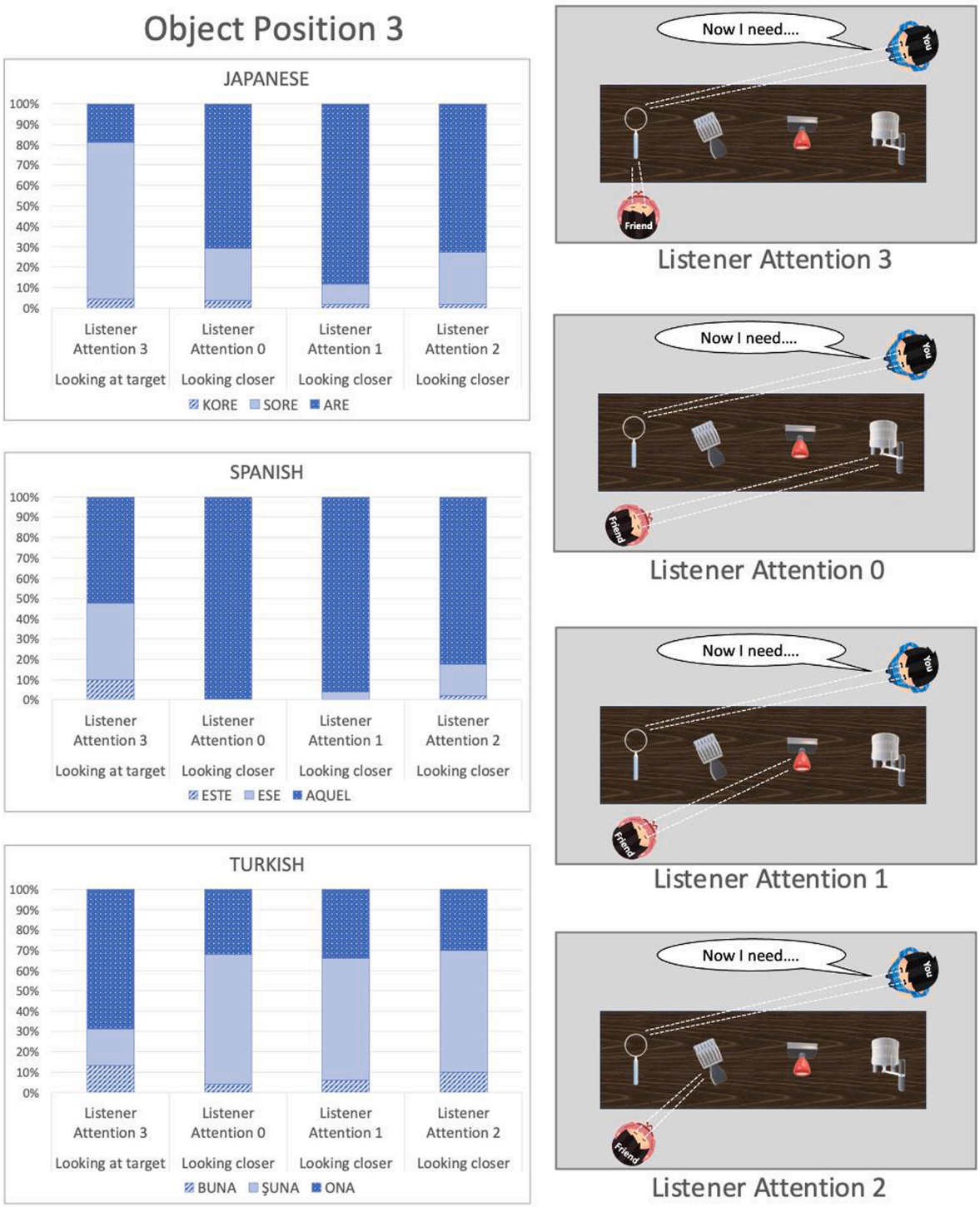

Fig. 17. Average percentage of demonstrative choice in Japanese, Spanish and Turkish (Experiment 4) in each listener-attention position when the target object was in Position 3. Displays were shown in the language of test.

Perspectives condition, and only 28 % of the time in the Misaligned Perspectives condition (on average). These percentages run counter to analyses comparing Japanese and Turkish for using their middle demonstrative for attention correction.

9.2.1.1. Object Position 1. Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice

('SORE' = 1, 'KORE' and 'ARE' = 0) was modelled for Object Position 1, with Perspectives (Aligned vs Misaligned) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Perspectives.

Perspectives had a significant effect on the choice of the medial demonstrative in Object Position 1 ($\beta = -1.6844$; SE = 0.6632; $p < 0.0112$), with 'sore' being selected more frequently in the Aligned Perspectives condition than in the Misaligned Perspectives condition (see Fig. 15).

Using logistic mixed effects regression, the binary outcome variable of Proximal Demonstrative choice ('KORE' = 1, 'SORE' and 'ARE' = 0) in the Misaligned Perspectives condition was modelled for Object Position 1, with Listener Gaze Direction (Closer vs Further) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Gaze Direction.

Listener Gaze Direction had a significant effect on the choice of 'kore' in Object Position 1 ($\beta = 3.6159$; SE = 1.7069; $p < 0.0342$), with the proximal demonstrative being selected more often when the listener was looking past the target object (see Fig. 15).

9.2.1.2. Object Position 2. Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('SORE' = 1, 'KORE' and 'ARE' = 0) was modelled for Object Position 2, with Perspectives (Aligned vs Misaligned) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Perspectives.

Perspectives had a significant effect on the choice of the medial demonstrative in Object Position 2 ($\beta = -8.783$; SE = 1.751; $p < 0.0001$), with 'sore' being selected more frequently in the Aligned Perspectives condition than in the Misaligned Perspectives condition (see Fig. 16).

Using logistic mixed effects regression, the binary outcome variable of Distal Demonstrative choice ('ARE' = 1, 'KORE' and 'SORE' = 0) in the Misaligned Perspectives condition was modelled for Object Position 2, with Listener Gaze Direction (Closer vs Further) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Gaze Direction.

