



Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives



David Peeters^{a,b,c,*}, Peter Hagoort^{a,c}, Aslı Özyürek^{a,c}

^a Max Planck Institute for Psycholinguistics, 6500 AH Nijmegen, The Netherlands

^b International Max Planck Research School for Language Sciences, 6500 AH Nijmegen, The Netherlands

^c Radboud University Nijmegen, Donders Institute for Brain, Cognition, and Behaviour, 6500 HE Nijmegen, The Netherlands

ARTICLE INFO

Article history:

Received 23 September 2013

Revised 9 May 2014

Accepted 21 October 2014

Keywords:

Multimodal reference

Social space

Spatial deixis

Demonstratives

ERPs

ABSTRACT

A fundamental property of language is that it can be used to refer to entities in the extra-linguistic physical context of a conversation in order to establish a joint focus of attention on a referent. Typological and psycholinguistic work across a wide range of languages has put forward at least two different theoretical views on demonstrative reference. Here we contrasted and tested these two accounts by investigating the electrophysiological brain activity underlying the construction of indexical meaning in comprehension. In two EEG experiments, participants watched pictures of a speaker who referred to one of two objects using speech and an index-finger pointing gesture. In contrast with separately collected native speakers' linguistic intuitions, N400 effects showed a preference for a proximal demonstrative when speaker and addressee were in a face-to-face orientation and all possible referents were located in the shared space between them, irrespective of the physical proximity of the referent to the speaker. These findings reject egocentric proximity-based accounts of demonstrative reference, support a sociocentric approach to deixis, suggest that interlocutors construe a shared space during conversation, and imply that the psychological proximity of a referent may be more important than its physical proximity.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

A fundamental property of language is that it can be used to refer to entities in the extra-linguistic physical context of a conversation. Using a spatial demonstrative such as “this” or “that” in combination with a pointing gesture allows one to establish a joint focus of attention on a referent (Diessel, 2006), and to directly anchor one's communication to the material world (Clark, 2003; Weissenborn & Klein, 1982). The production and comprehension of such referring expressions entails a complex interaction between language, gesture, space, and attention. This has

long been a topic of interest in not only linguistics (e.g., Bühler, 1934; Jakobson, 1971; Jespersen, 1922; Levinson, 1983; Rommetveit, 1968) but also in anthropology (see Hanks, 1990, 2005), cognitive science (e.g., Kemmerer, 1999), and philosophy (e.g., Peirce, 1931; Russell, 1940). Furthermore, establishing and understanding triadic reference is a milestone in the acquisition of language and social skills (Bakeman & Adamson, 1984; Butterworth, 2003; Carpenter, Nagell, & Tomasello, 1998; Clark & Sengul, 1978; Tomasello, Carpenter, & Liszkowski, 2007). Here we focus on the neural and cognitive mechanisms underlying the comprehension of referential speech acts that contain a spatial demonstrative and a concurrently produced pointing gesture. The present study investigates this issue by contrasting and testing two different theoretical views on demonstrative reference. To set the stage for

* Corresponding author at: DCCN, P.O. Box 9101, NL-6500 HB Nijmegen, The Netherlands. Tel.: +31 (0)24 36 10980; fax: +31 (0)24 36 10989.

E-mail address: david.peeters@mpi.nl (D. Peeters).

the description of the present study, we will first outline these two views, which were put forward on the basis of previous typological and experimental studies on the production and comprehension of such spatial demonstratives.

Demonstrative systems in the world's languages have long mainly been described in egocentric and speaker-anchored terms. Speakers would use demonstrative pronouns and determiners on the basis of the proximity of the referent from their own physical location (cf. Anderson & Keenan, 1985; Diessel, 2005; Fillmore, 1982; Halliday & Hasan, 1977; Hottenroth, 1982; Lakoff, 1974; Lyons, 1977; Rauh, 1983) and as such do not consider demonstrative reference a joint activity but rather an egocentric act. In a survey of reference grammars and other typological sources on 234 languages from a large range of different language families and geographical areas, Diessel (1999, 2005) found that the large majority (i.e., 227) of languages in his corpus used a proximal–distal contrast in the description of their adnominal demonstrative system. In English, for instance, the adnominal demonstratives “this” and “that” would be used comparably to the adverbials “here” and “there”, in that a distance scale is used with the speaker as its deictic center (Anderson & Keenan, 1985; Clark & Sengul, 1978; Diessel, 1999; Fillmore, 1982; Halliday & Hasan, 1977; Lakoff, 1974; Lyons, 1977). When the two demonstratives are used contrastively, “this” is used for referents in relative proximity to the deictic center (i.e., the speaker), whereas “that” is used for referents relatively remote from the speaker's location (Diessel, 2005; Lakoff, 1974; Levinson, 1983). Also three-term systems, such as Spanish, have been described as speaker-anchored systems, with a “middle” demonstrative that would be used for entities at medial distance from the speaker (Anderson & Keenan, 1985; Diessel, 1999; Fillmore, 1982; Levinson, 2004). Such an egocentric speaker-anchored explanation of deictic reference is found not only in linguistic typology, but also in philosophical approaches to indexicality (e.g., Russell, 1940). We will call this first theoretical view the *egocentric proximity account*. Recently, this account has been taken as a starting point in psycholinguistic experimental work.

Coventry, Valdés, Castillo, and Guijarro-Fuentes (2008), for instance, experimentally investigated the production of English and Spanish demonstratives in a controlled “memory game” paradigm. Participants were seated at a table with twelve spatial locations marked by colored dots in front of them at different distances. They were instructed to name objects that were placed on the different dots by using a three-term structure (e.g., *that red triangle*). When objects were within the participant's reach, both English and Spanish participants preferred to use a proximal demonstrative (*this* in English; *este* in Spanish). When the objects were placed outside of the participants' reach, a non-proximal demonstrative was preferred (*that* in English; *ese* or *aquel* in Spanish). The authors conclude that spatial demonstratives are used on the basis of a perceptual distinction between peripersonal space near the speaker and extrapersonal space far away from the speaker (but see Bonfiglioli, Finocchiaro, Gesierich, Rositani, & Vescovi, 2009; Kemmerer, 1999).

Recently, Stevens and Zhang (2013) studied the comprehension of demonstratives by showing participants (native speakers of English) images of a speaker, an addressee, and several potential referents. In all the images, the speaker pointed and gazed toward a referent, which was always a blue cat. In a 2×3 design, the locus of attention of the addressee (on the referent, not on the referent), and the location of the referent (in speaker-associated space, i.e., close to speaker, in hearer-associated space, i.e., close to addressee, or in non-associated space, i.e., remote from both) were manipulated. Each image was then combined once with the auditory expression “this one” and once with the expression “that one”. Participants looked at the images and listened to the referential utterances, and were asked to judge by pressing one of two buttons whether the audiovisual scene (*this one* or *that one* in combination with the image) was semantically congruent or incongruent. In addition, their electroencephalogram (EEG) was recorded continuously during the experiment.

Stevens and Zhang found that participants judged a proximal demonstrative (*this*) congruent for referents in speaker-associated space, and a distal demonstrative (*that*) congruent for referents in hearer-associated and non-associated space. Thus, the participants' linguistic intuitions as expressed in their congruency judgments were in line with the egocentric proximity account outlined above. In addition, it was found that reaction times were faster for congruent demonstrative usages, but only when speaker and hearer shared gaze at the referent. Thus, the egocentric proximity account would only hold for situations in which speaker and addressee both already focus their gaze on the referent. The event-related potential (ERP) data further complicated the story, in only showing an effect in one of the six conditions. When the referent was in hearer-associated space, and there was shared gaze between speaker and hearer on the referent, a late negative deflection was found for the proximal demonstrative compared to the distal demonstrative. The authors name this an N600 effect, which would be a delayed variant of the canonical N400 effect for semantically anomalous words (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980).

The egocentric proximity account has been challenged by a number of studies that argued that a referent's relative proximity to the speaker is not the only or not the most important parameter in determining the selection of demonstratives in a particular context (Anderson & Keenan, 1985; Burenhult, 2003; Da Milano, 2007; Himmelmann, 1996; Küntay & Özyürek, 2006; Levinson, 1983; Piwek, Beun, & Cremers, 2008). Some three-term systems would have the medial demonstrative refer to entities close to the addressee (e.g., Japanese: Diessel, 1999, 2005; Levinson, 2004) thus taking the addressee's location into account. In other languages, the visibility, geographical location (e.g., uphill or downhill), and height of the referent (Anderson & Keenan, 1985; Diessel, 1999), or the presence or absence of joint attention between speaker and addressee (Küntay & Özyürek, 2006; Levinson, 2004) on the referent would also play a role in which demonstrative a speaker would select. Nevertheless, the egocentric proximity account is omnipresent in typologies of demonstrative systems (Diessel, 2005).

A theoretical alternative to the egocentric proximity account was put forward by Jungbluth (2003) in line with Weinrich (1988) and Laury (1996). This second theoretical view has been termed *the dyad-oriented account*. On the basis of fieldwork observations on Spanish, Jungbluth (2003) proposes that the physical orientation of speaker and addressee relative to each other plays a crucial role in explaining which demonstrative a Spanish speaker would select to use. More specifically, when speaker and addressee are face-to-face, as in a conversational dyad, a proximal demonstrative (*este* in Spanish) would be used to refer to an object within the shared space between speaker and addressee, irrespective of whether the referent is close to the speaker or not. In other words, every referent inside the conversational dyad is treated as proximal without any further differentiation (Jungbluth, 2003, p. 19). Referents outside of the dyad, for instance when located behind the addressee, would elicit a distal demonstrative (*aquel* in Spanish). Thus, this view on demonstratives moves away from the egocentric, ‘monadic’ typologies that underline a referent’s relative physical proximity to the speaker, and supports a more sociocentric, dyadic view of multimodal reference in which the spatial locations of speaker, addressee, and referent play an important role (Weinrich, 1988). This sociocentric approach to deixis is in line with other work arguing that the egocentric speaker-anchored proximity view is too simplistic because it does not acknowledge that establishing reference is a social, interactive process that always takes place in a socio-cultural framework (Enfield, 2003; Hanks, 1990, 2005).

In the current study we will contrast and empirically test these two theoretical accounts on spatial demonstratives, which were mainly based on the study of demonstrative production, in two electrophysiological experiments in demonstrative comprehension.

1.1. The present study

The first aim of the present study is to contrast and test the two theoretical accounts outlined above by recording participants’ electroencephalogram (EEG) and time-locked event-related potentials (ERPs) to the onset of an auditorily presented demonstrative. In order to create a visual, referential context, we presented participants with pictures in which there were two objects and a speaker. One object was always close to the speaker and another at a larger relative distance from the speaker, and on each trial the speaker referred to one of the two objects by manually pointing at it and using a demonstrative (*this* or *that*) embedded in a sentence. The two objects were located in one of two differentially oriented axes (i.e., lateral vs. sagittal) which allowed us to contrast the two theoretical views. Fig. 1 below shows a subset of the stimuli.

The egocentric proximity account (e.g., Anderson & Keenan, 1985; Clark & Sengul, 1978; Diessel, 1999; Fillmore, 1982; Halliday & Hasan, 1977; Lakoff, 1974; Lyons, 1977; Rauh, 1983; Russell, 1940) predicts that perceiving a distal demonstrative (“that”) in reference to an object close to the speaker should be reflected in a higher processing cost compared to perceiving a proximal demonstrative (“this”) in reference to the same object. Similarly, perceiving a proximal demonstrative in reference to an object that is relatively remote from the speaker should also be experienced as a violation. According to this account, it should not matter whether objects are oriented on a sagittal or on a lateral plane between speaker and addressee. More specifically, this account thus predicts a Demonstrative \times Distance interaction, and no Demonstrative \times Orientation interaction.