Listener Gaze Direction had a significant effect on the choice of 'are' in Object Position 2 ($\beta = -2.0261$; SE = 0.9322; $p < 0.0298$), with the distal demonstrative being selected more often when the listener was looking closer from the target position (see Fig. 16).

9.2.1.3. Object Position 3. Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('SORE' = 1, 'KORE' and 'ARE' = 0) was modelled for Object Position 3, with Perspectives (Aligned vs Misaligned) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Perspectives.

Perspectives had a significant effect on the choice of the medial demonstrative in Object Position 3 ($\beta = -9.031$; SE = 1.926; $p < 0.0001$), with 'sore' being selected more frequently in the Aligned Perspectives condition than in the Misaligned Perspectives condition (see Fig. 17).

In Object Position 3, the listener was looking closer from the target position in all the Misaligned Perspectives trials (see Fig. 17), so the second analysis could not be performed in this object position.

9.2.2. Spanish data

The proximal demonstrative 'este' was the most frequent choice in Object Position 1 (64 % on average), while the medial demonstrative 'ese' was the preferred option in Object Position 2 (63 % on average) and the distal demonstrative 'aquel' was the most frequent choice in Object Position 3 (83 % on average).

Regarding the use of the medial demonstrative to correct the listener's attention, 'ese' was selected 49 % of the time in the Aligned Perspectives condition and only 32 % of the time in the Misaligned Perspectives condition.

9.2.2.1. Object Position 1. Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('ESE' = 1, 'ESTE' and 'AQUEL' = 0) was modelled for Object Position 1, with Perspectives (Aligned vs Misaligned) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Perspectives.

Perspectives did not have a significant effect on the choice of the medial demonstrative in Object Position 1 ($\beta = -0.3684$; SE = 0.7526; $p = 0.6245$; see Fig. 15).

Using logistic mixed effects regression, the binary outcome variable of Proximal Demonstrative choice ('ESTE' = 1, 'ESE' and 'AQUEL' = 0) in the Misaligned Perspectives condition was modelled for Object Position 1, with Listener Gaze Direction (Closer vs Further) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Gaze Direction.

Listener Gaze Direction had a significant effect on the choice of 'este' in Object Position 1 ($\beta = 3.17024$; SE = 1.06553; $p < 0.00294$), with the proximal demonstrative being selected more often when the listener was looking past the target object (see Fig. 15).

9.2.2.2. Object Position 2. Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('ESE' = 1, 'ESTE' and 'AQUEL' = 0) was modelled for Object Position 2, with Perspectives (Aligned vs Misaligned) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Perspectives.

Perspectives did not have a significant effect on the choice of the medial demonstrative in Object Position 2 ($\beta = -0.9303$; SE = 0.8455; $p = 0.2712$; see Fig. 16).

Using logistic mixed effects regression, the binary outcome variable of Distal Demonstrative choice ('AQUEL' = 1, 'ESTE' and 'ESE'

= 0) in the Misaligned Perspectives condition was modelled for Object Position 2, with Listener Gaze Direction (Closer vs Further) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Gaze Direction.

Listener Gaze Direction had a significant effect on the choice of 'aquel' in Object Position 2 ($\beta = -17.1449$; SE = 6.1500; $p < 0.00532$), with the distal demonstrative being selected more often when the listener was looking closer from the target object position (see Fig. 16).

9.2.2.3. Object Position 3. Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('ESE' = 1, 'ESTE' and 'AQUEL' = 0) was modelled for Object Position 3, with Perspectives (Aligned vs Misaligned) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Perspectives.

Perspectives had a significant effect on the choice of 'ese' in Object Position 3 ($\beta = -31.964$; SE = 11.535; $p < 0.00560$), with the medial demonstrative being selected more frequently in the Aligned Perspectives condition than in the Misaligned Perspectives condition (see Fig. 17).

In Object Position 3, the listener was looking closer from the target position in all the Misaligned Perspectives trials (see Fig. 17), so the second analysis could not be performed in this object position.

9.2.3. Turkish data

The proximal demonstrative 'buna' was the most frequent choice in Object Position 1 (52 % on average), while the medial form 'şuna' was the preferred option in Object Position 2 (69 % on average) and Object Position 3 (51 % on average).

Regarding the use of the medial demonstrative to correct the listener's attention, 'şuna' was selected 22 % of the time in the Aligned Perspectives condition and 64 % of the time in the Misaligned Perspectives condition, revealing the reverse pattern to the one observed in Japanese and Spanish.

9.2.3.1. Object Position 1. Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('ŞUNA' = 1, 'BUNA' and 'ONA' = 0) was modelled for Object Position 1, with Perspectives (Aligned vs Misaligned) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Perspectives.

Perspectives had a significant effect on the choice of 'şuna' in Object Position 1 ($\beta = 6.047$; SE = 1.953; $p < 0.00197$), with the medial demonstrative being selected more often in the Misaligned Perspectives condition than in the Aligned Perspectives condition (see Fig. 15).

Using logistic mixed effects regression, the binary outcome variable of Proximal Demonstrative choice ('BUNA' = 1, 'ŞUNA' and 'ONA' = 0) in the Misaligned Perspectives condition was modelled for Object Position 1, with Listener Gaze Direction (Closer vs Further) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Listener Gaze Direction.

Listener Gaze Direction did not have a significant effect on the choice of the proximal demonstrative in Object Position 1 ($\beta = -0.9370$; SE = 1.6298; $p = 0.565$; see Fig. 15).

9.2.3.2. Object Position 2. Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('ŞUNA' = 1, 'BUNA' and 'ONA' = 0) was modelled for Object Position 2, with Perspectives (Aligned vs Misaligned) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Perspectives.

Perspectives had a significant effect on the choice of the medial demonstrative in Object Position 2 ($\beta = 4.0819$; SE = 0.9247; $p < 0.0001$), with 'şuna' being selected more often in the Misaligned Perspectives condition than in the Aligned Perspectives condition (see Fig. 16).

Using logistic mixed effects regression, the binary outcome variable of Distal Demonstrative choice ('ONA' = 1, 'ŞUNA' and 'BUNA' = 0) in the Misaligned Perspectives condition was modelled for Object Position 2, with Listener Gaze Direction (Closer vs Further) as a predictor variable. This model included random intercepts for Participants and Items. Because of convergence issues, this model did not include random by-participant slopes for Listener Gaze Direction.