In contrast, the dyad-oriented account (Jungbluth, 2003; Weinrich, 1988) predicts that people prefer a

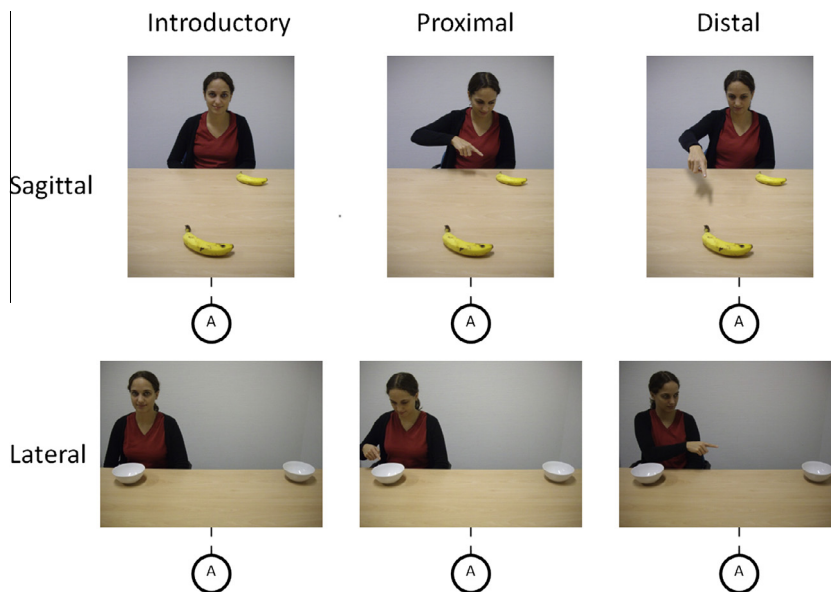


Fig. 1. Subset of stimuli used in Experiment 1. Each row contains an introductory picture, a target picture in which the speaker refers to the proximate object, and a target picture in which she refers to the distal object. The top row shows pictures with a sagittal orientation of objects. The bottom row shows pictures with a lateral orientation of objects. The position of the participant (i.e. the implicit Addressee) is marked below the pictures.

proximal demonstrative when speaker and addressee are oriented in a dyad and the referent is positioned in any location within the shared space in that dyad. This is the case for both referents in our sagittally oriented picture stimuli because the speaker and the participant addressee are sitting face-to-face, with the referents between them. It further predicts a preference for a proximal demonstrative for the object close to the speaker in the lateral orientation of objects, because this object may be considered to be inside the shared dyad space. Conversely, it predicts a preference for a distal demonstrative for the distal object in the lateral orientation, because this object is clearly outside of the dyad. In contrast with the egocentric account, this account thus predicts no Demonstrative \times Distance interaction, but a Demonstrative \times Orientation interaction effect instead.

The current study goes beyond the previous electrophysiological study by Stevens and Zhang (2013) on demonstrative comprehension in several ways. Firstly, it is the first study investigating the neural and cognitive mechanisms underlying demonstrative comprehension by an implicit addressee. It may be the case that overhearing a referential speech act uttered by a speaker for an addressee, as in Stevens and Zhang (2013), taps into different neural and cognitive mechanisms than being the (implicit) addressee of a referential speech act (e.g., Schober & Clark, 1989). Second, it is likely that people's linguistic intuitions about demonstratives do not align with their actual everyday usage and comprehension of demonstratives (Enfield, 2003; Piwek et al., 2008). Therefore, we will avoid using an experimental task that gives away the aim of the study to the participants. In addition, we will use a large number of different referents throughout the experiment and embed the demonstratives in a wider sentence context, aspects which were both absent in Stevens and Zhang (2013).

Finally, Stevens and Zhang (2013) argue that the N600 effect they found is similar to the canonical N400 effect generally found to semantic violations on the noun. However, they did not include a condition in their experiment in which they compared the two types of violation. Friedrich and Friederici (2004) have shown that hearing an incorrect label to an object elicits a widespread semantic N400 effect on the noun. In the present study, we compare the time-course and topography of such an N400 effect to an effect of a "demonstrative violation" in the same study to the same visual materials.

2. Experiment 1

2.1. Method

2.1.1. Participants

Twenty-seven participants (mean age 20.9, range 18–27, three male) took part in the experiment. They were all right-handed as assessed by a Dutch translation of the Edinburgh Inventory for hand dominance (Oldfield, 1971). Data from four participants were discarded due to a large number of trials that contained eye blinks and movement artifacts. All participants were Dutch, studying

in Nijmegen, and Dutch was their native language. They had normal or corrected-to-normal vision, no language or hearing impairments or history of neurological disease. They provided written informed consent and were paid for participation.

2.1.2. Materials and pretest of materials

The experimental materials consisted of 120 introductory pictures, 240 target pictures, and 240 spoken Dutch sentences. The pictures were selected after pretesting from an initial set of 140 introductory and 280 target pictures (see below). Half of the critical pictures were sagittally oriented and the other half of the pictures were laterally oriented (see Fig. 1). In the laterally oriented pictures, the position of the speaker (left or right) was counterbalanced. All pictures contained two similar objects, one close to the speaker and one further away from the speaker. In total, 120 different object pairs were used. In half of the critical pictures, the speaker was looking and pointing at the proximate object (the *proximate* condition). In the other half, she was looking and pointing at the remote object (the *distal* condition). The size of the laterally oriented pictures was 15×9 cm and the sagittally oriented pictures 11×12 cm, subtending visual angles of respectively $7.8 \times 4.7^\circ$ and $5.7 \times 6.2^\circ$.

The 240 sentences were spoken by a female native speaker of Dutch and digitized at a sample frequency of 44.1 kHz. They consisted of 120 sentence pairs. The two sentences in each pair only differed in the demonstrative (proximal vs. distal) they contained. All sentences started as follows: "*Ik heb* [demonstrative] [noun] *etc.*" (see Table 1 for examples translated into English) such that the words preceding the demonstrative were always the same across all conditions. For every sentence, the onsets of the demonstrative and the following noun were determined by using a speech analysis software package (Praat, version 5.2; www.praat.org). The sentences had an average duration of 2176 ms ($SD = 224$). The critical demonstratives had an average duration of 197 ms ($SD = 67$) and the critical nouns had an average duration of 409 ms ($SD = 132$). Sentences were equalized in maximum amplitude.

The experiment contained a demonstrative manipulation and a noun manipulation. The main independent variables in the demonstrative manipulation were Object Orientation (lateral or sagittal), Distance of the referent from the speaker (proximate or distal), and Demonstrative (proximal or distal). The demonstrative manipulation was done by having each picture paired with an auditorily presented sentence that could contain a proximal demonstrative (*dit* or *deze*, "this") or a distal demonstrative (*dat* or *die*, "that"). The experiment was carried out in Dutch, in which generally a distinction is made between proximal (*dit*, *deze*) and distal (*dat*, *die*) demonstratives, with separate terms for neuter (*dit*, *dat*) and common gender (*deze*, *die*) nouns following the demonstrative.

The main independent variables in the noun manipulation were Object Orientation (lateral or sagittal), Distance of the referent from the speaker (proximate or distal), and Noun (congruent or incongruent). The noun manipulation was done by having each picture paired with an auditorily presented sentence that contained a correct

Table 1

Design of both Experiments 1 and 2 with examples of the auditory stimuli (translated from Dutch).

Condition	Orientation	Distance	Example sentences	Picture
Proximal demonstrative	Sagittal	Proximate	"I have bought this plate at the market"	Plate-close
		Distal	"I have bought this plate at the market"	Plate-far
	Lateral	Proximate	"I have found this vase in the cupboard"	Vase-close
		Distal	"I have found this vase in the cupboard"	Vase-far
Distal demonstrative	Sagittal	Proximate	"I have bought that plate at the market"	Plate-close
		Distal	"I have bought that plate at the market"	Plate-far
	Lateral	Proximate	"I have found that vase in the cupboard"	Vase-close
		Distal	"I have found that vase in the cupboard"	Vase-far
Congruent noun	Sagittal	Proximate	"I have bought this plate at the market"	Plate-close
		Distal	"I have bought that plate at the market"	Plate-far
	Lateral	Proximate	"I have found this vase in the cupboard"	Vase-close
		Distal	"I have found that vase in the cupboard"	Vase-far
Incongruent noun	Sagittal	Proximate	"I have bought this mango at the market"	Plate-close
		Distal	"I have bought that mango at the market"	Plate-far
	Lateral	Proximate	"I have found this spoon in the cupboard"	Vase-close
		Distal	"I have found that spoon in the cupboard"	Vase-far

noun-referent (e.g., "plate" for a plate) or an incorrect noun-referent (e.g., "mango" for a plate). In the noun manipulation, the demonstrative preceding the noun was always similar across noun congruency conditions. Table 1 shows an overview of the different manipulations.

To select the best materials, all pictures were presented in a pre-test to 16 native Dutch participants (mean age 20.4, range 18–26) who did not participate in the main experiment. They were shown the critical pictures, one by one, on a computer screen in a soundproof booth, and were asked to type the demonstrative and the noun that they would use imagining they were the speaker in the picture. This provided us with the demonstratives that participants intuitively prefer in different conditions and the labels that participants use for the depicted objects. Whenever the speaker pointed to the close object, participants used a proximal demonstrative ("this") in 97.86% of the cases. When the speaker pointed to the remote object, participants used a distal demonstrative ("that") in 98.21% of the cases. This preference was irrespective of the orientation (lateral vs. sagittal) of objects. From the 140 pictures, the 120 pictures that were labeled most consistently were selected for the main experiment.

2.1.3. Procedure

Participants were seated in a comfortable chair in a dimly lit room at a distance of 110 cm from a computer screen. Pictures were presented on the screen using *Presentation* software (Neurobehavioral Systems) and speech was presented through EEG-compatible headphones.

Fig. 2 shows the structure of a trial. A trial consisted of a fixation cross (200 ms), an introductory picture (800 ms), another fixation cross (200 ms), the target picture (2500 ms), and a symbol (- -) during which participants could blink their eyes (3000 ms). The spoken sentence was presented 500 ms after target picture onset and lasted, on average, 2176 ms. The introductory picture was presented to familiarize participants with the objects in the picture and to mark the speaker's referential, communicative intention. Directly looking at someone is often interpreted as an ostensive signal that establishes the speaker's intention to communicate (e.g., Csibra, 2010). As such, we hoped that participants would consider themselves the addressee of the speaker's utterance, as in a referential triadic context in which speakers alternate their gaze between the referent and their addressee (e.g., Bakeman & Adamson, 1984).

Participants were instructed to carefully look at the pictures and listen to the sentences. Their task was to push a button whenever the speaker in the pictures referred to both objects in the picture (*catch trials*, 10% of all trials).

When the speaker was referring to only one object, participants did not push a button (*target trials*, 90% of all trials). They were instructed to blink their eyes only during the presentation of the symbol (- -). They were instructed to fixate on a central point on the screen and not to blink or make any saccades during the presentation of the target pictures. Participants correctly identified 89% of all catch trials.

The experiment consisted of four blocks of 60 trials. A different randomized list was used for each participant.

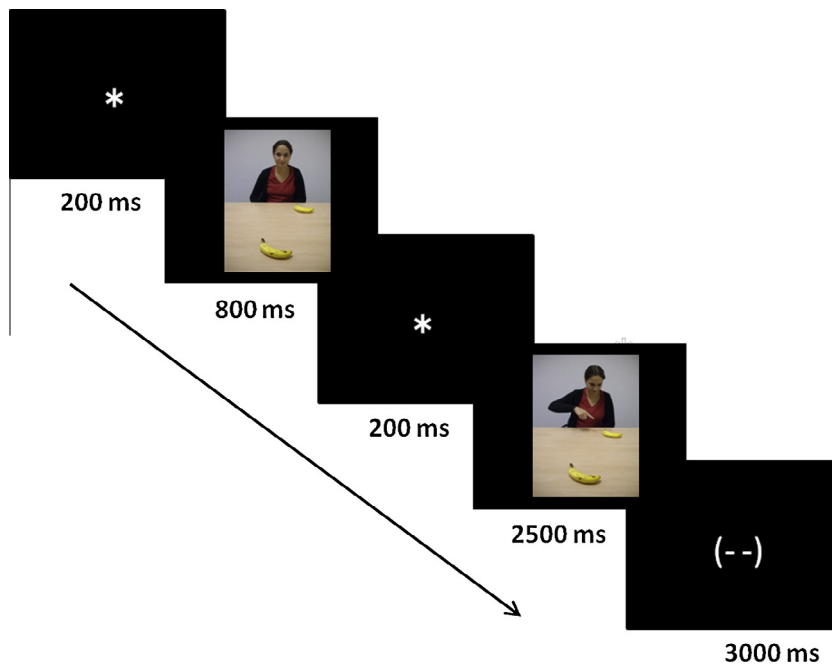


Fig. 2. Overview of the trial structure used in Experiments 1 and 2. During the presentation of the second picture a spoken sentence was presented.