Listener Gaze Direction did not have a significant effect on the choice of the distal demonstrative in Object Position 2 ($\beta = -0.9504$; SE = 0.9905; $p = 0.33730$; see Fig. 16).

9.2.3.3. Object Position 3. Using logistic mixed effects regression, the binary outcome variable of Medial Demonstrative choice ('ŞUNA' = 1, 'BUNA' and 'ONA' = 0) was modelled for Object Position 3, with Perspectives (Aligned vs Misaligned) as a predictor variable. This model included random intercepts for Participants and Items and random by-participant slopes for Perspectives.

Perspectives had a significant effect on the choice of 'şuna' in Object Position 3 ($\beta = 6.428$; SE = 2.646; $p < 0.0152$), with the medial demonstrative being selected more often in the Misaligned Perspectives condition than in the Aligned Perspectives condition (see Fig. 17).

Once again, in Object Position 3, the listener was looking closer from the target position in all the Misaligned Perspectives trials, so the second analysis could not be performed in this object position.

9.3. Discussion

The results of Experiment 4 supported a number of typological analyses compatible with Özyürek's characterization of the Turkish medial demonstrative 'şu' as the form that is used to align the interlocutors' perspectives on the intended referent (Underhill, 1976; Bastuji, 1976; Hayasi, 1985, 1988, 1989; Özyürek, 1998; Küntay & Özyürek, 2006; Hayasi & Özsoy, 2015). In all three object positions, the medial demonstrative was the preferred option in the Misaligned Perspectives condition. In addition, the proximal and distal forms were not used flexibly in the Misaligned condition, depending on whether the listener was looking closer or further from the target object. These patterns of results suggest that the Turkish medial demonstrative has been lexicalized as the directive used for attention correction.

Contrary to the expectation that Japanese would reveal similar results to Turkish, demonstrative choice in Japanese was closer to the patterns observed in Spanish. Speakers of both languages selected the medial demonstrative more often in the Aligned Perspectives condition than in the Misaligned perspectives condition (with this difference reaching significance in Object Positions 1–3 in Japanese and in Object Position 3 in Spanish). This preference is therefore the *reverse* of the one observed in Turkish. In addition, both Japanese and Spanish speakers selected their proximal demonstrative more often when the listener was looking past the target object (gloss: *Look over here!*) than when she was looking closer. Reversely, their choice of the distal form was more frequent when the listener was looking closer from the target (gloss: *Look over there!*). This flexible use of the proximal and distal demonstratives depending on where the listener is looking suggests that in both Japanese and Spanish, there is not a specific demonstrative which is used by default to reorient the listener towards the intended referent – unlike in Turkish.

10. General discussion

The goal of this study was to experimentally test different typological analyses of various demonstrative systems as a way to explore their socio-cognitive demands. While the use of online methods for language elicitation cannot replace interactive experimentation and naturalistic observations, the results of the demonstrative-choice task employed in this study revealed sufficient cross-linguistic variation to arbitrate between different typological analyses of various demonstrative systems. Further supporting the reliability of this online paradigm, Experiment 1 replicated and extended the results of a lab-based experiment by Coventry et al. (2008), which had revealed addressee effects in demonstrative use in Spanish, but could not determine the exact nature of this demonstrative system since referent and listener position did not vary parametrically.

Here the results of Experiment 1 offered strong support to the person-oriented view of Spanish demonstratives (Alonso, 1968; Cifuentes- Honrubia, 1989; Eguren, 1999), with the proximal form 'este' indicating proximity to the speaker, the medial form 'ese' indicating distance from the speaker but proximity to the listener, and the distal form 'aquel' indicating distance from both the speaker and the listener. Unlike Spanish speakers, the English-speaking participants in Experiment 1 did not show sensitivity to listener position in their demonstrative choice, confirming that the English system is distance oriented.

The results of Experiment 1 also confirmed that Spanish speakers draw spatial distinctions in face-to-face communication, contrary to Jungbluth's (2003) claim that everything inside the conversational dyad is treated as proximal. As discussed in the introduction (see Table 1), Jungbluth's conclusion was possibly a result of her not distinguishing the use of proximal demonstratives as deictic expressions vs fillers (Kouteva et al., 2019). This would explain the inconsistent results of Jungbluth (2003) vs Coventry et al. (2008), as Peeters et al. (2021) recently pointed out (see also Shin et al. (2020) for inconsistent results).

The findings of Experiment 2 confirmed that not all three-way demonstrative systems are sensitive to listener position. Japanese speakers showed a preference for the medial form for referents far from the speaker but close to the listener, while they selected the distal form for referents far from both interlocutors. Turkish speakers, on the other hand, did not show sensitivity to listener position in their choice of the medial and distal forms, suggesting that their demonstrative system is distance oriented.

The results of Experiment 3 showed that the Catalan demonstrative system is distance oriented, revealing analogous results to English. Interestingly, Catalan-Spanish bilinguals revealed a different pattern of results when tested in Spanish relative to the Spanish monolinguals in Experiment 1, only showing sensitivity to listener position in their choice of the distal form 'aquel' but not in their choice of the medial form 'ese'. In addition, Catalan-Spanish bilinguals selected the medial form less frequently than Spanish monolinguals in Object Position 1. The different results observed with Spanish monolinguals and Catalan-Spanish bilinguals were interpreted as a transfer effect from Catalan, which lost the middle form in diachrony (Badia i Margarit, 1981; Perez Saldanya, 2015; Todisco et al., 2021a). Thus, monolingual Spanish speakers use the medial demonstrative 'ese' to indicate distance from the speaker and proximity to the listener, while Catalan-Spanish bilinguals use this form less often, and to indicate distance from the speaker – but not proximity to the listener.