The experiment started with a practice session of twelve practice items, not used in the main experiment, to familiarize the subjects with the experiment. The experimental session lasted approximately 55 min.

2.1.4. EEG recording and analysis

The electroencephalogram (EEG) was recorded continuously from 59 active electrodes held in place on the scalp by an elastic cap. Fig. 3 shows the electrode montage. In addition to the 59 scalp sites, three external electrodes were attached to record EOG, one below the left eye (to monitor for vertical eye movement/blinks), and two on the lateral canthi next to the left and right eye (to monitor for horizontal eye movements/saccades). Finally, one electrode was placed over the left mastoid bone and one over the right mastoid bone. All electrode impedances were kept below 20 K Ω . The continuous EEG was recorded with a sampling rate of 500 Hz, a low cut-off filter of 0.01 Hz and a high cut-off filter of 200 Hz. EEG was filtered offline (high-pass at 0.01 Hz and low-pass at 40 Hz). All electrode sites were online referenced to the electrode placed over the left mastoid and re-referenced offline to the average of the right and left mastoids.

Epochs from 200 ms preceding the onset of the demonstrative (“this” or “that”) to 800 ms after the onset of the demonstrative, and epochs from 200 ms preceding the onset of the noun to 800 ms after the onset of the noun were selected using Brain Vision Analyzer software (Version 2.0, Brain Products, Munich). The 200 ms pre-stimulus period was used as a baseline. Trials containing ocular or muscular artifacts were not taken into consideration in the averaging process (16.8% of all data). The mean amplitudes of the ERP waveforms for each condition per subject were entered into repeated measures ANOVAs. In the

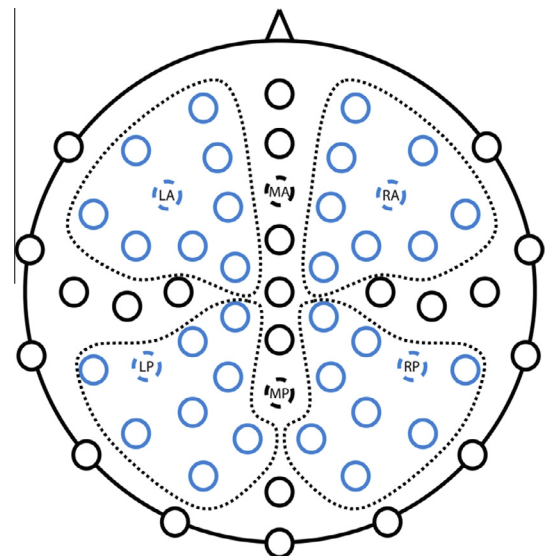


Fig. 3. Electrode montage. The four quadrants used in the analysis of the electrophysiological data (in addition to the vertical midline column) are circled and highlighted in blue. The electrode locations indicated with dashed lines refer to the electrode sites displayed in Figs. 4–7 and 9–12 (left anterior, LA; middle anterior, MA; right anterior, RA; left posterior, LP; middle posterior, MP; right posterior, RP). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

demonstrative manipulation, the independent variables were Object Orientation (lateral or sagittal; henceforth: Orientation), Distance of the referent from the speaker (proximate or distal; henceforth: Distance), Demonstrative (proximal or distal), Quadrant (left anterior, LA; right

anterior, RA; left posterior, LP; right posterior, RP; vertical midline, VM), and Electrode. In the noun manipulation, the independent variables were Orientation (lateral or sagittal), Distance (proximate or distal), Noun (congruent or incongruent), Quadrant (LA, RA, LP, RP, VM), and Electrode. For both manipulations, we performed a time-window analysis for subsequent 100 ms time-windows in the 800 ms following noun onset (in the noun manipulation) and in the 800 ms following demonstrative onset (in the demonstrative manipulation) respectively. The [Greenhouse and Geisser \(1959\)](#) correction was applied to all analyses with more than one degree of freedom in the numerator (corrected degrees of freedom and *p*-values are reported). Only significant results at the 5% level are reported.

2.2. Results

[Table 2](#) gives an overview of the results of the time-window analysis of the ERPs time-locked to the onset of the demonstrative. The overall analysis revealed most importantly a significant Demonstrative \times Orientation interaction effect in time-windows between 100 and 600 ms after demonstrative onset. Follow-up analyses revealed a significant main effect of Demonstrative in these time-windows for the picture stimuli with a sagittal orientation of referents. [Fig. 4](#) shows this effect, which denotes a more negative wave for the distal demonstrative compared to the proximal demonstrative and is widespread over the scalp (i.e., it does not interact with Quadrant). Importantly, no interaction was found with Distance, implying that the effect was similar for the referent close to the speaker compared to the referent remote from the speaker. In contrast, no significant main effect of Demonstrative was found in any of the time-windows for the pictures with a lateral orientation of referents. [Fig. 5](#) shows the grand average waveforms and the topographic plots corresponding to the difference waves for different time-windows for the pictures with a lateral orientation of referents. [Table 2](#) and [Fig. 6](#) show that the

Demonstrative \times Distance interaction predicted by the egocentric proximity account was not present in the data.

[Table 3](#) shows the outcome of the ERP analyses time-locked to the onset of the noun. A large N400 effect was found comparing incongruent to congruent referential nouns, which starts out as an effect that is larger in centro-parietal regions in early time-windows and is larger in anterior regions in late time-windows. [Fig. 7](#) shows the grand average waveforms and the topographic plots.

Finally, we directly compared the effect found time-locked to the demonstrative (as presented in [Fig. 4](#)) to the congruity effect found on the noun (as presented in [Fig. 7](#)). The dependent variable in this analysis was the difference wave, both for the demonstrative effect and for the noun effect, in the different time-windows and the independent variables used were Word Class (Demonstrative vs. Noun) and Quadrant. This analysis yielded only a significant Word Class \times Quadrant interaction effect in the 200–300 ms time-window, $F(2, 39) = 4.43$, $p = .022$. However, follow-up analyses for the separate quadrants did not reveal any significant main effect of Word Class (all p 's $> .15$). Thus, both the ERP effect on the demonstrative and the ERP effect on the noun had a negative directionality, were widespread over the scalp, and took place in similar time-windows. Due to the similar time course and negative directionality, we interpret a more negative wave in the demonstrative manipulation as reflecting a higher processing cost (cf. [Stevens & Zhang, 2013](#)).

2.3. Discussion

Experiment 1 contrasted and tested two theoretical accounts of demonstrative reference by simultaneously presenting visual and auditory stimuli while recording the EEG of native speakers of Dutch.

The results from the pretest, which tapped into participants' linguistic intuitions on demonstrative use, were in line with the egocentric proximity account in that a proximal demonstrative was used for objects relatively close to the speaker, and a distal demonstrative for objects

Table 2

Results of the ERP analyses time-locked to the onset of the demonstrative in Experiment 1. This and subsequent tables present *F*-values and significance levels for separate analyses performed for 100 ms time-windows.

	100–200	200–300	300–400	400–500	500–600	600–700	700–800
EXPERIMENT 1							
<i>Overall</i>							
Demonstrative \times Orientation \times Quadrant	1.49	1.91	1.19	<1	<1	1.31	<1
Demonstrative \times Orientation \times Distance	2.56	<1	<1	<1	<1	<1	<1
Demonstrative \times Orientation	13.91**	6.08*	5.48*	5.00*	5.45*	1.44	<1
Demonstrative \times Quadrant	<1	<1	4.11*	2.96	1.26	2.45	2.08
Demonstrative \times Distance	1.22	<1	<1	<1	<1	<1	4.05
Demonstrative	<1	14.42**	1.30	2.55	4.91*	<1	<1
<i>Lateral orientation</i>							
Demonstrative	2.14	<1	<1	<1	<1	1.15	<1
<i>Sagittal orientation</i>							
Demonstrative	11.67**	21.24***	4.61*	7.91*	9.70*	<1	<1

* $p < .05$.

** $p < .01$.

*** $p < .001$.

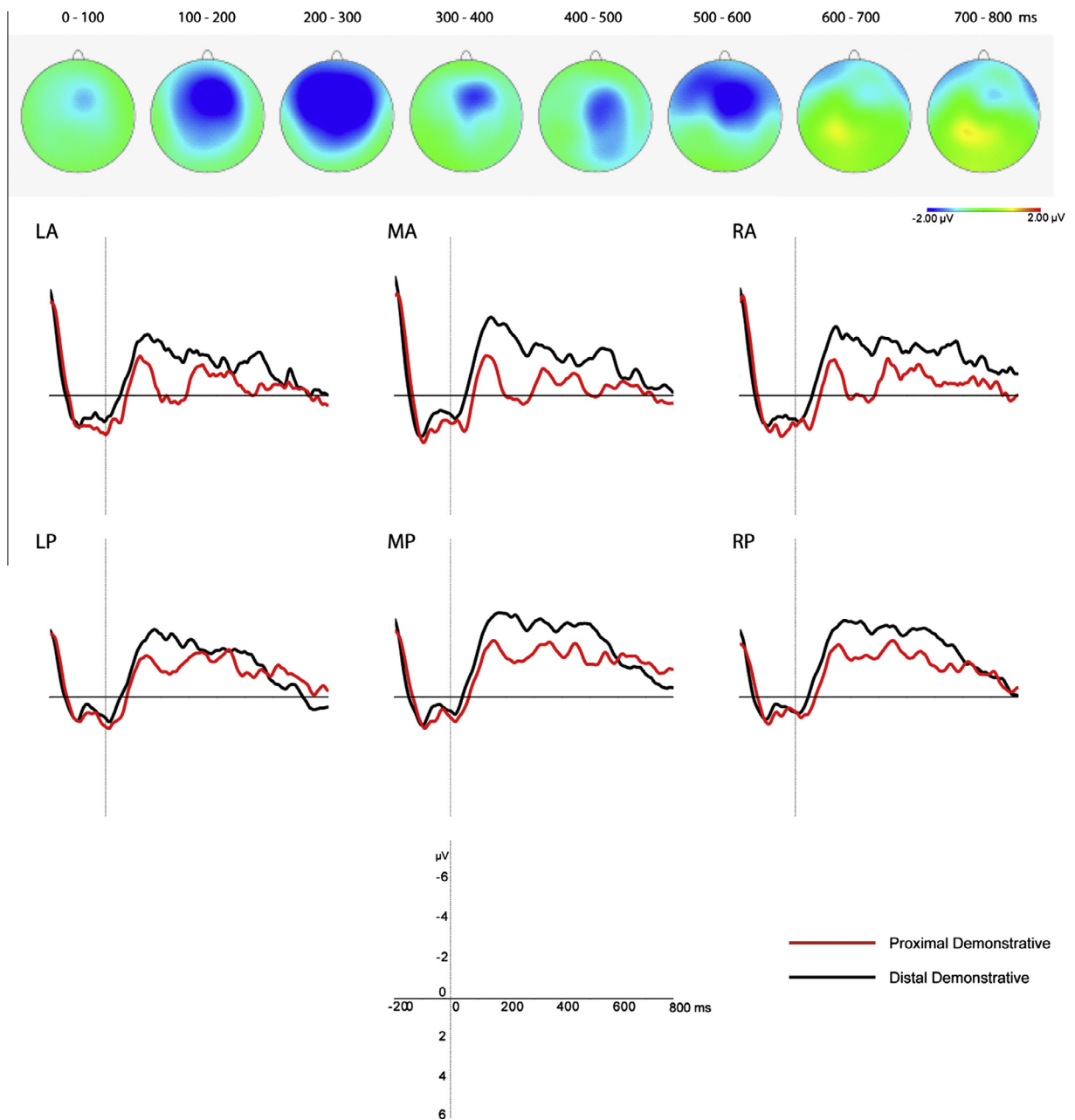


Fig. 4. Grand average waveforms time-locked to the onset of the demonstrative (proximal vs. distal) for the picture stimuli with a sagittal orientation of referents in Experiment 1. The electrode locations are indicated and can be found in Fig. 3. The topographic plots show the corresponding voltage differences between the two conditions, for 100 ms time-windows.