Finally, the results of Experiment 4 confirmed that Turkish speakers use the medial demonstrative 'şu' to redirect the listener's attention to the intended referent, regardless of whether the listener is looking closer or further from the target (Underhill, 1976; Bastuji, 1976; Hayasi, 1985, 1988, 1989; Özyürek, 1998; Küntay & Özyürek, 2006; Hayasi & Özsoy, 2015). By contrast, Japanese and Spanish speakers chose flexibly between the proximal and distal demonstratives when the interlocutors' perspectives were misaligned, showing a preference for the proximal demonstrative when the listener was looking further (gloss: *Look over here!*) and for the distal demonstrative when the listener was looking closer (gloss: *Look over there!*).

Overall, the results of this study confirm that different demonstrative systems pose different demands on social cognition, resulting in cross-linguistic differences in speakers' sensitivity to the listener's spatial location and attention focus (Rubio-Fernandez, 2020; see also Evans et al., 2018a, 2018b). Importantly, other experimental studies using interactive tasks have revealed how both spatial and social factors affect demonstrative use across languages (e.g., Coventry et al., 2008; Piwek et al., 2008; Shin et al., 2020; Skilton &

Peeters, 2021). Future studies should therefore try to replicate the results of this study using more naturalistic methods of demonstrative elicitation in real-time interaction.

10.1. On the egocentric vs interactive debate on demonstratives

An important question that has somewhat polarized recent cross-linguistic research on demonstratives is whether speakers' demonstrative choice is based on their own peripersonal space (i.e. what is within the speaker's reach), or is interactive in nature (i.e. relative to both interlocutors and their ongoing activity; for arguments and empirical evidence supporting these two positions, see Coventry et al., 2008; Piwek et al., 2008; Diessel, 2014; Peeters et al., 2014, 2015; Peeters & Özyürek, 2016; Rocca et al., 2018, 2019a, 2019b; Caldano & Coventry, 2019; Diessel & Coventry, 2020; Reile et al., 2020; Shin et al., 2020; Stukenbrock, 2020; Skilton & Peeters, 2021; Todisco et al., 2021c).

The results of Experiment 1 revealed that listener position affected demonstrative choice in Spanish, but only when the referent was two or three positions away from the speaker. This pattern of results supports Coventry et al.'s (2008) conclusion that there is a 'basic perceptual distinction between near and far space' (p.895), with listener position only having an effect in the more distal zones. However, once the referent is beyond the peripersonal space of the speaker, the interactive nature of demonstratives is clear in the Spanish system. Interestingly, the same pattern of results was observed in Japanese in Experiment 2, with listener position affecting demonstrative choice more strongly when the choice was between the medial and distal forms, than when it was between the proximal and medial forms. The results of Experiments 1 and 2 therefore support the social view of demonstrative choice as interactive in nature (Peeters et al., 2014, 2021; Peeters & Özyürek, 2016; Shin et al., 2020), while also revealing a fundamental distinction between near and far space from the speaker's position (Coventry et al., 2008).

Seen this way, the results of the present study are in line with the findings of Shin et al. (2020), who used an interactive task with native speakers of Spanish and observed that both spatial and interactive factors affected their demonstrative use. Importantly, neither type of factor fully determined demonstrative use in Spanish, with the results of this recent study failing to support purely spatial analyses of demonstratives (e.g., Diessel, 1999, 2014) but also purely interactive accounts (e.g., the idea that proximal demonstratives indicate shared space; Jungbluth, 2003; Peeters & Özyürek, 2016). In conclusion, even if typological analyses of demonstratives have received some 'bad press' for relying on reference grammars and linguistic intuitions (Peeters et al., 2015; see also Section 4 above), the results of Shin et al. (2020) as well as the present study confirm the validity and value of such semantic analyses and their compatibility with pragmatic factors observed in interaction.

10.2. Implications for the relationship between language and social cognition

The experiments in this study investigated two socio-cognitive abilities that may be recruited in demonstrative use: namely, monitoring the listener's spatial location and monitoring the listener's focus of attention. The results of Experiments 1–3 showed clear cross-linguistic differences in demonstrative choice depending on the listener's distance to the referent. These cross-linguistic differences reveal what Slobin (1996) called *thinking for speaking*: speakers automatically monitor those features of the environment that are encoded in their grammars (Papafragou et al., 2008; Boroditsky, 2011; Wolff & Holmes, 2011).

For instance, if a demonstrative system signals the visibility of a referent for the listener, speakers of that language will monitor their listener's visual perspective in order to accurately use their demonstratives. By contrast, if a demonstrative system signals distance to the speaker, using that demonstrative system will not require monitoring the listener's visual perspective. The results with Catalan-Spanish bilinguals in Experiment 3 are particularly revealing in this respect, showing that speaking two languages with different demonstrative systems can reduce speakers' sensitivity to listener position in the language with a person-oriented system.

Regarding speakers' monitoring of the listener's attention, the results of Experiment 4 showed two different strategies to redirect a listener towards the intended referent: Turkish seems to have lexicalized their medial demonstrative for attention correction (for related studies of Yucatec Mayan, see Bohnemeyer (2018); Hanks (2009)), whereas Japanese and Spanish speakers make flexible use of their proximal and distal demonstratives depending on where the listener is looking. Since all languages have demonstratives (Levinson, 2018) and their directive function is also a universal (Diessel, 1999a, 1999b), speakers of all languages should be sensitive to their listeners' focus of attention in their use of demonstratives, regardless of the specific strategy used in their language.

Overall, the results of this study suggest that investigating the connection between language and social cognition requires studying both universals and cross-linguistic differences in reference systems (for extensive discussion, see Rubio-Fernandez, 2020). Thus, to better understand the positive effect of language on social cognition, I propose to investigate how these two cognitive capacities are *jointly used* in communication, focusing on those specific forms of social cognition that are deployed in verbal interaction. This methodological approach differs markedly from previous studies on the effect of language on thought, which traditionally employed non-verbal tasks with speakers of different languages (for a review, see Enfield, 2015). The artificial nature of those tasks, however, has been argued to undermine the depth and even the validity of the effects of language on thought that these tasks revealed (Gentner & Goldin-Meadow, 2003). I therefore propose to investigate how universals and cross-linguistic differences in reference systems may recruit and train communicative social cognition, leaving future studies to test the effect of language on social cognition outside of communication.

To conclude, the results of this study confirm that the cross-linguistic study of the acquisition and mature use of demonstrative systems offers a window into the deep connection between language and social cognition.

Open practice statement

Data and analysis script are available at the Open Science Framework repository (https://osf.io/tbrzj/). The experiment was not

preregistered.

Declaration of Competing Interest

The author declares that she has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data and analysis files are available at an OSF repository (link provided in manuscript).

Acknowledgements

This research was supported by a FRIPRO grant (#275505) from the Research Council of Norway. The author gratefully acknowledges this funding. Access to Qualtrics was granted by MIT through a Research Affiliation with the Brain and Cognitive Sciences Department. Special thanks to Nazli Altinok, Lluís Barcelo i Coblijn, Aurora Bell, Toni Gomila, Chigusa Kurumada, Louise McNally and Satomi Sugaya for their valuable input in understanding the various demonstrative systems investigated in the study and for help preparing the linguistic materials. Thanks also to Swastika Jajoo for her help with participant recruitment at Tohoku University (Japan).

Appendix. Mean proportion of demonstrative choice in Experiments 1–4

Exp.	Language	Object position	Listener position	Proximal	Medial	Distal
1	English	1	1	90.6	N.A.	9.4
1	English	1	2	90.6	N.A.	9.4
1	English	1	3	86.8	N.A.	13.2
1	English	1	4	84.9	N.A.	15.1
1	English	2	1	50.9	N.A.	49.1
1	English	2	2	62.3	N.A.	37.7
1	English	2	3	52.8	N.A.	47.2
1	English	2	4	58.5	N.A.	41.5
1	English	3	1	20.8	N.A.	79.2
1	English	3	2	17.0	N.A.	83.0
1	English	3	3	32.1	N.A.	67.9
1	English	3	4	26.4	N.A.	73.6
1	English	4	1	20.8	N.A.	79.2
1	English	4	2	15.1	N.A.	84.9
1	English	4	3	24.5	N.A.	75.5
1	English	4	4	15.1	N.A.	84.9
1	Spanish (monolingual)	1	1	98.0	2.0	0.0
1	Spanish (monolingual)	1	2	92.2	7.8	0.0
1	Spanish (monolingual)	1	3	92.2	5.9	2.0
1	Spanish (monolingual)	1	4	90.2	5.9	3.9
1	Spanish (monolingual)	2	1	23.5	74.5	2.0
1	Spanish (monolingual)	2	2	43.1	56.9	0.0
1	Spanish (monolingual)	2	3	31.4	66.7	2.0
1	Spanish (monolingual)	2	4	29.4	64.7	5.9
1	Spanish (monolingual)	3	1	0.0	45.1	54.9
1	Spanish (monolingual)	3	2	0.0	74.5	25.5
1	Spanish (monolingual)	3	3	2.0	80.4	17.6
1	Spanish (monolingual)	3	4	0.0	80.4	19.6
1	Spanish (monolingual)	4	1	0.0	0.0	100.0
1	Spanish (monolingual)	4	2	0.0	2.0	98.0
1	Spanish (monolingual)	4	3	2.0	37.3	60.8
1	Spanish (monolingual)	4	4	0.0	37.3	62.7
2	Japanese	1	1	96.3	3.7	0.0
2	Japanese	1	2	90.7	7.4	1.9
2	Japanese	1	3	94.4	3.7	1.9
2	Japanese	1	4	98.1	1.9	0.0
2	Japanese	2	1	24.1	74.1	1.9
2	Japanese	2	2	29.6	70.4	0.0
2	Japanese	2	3	33.3	64.8	1.9
2	Japanese	2	4	37.0	53.7	9.3
2	Japanese	3	1	0.0	37.0	63.0
2	Japanese	3	2	0.0	66.7	33.3

(*continued on next page*)

(*continued*)

Exp.	Language	Object position	Listener position	Proximal	Medial	Distal
2	Japanese	3	3	0.0	88.9	11.1
2	Japanese	3	4	0.0	87.0	13.0
2	Japanese	4	1	0.0	3.7	96.3
2	Japanese	4	2	0.0	5.6	94.4
2	Japanese	4	3	0.0	31.5	68.5
2	Japanese	4	4	1.9	59.3	38.9
2	Turkish	1	1	90.2	7.8	2.0
2	Turkish	1	2	86.3	11.8	2.0
2	Turkish	1	3	86.3	11.8	2.0
2	Turkish	1	4	92.2	5.9	2.0
2	Turkish	2	1	25.5	64.7	9.8
2	Turkish	2	2	35.3	49.0	15.7
2	Turkish	2	3	17.6	72.5	9.8
2	Turkish	2	4	23.5	70.6	5.9
2	Turkish	3	1	0.0	68.6	31.4
2	Turkish	3	2	2.0	78.4	19.6
2	Turkish	3	3	2.0	52.9	45.1
2	Turkish	3	4	2.0	72.5	25.5
2	Turkish	4	1	2.0	35.3	62.7
2	Turkish	4	2	5.9	39.2	54.9
2	Turkish	4	3	2.0	39.2	58.8
2	Turkish	4	4	3.9	23.5	72.5
3	Catalan	1	1	98.1	N.A.	1.9
3	Catalan	1	2	100.0	N.A.	0.0
3	Catalan	1	3	98.1	N.A.	1.9
3	Catalan	1	4	100.0	N.A.	0.0
3	Catalan	2	1	88.5	N.A.	11.5
3	Catalan	2	2	94.2	N.A.	5.8
3	Catalan	2	3	86.5	N.A.	13.5
3	Catalan	2	4	82.7	N.A.	17.3
3	Catalan	3	1	1.9	N.A.	98.1
3	Catalan	3	2	7.7	N.A.	92.3
3	Catalan	3	3	19.2	N.A.	80.8
3	Catalan	3	4	13.5	N.A.	86.5
3	Catalan	4	1	1.9	N.A.	98.1
3	Catalan	4	2	3.8	N.A.	96.2
3	Catalan	4	3	0.0	N.A.	100.0
3	Catalan	4	4	11.5	N.A.	88.5
3	Spanish (bilingual)	1	1	100.0	0.0	0.0
3	Spanish (bilingual)	1	2	100.0	0.0	0.0
3	Spanish (bilingual)	1	3	100.0	0.0	0.0
3	Spanish (bilingual)	1	4	98.0	2.0	0.0
3	Spanish (bilingual)	2	1	52.0	48.0	0.0
3	Spanish (bilingual)	2	2	58.0	42.0	0.0
3	Spanish (bilingual)	2	3	52.0	48.0	0.0
3	Spanish (bilingual)	2	4	52.0	48.0	0.0
3	Spanish (bilingual)	3	1	0.0	70.0	30.0
3	Spanish (bilingual)	3	2	2.0	82.0	16.0
3	Spanish (bilingual)	3	3	8.0	82.0	10.0
3	Spanish (bilingual)	3	4	4.0	84.0	12.0
3	Spanish (bilingual)	4	1	0.0	4.0	96.0
3	Spanish (bilingual)	4	2	0.0	4.0	96.0
3	Spanish (bilingual)	4	3	0.0	22.0	78.0
3	Spanish (bilingual)	4	4	4.0	32.0	64.0