relatively remote from the speaker's location, irrespective of the spatial orientation of objects. These results also resemble the congruency judgments participants made in the study by Stevens and Zhang (2013). Piwek et al. (2008) termed this intuitive belief on demonstrative use "the folk view" on proximal and distal demonstratives. Interestingly, the current ERP results tell a different story.

The egocentric proximity account (e.g., Anderson & Keenan, 1985; Clark & Sengul, 1978; Diessel, 1999; Fillmore, 1982; Halliday & Hasan, 1977; Lakoff, 1974;

Lyons, 1977; Rauh, 1983; Russell, 1940) predicted a Demonstrative \times Distance interaction. According to this account, perceiving a proximal demonstrative in reference to an object remote from the speaker, and perceiving a distal demonstrative in reference to an object close to the speaker, would be experienced as caused by a violation, leading to a higher processing cost. The orientation of objects in a lateral or sagittal plane would not influence demonstrative comprehension. The current ERP data reject this account, because there was no Demonstrative \times Distance

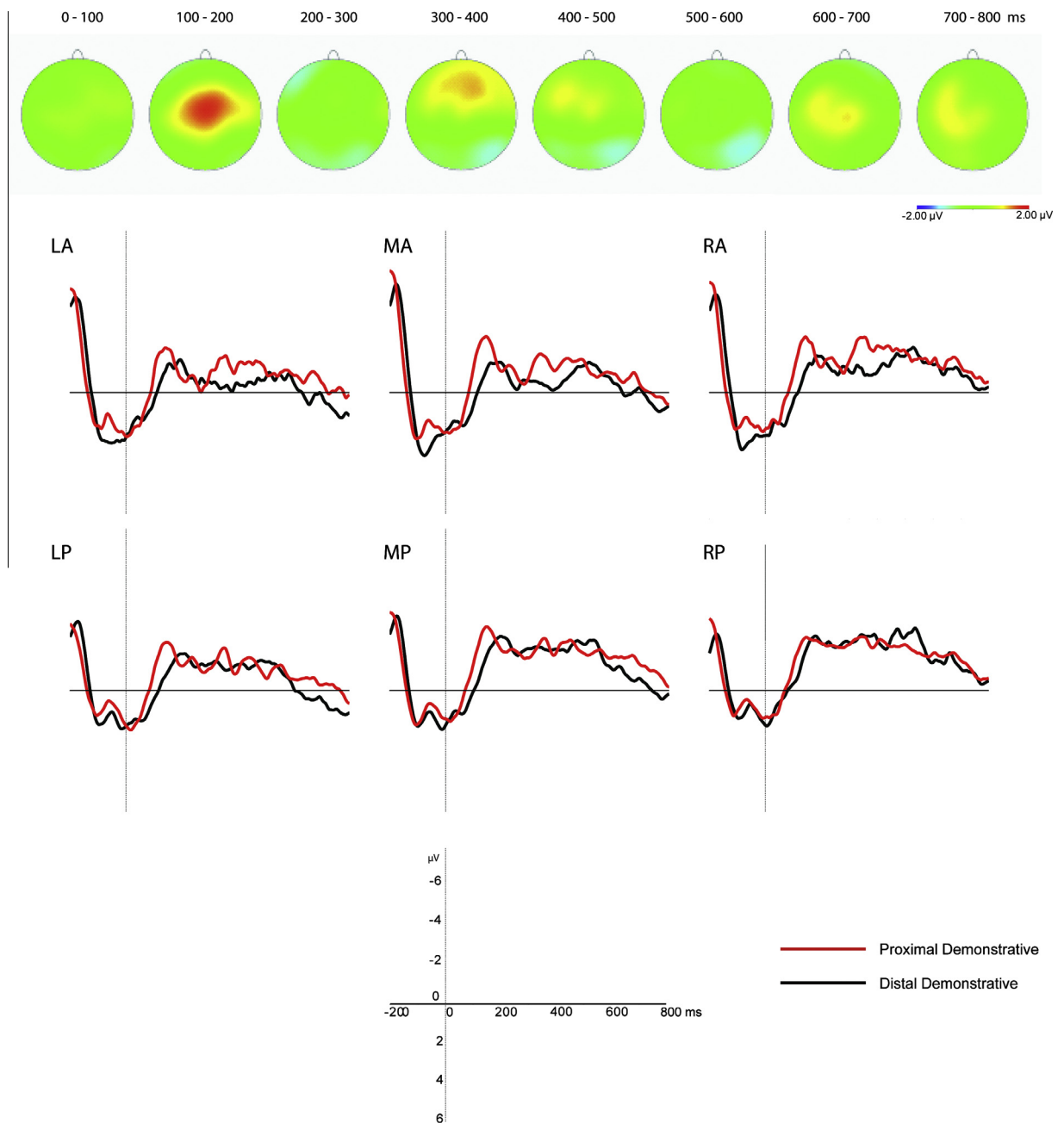


Fig. 5. Grand average waveforms time-locked to the onset of the demonstrative (proximal vs. distal) for the picture stimuli with a lateral orientation of referents in Experiment 1. The topographic plots show the corresponding voltage differences between the two conditions, for 100 ms time-windows.

interaction, and the orientation of objects played a crucial role. When objects were oriented in a sagittal plane between speaker and participant, a more negative wave was found for the distal compared to the proximal demonstrative, irrespective of whether the referent was relatively close to or remote from the speaker. On the other hand, when objects were oriented on a lateral plane, no difference was found for the ERPs to the proximal compared to the distal demonstratives. The egocentric proximity account cannot explain these findings.