Exp.	Language	Obj. / List. position	Listener attention	Listener looking	Proximal	Medial	Distal
4	Japanese	2	2	Target	49.7	49.7	0.7
4	Japanese	2	1	Closer	54.9	39.2	5.9
4	Japanese	2	3	Further	80.4	15.7	3.9
4	Japanese	2	4	Further	76.5	19.6	3.9
4	Japanese	3	3	Target	5.9	89.5	4.6
4	Japanese	3	1	Closer	5.9	31.4	62.7
4	Japanese	3	2	Closer	3.9	39.2	56.9
4	Japanese	3	4	Further	27.5	49.0	23.5
4	Japanese	4	4	Target	4.6	76.5	19.0
4	Japanese	4	1	Further	3.9	25.5	70.6
4	Japanese	4	2	Further	2.0	9.8	88.2
4	Japanese	4	3	Further	2.0	25.5	72.5

(*continued on next page*)

(*continued*)

Exp.	Language	Obj. / List. position	Listener attention	Listener looking	Proximal	Medial	Distal
4	Spanish (mono)	2	2	Target	63.4	36.6	0.0
4	Spanish (mono)	2	1	Closer	51.0	39.2	9.8
4	Spanish (mono)	2	3	Further	68.6	29.4	2.0
4	Spanish (mono)	2	4	Further	74.5	21.6	3.9
4	Spanish (mono)	3	3	Target	15.7	73.9	10.5
4	Spanish (mono)	3	1	Closer	0.0	60.8	39.2
4	Spanish (mono)	3	2	Closer	5.9	54.9	39.2
4	Spanish (mono)	3	4	Further	15.7	62.7	21.6
4	Spanish (mono)	4	4	Target	9.8	37.9	52.3
4	Spanish (mono)	4	1	Further	0.0	0.0	100.0
4	Spanish (mono)	4	2	Further	0.0	3.9	96.1
4	Spanish (mono)	4	3	Further	2.0	15.7	82.4
4	Turkish	2	2	Target	55.3	16.7	28.0
4	Turkish	2	1	Closer	42.0	56.0	2.0
4	Turkish	2	3	Further	56.0	44.0	0.0
4	Turkish	2	4	Further	56.0	44.0	0.0
4	Turkish	3	3	Target	21.3	30.0	48.7
4	Turkish	3	1	Closer	4.0	80.0	16.0
4	Turkish	3	2	Closer	6.0	86.0	8.0
4	Turkish	3	4	Further	12.0	80.0	8.0
4	Turkish	4	4	Target	13.3	18.0	68.7
4	Turkish	4	1	Further	4.0	64.0	32.0
4	Turkish	4	2	Further	6.0	60.0	34.0
4	Turkish	4	3	Further	10.0	60.0	30.0

References

Alonso, M. (1968). *Gramatica del español contemporaneo*. Madrid: Guadarrama.

Badia i Margarit, A. (1981). *Gramàtica històrica catalana*. Biblioteca d'Estudis i Investigacions, 4. Benetússer (València): Garffitti XXI.

Balpinar, M. (2019a). Some limitations on the controllability of the referent in Turkish demonstratives (Türkçe işaret sözcüklerinde göndergenin kontrol edilebilirliği üzerindeki bazi kisitlamalar). *Selçuk Üniversitesi Türkiyat Araştırmaları Dergisi, 47*, 73–84.

Balpinar, M. (2019b). *Demonstratives and Grammaticalization: A Perspective from Modern Turkish*. London and New York: Routledge.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.

Bastuji, J. (1976). *Les relations spatiales en Turc contemporain: Etude semantique*. Editions Klincksieck.

Bazin, L. (1968/1987). *Introduction à l'étude pratique de la langue turque (Troisième édition revue et corrigée, 1987)*. Paris: Librairie d'Amérique et d'Orient, Adrien Maisonneuve.

Bergqvist, H., & Knuchel, D. (2019). Explorations of engagement: Introduction. *Open Linguistics, 5*(1), 650–665.

Bohnemeyer, J. (2018). Yucatec demonstratives in interaction: Spontaneous versus elicited data. In S. Levinson, et al. (Eds.), *Demonstratives in cross-linguistic perspective* (pp. 176–205). Cambridge: Cambridge University Press.

Boroditsky, L. (2011). How language shapes thought. *Scientific American, 304*, 62–65.

Caldano, M., & Coventry, K. R. (2019). Spatial demonstratives and perceptual space: To reach or not to reach? *Cognition, 191*, Article 103989. https://doi.org/10.1016/j.cognition.2019.06.001

Cifuentes-Honrubia, J. L. (1989). *Lengua y espacio: Introducción al problema de la deixis en español*. Alicante: Universidad de Alicante.

Clark, H. H., & Tree, J. E. F. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*(1), 73–111.

Cooperrider, K. (2016). The co-organization of demonstratives and pointing gestures. *Discourse Processes, 53*(8), 632–656.

Coventry, K. R., Valdés, B., Castillo, A., & Guijarro-Fuentes, P. (2008). Language within your reach: Near–far perceptual space and spatial demonstratives. *Cognition, 108*(3), 889–895.

Diessel, H. (1999a). The morphosyntax of demonstratives in synchrony and diachrony. *Linguistic Typology, 3*(1).

Diessel, H. (1999b). *Demonstratives. Form, Function, and Grammaticalization*. Amsterdam: John Benjamins.

Diessel, H. (2003). The relationship between demonstratives and interrogatives. *Studies in Language, 27*(3), 635–655.

Diessel, H. (2005). Distance contrasts in demonstratives. In M. Haspelmath, M. Dryer, D. Gil, & B. Comrie (Eds.), *World atlas of language structures* (pp. 170–173). Oxford: Oxford University Press.

Diessel, H. (2006). Demonstratives, joint attention, and the evolution of grammar. *Cognitive Linguistics, 17*, 463–489.

Diessel, H. (2012a). Deixis and demonstratives. In C. Maienborn, K.v. Heusinger, & P. Portner (Eds.), *An international handbook of natural language meaning* (pp. 1–25). Berlin: Mouton de Gruyter.

Diessel, H. (2012b). Buehler's two-field theory of pointing and naming and the deictic origins of grammatical morphemes. In T. Breban, L. Brems, K. Davidse, & T. Mortelmans (Eds.), *New perspectives on grammaticalization: Theoretical understanding and empirical description* (pp. 35–48). Amsterdam: John Benjamins.

Diessel, H. (2014). Demonstratives, frames of reference, and semantic universals of space. *Language and Linguistics Compass, 8*(3), 116–132.

Diessel, H., & Coventry, K. R. (2020). Demonstratives in spatial language and social interaction: An interdisciplinary review. *Frontiers in Psychology, 11*.

Eguren, L. J. (1999). Pronombres y adverbios demonstrativos. Las relaciones deícticas. In I. Bosque, & V. Demonte (Eds.), *Gramática Descriptiva de la Lengua Española* (pp. 929–974). Madrid: Espasa.

Enfield, N. J. (2015). Linguistic relativity from reference to agency. *Annual Review of Anthropology, 44*, 207–224.

Evans, N., Bergqvist, H., & San Roque, L. (2018a). The grammar of engagement I: Framework and initial exemplification. *Language and Cognition, 10*(1), 110–140.

Evans, N., Bergqvist, H., & San Roque, L. (2018b). The grammar of engagement II: Typology and diachrony. *Language and Cognition, 10*(1), 141–170.

Gentner, D., & Goldin-Meadow, S. (2003). *Language in mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press.

Goody, E. (1995). Introduction: Some implications of a social origin of intelligence. In E. Goody (Ed.), *Social Intelligence and Interaction: Expressions and Implications of the Social Bias in Human Intelligence* (pp. 1–33). Cambridge: Cambridge University Press.

Graham, L. A. (2013). Comparing hesitation markers in Sanjuanero Spanish. *Diálago de la Lengua, 5*, 66–77.

Graham, L. A. (2018). Variation in hesitation: The case of *este* vs. *eh* in Latin American Spanish. *Spanish in Context, 15*(1), 1–26.

Hanks, W. F. (2009). Fieldwork on deixis. *Journal of Pragmatics, 41*(1), 10–24.
Hayasi, T. (1985). Turkish demonstratives (Torukogo no shijishi). *Ajia & Afurika Gengo Bunka Kenkyūjo Tsūshin, 53*, 55–57. Cited in Balpinar (2019b).
Hayasi, T. (1988). On Turkish demonstratives. *Tokyo University Linguistic Papers, 88*, 229–238. Cited in Balpinar (2019b).
Hayasi, T. (1989). An invitation to Turkish 3 (Torukogo no susume 3: 'kore-sore-are' arekore). *Gengo, 18*(1), 96–101. Cited in Balpinar (2019b.
Hayasi, T., & Özsoy, A. S. (2015). *Şu* or *bu/o*: Turkish nominal demonstratives with concrete referents. In D. Zeyrek, Ç. S. Şimşek, U. Ataş, & J. Rehbein (Eds.), *Turcologica* (pp. 389–401). Wiesbaden: Harrassowitz Verlag.
Heine, B., Kuteva, T., Long, H., Narrog, H., & Wu, F. (2020). Where do demonstratives come from? *STUF-Language Typology and Universals, 73*(3), 403–434.
Hottenroth, P. M. (1982). The system of local deixis in Spanish. In J. Weissenborn, & W. Klein (Eds.), *Here and There: Crosslinguistic Studies on Deixis and Demonstration* (pp. 133–154). Amsterdam: John Benjamins.
Jungbluth, K. (2003). Deictics in the conversational dyad. In F. Lenz (Ed.), *Deictic conceptualisation of space, time and person* (pp. 13–40). Amsterdam, The Netherlands: John Benjamins Publishing.
Kemmerer, D. (1999). "Near" and "far" in language and perception. *Cognition, 73*(1), 35–63.
Kita, S. (Ed.). (2003). *Pointing: Where language, culture, and cognition meet*. Psychology Press.
Kornfilt, J. (1997). *Turkish*. London: Routledge.
Kouteva, T., Heine, B., Hong, B., Long, H., Narrog, H., & Rhee, S. (2019). *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
Küntay, A. C. (2012). Crosslinguistic research. In E. Hoff (Ed.), *The Blackwell guide to research methods in child language* (pp. 287–299). Malden, MA: Wiley-Blackwell.
Küntay, A. C., & Özyürek, A. (2006). Learning to use demonstratives in conversation: What do language specific strategies in Turkish reveal? *Journal of Child Language, 33*(2), 303–320.
Küntay, A. C., Nakamura, K., & Ates Sen, B. (2014). Crosslinguistic and crosscultural approaches to pragmatic development. In D. Matthews (Ed.), *Pragmatic development in first language acquisition* (pp. 317–341). Amsterdam: John Benjamins.
Levinson, S. C. (2004). Deixis and pragmatics. In L. Horn, & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 97–121). Oxford: Blackwell.
Levinson, S. C. (2018). Demonstratives: Patterns in diversity. In M. J. Dunn, N. J. Enfield, S. Cutfield, S. C. Levinson, & S. Meira (Eds.), *Demonstratives in cross-linguistic perspective*. Cambridge: Cambridge University Press, 0-32.
Lewis, G. L. (1967). *Turkish Grammar*. Oxford: Oxford University Press.
Lyons, J. (1977). *Semantics 2*. London: Cambridge University Press.
Meira, S. (2003). Addressee effects' in demonstrative systems: The cases of Tiriyó and Brazilian Portuguese. In F. Lenz (Ed.), *Deictic conceptualization of space, time and person* (pp. 3–11). Amsterdam: John Benjamins Publishing Company.
Özyürek, A. (1998). An analysis of the basic meaning of Turkish demonstratives in face-to-face conversational interaction. In S. Santi, I. Guaitella, C. Cave, & G. Konopczynski (Eds.), *Oralite et Gestualite: Communication multimodale, interaction*. Paris: L'Harmattan.
Özyürek, A., & Kita, S. (2000). Attention manipulation in the situational use of Turkish demonstratives. In *Paper Presented at the Linguistic Society of America Conference, Chicago*.
Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition, 108*, 55–184.
Peeters, D., Azar, Z., & Özyürek, A. (2014). In *The interplay between joint attention, physical proximity, and pointing gesture in demonstrative choice* (pp. 1144–1149).
Peeters, D., Hagoort, P., & Özyürek, A. (2015). Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives. *Cognition, 136*, 64–84.
Peeters, D., Krahmer, E., & Maes, A. (2021). A conceptual framework for the study of demonstrative reference. *Psychonomic Bulletin & Review, 28*, 409–433.
Peeters, D., & Özyürek, A. (2016). This and that revisited: A social and multimodal approach to spatial demonstratives. *Frontiers in Psychology, 7*, 222.
Perez Saldanya, M. (2015). Paradigms as triggers of semantic change: Demonstrative adverbs in Catalan and Spanish. *Catalan Journal of Linguistics, 14*, 113–135.
Piwek, P., Beun, R. J., & Cremers, A. (2008). 'Proximal' and 'distal' in language and cognition: Evidence from deictic demonstratives in Dutch. *Journal of Pragmatics, 40*, 694–718.
R Development Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Reile, M., Plado, H., Gudde, H. B., & Coventry, K. R. (2020). Demonstratives as spatial deictics or something more? Evidence from Common Estonian and Võro. *Folia Linguistica, 54*(1), 167–195.
Rocca, R., Tylén, K., & Wallentin, M. (2019). This shoe, that tiger: Semantic properties reflecting manual affordances of the referent modulate demonstrative use. *PloS one, 14*(1), Article e0210333.
Rocca, R., & Wallentin, M. (2020). Demonstrative reference and semantic space: A large-scale demonstrative choice task study. *Frontiers in Psychology, 11*, 629.
Rocca, R., Wallentin, M., Vesper, C., & Tylén, K. (2018). This and that back in context: Grounding demonstrative reference in manual and social affordances. In *Proceedings of the 40th Annual Meeting of the Cognitive Science Society* (pp. 1522–1527). Berlin: Springer.
Rocca, R., Wallentin, M., Vesper, C., & Tylén, K. (2019). This is for you: Social modulations of proximal versus distal space in collaborative interaction. *Scientific Reports, 9*, 14967.
Rubio-Fernandez, P. (2020). Pragmatic markers: The missing link between language and theory of mind. *Synthese, 199*(1), 1125–1158. Invited contribution to the Special Issue The cultural origins of human social cognition. https://link.springer.com/article/10.1007/s11229-020-02768-z.
Shin, N., Hinojosa-Cantú, L., Shaffer, B., & Morford, J. P. (2020). Demonstratives as indicators of interactional focus: Spatial and social dimensions of Spanish *esta* and *esa*. *Cognitive Linguistics, 31*(3), 485–514.
Shin, N. L., & Morford, J. P. (2020). Demonstratives in Spanish: Children's developing conceptualization of interactive space. In J. J. Colomina-Almiñana, & S. Sessarego (Eds.), *Language Patterns in Spanish and Beyond* (pp. 285–301). Routledge.
Skilton, A. H., & Peeters, D. (2021). Cross-linguistic differences in demonstrative systems: Comparing spatial and non-spatial influences on demonstrative use in Ticuna and Dutch. *Journal of Pragmatics, 180*, 248–265.
Slobin, D. I. (1996). From 'thought and language' to 'thinking for speaking'. In J. Gumperz, & S. C. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 70–96). Cambridge, MA: Cambridge University Press.
Soler Arechalde, M.Á. (2008). Algunos factores determinantes y contextos de uso para el marcador discursivo este... en el habla de la ciudad de México. *Anuario de Letras: Lingüística y Filología, 46*, 155–168.
Stevens, J., & Zhang, Y. (2013). Relative distance and gaze in the use of entity-referring spatial demonstratives: An event-related potential study. *Journal of Neurolinguistics, 26*, 31–45.
Stoesslein, H. E. (2014). Los marcadores del discurso más frecuentes en el castellano argentino. *ИбероаМериканские тетради, 2*, 138–146.
Stukenbrock, A. (2020). Deixis, meta-perceptive gaze practices, and the interactional achievement of joint attention. *Frontiers in Psychology, 11*, 1779.
Swift, L. B. (1963). *A Reference Grammar of Modern Turkish*. Bloomington: Indiana University.
Todisco, E., Guijarro-Fuentes, P., Collier, J., & Coventry, K. R. (2021). The temporal dynamics of deictic communication. *First Language, 41*(2), 154–178.
Todisco, E., Guijarro-Fuentes, P., & Coventry, K. R. (2021). Analogical Levelling in the Majorcan Catalan Demonstrative System. *Probus International Journal of Romance Linguistics*.
Todisco, E., Rocca, R., & Wallentin, M. (2021). The semantics of spatial demonstratives in Spanish: A demonstrative-choice task study. *Language and Cognition, 13*(4), 503–533.
Underhill, R. (1976). *Turkish grammar*. Cambridge, MA: MIT press.
Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*, 253–265.