The results of Experiment 1 are more in line with the dyad-oriented account, as put forward by Jungbluth (2003). Indeed, when speaker and participant were face-to-face as in a conversational dyad, and both potential referents were located in the shared space between them, the ERPs showed a more negative wave for the distal compared to the proximal demonstrative irrespective of the location of the referent. This is exactly what the dyad-oriented account predicted: distal demonstratives are more inappropriate in reference to an object in shared space, and

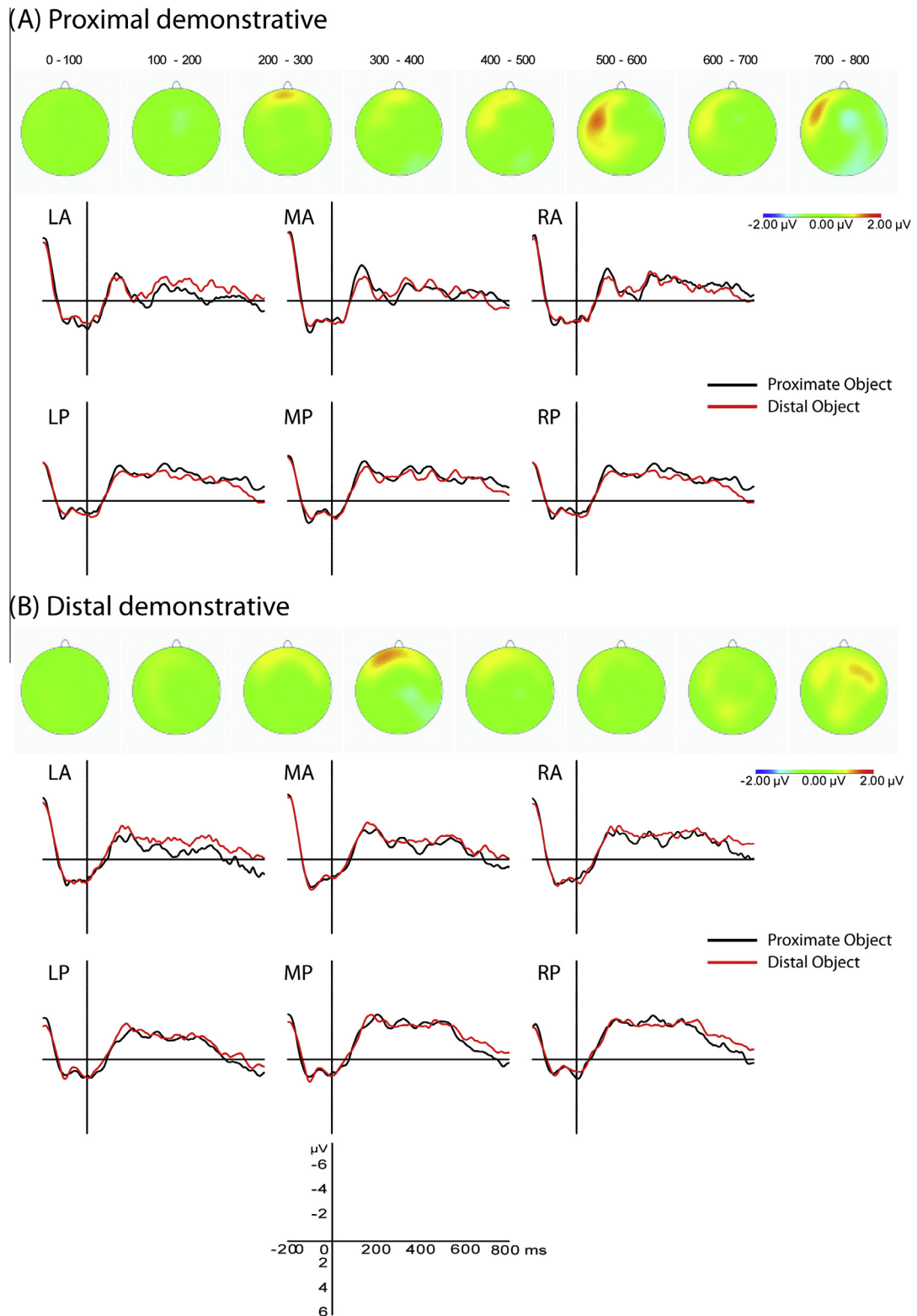


Fig. 6. Grand average waveforms time-locked to the onset of the demonstrative (panel A: proximal demonstrative; panel B: distal demonstrative) in Experiment 1 for the picture stimuli in which the referent was close to the speaker (black line) or relatively distal (red line), collapsed across Orientation. The topographic plots show the corresponding voltage differences between the two conditions, for 100 ms time-windows. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Results of the ERP analyses time-locked to the onset of the noun in Experiment 1.

	100–200	200–300	300–400	400–500	500–600	600–700	700–800
EXPERIMENT 1							
<i>Overall</i>							
Noun × Quadrant	<1	2.09	4.76*	<1	5.30*	3.47*	1.21
Noun main effect	4.14	27.73***	44.87***	18.20***	2.27	1.33	<1
<i>Follow-up</i>							
Noun main effect							
In LA quadrant	1.96	12.86**	29.01***	22.86***	9.05**	6.33*	2.54
In RA quadrant	4.17	13.16**	29.04***	15.21**	4.25	1.67	<1
In LP quadrant	8.03*	38.50***	45.70***	18.89***	<1	<1	3.44
In RP quadrant	<1	6.57*	36.29***	6.21*	<1	<1	<1
In vertical midline	10.72**	30.93***	41.92***	20.43***	2.57	<1	<1

* $p < .05$.

** $p < .01$.

*** $p < .001$.

therefore recruit more processing resources. It may seem surprising that no ERP difference was found when objects were oriented laterally, particularly for the proximate object which was located in the space between speaker and participant. However, because the picture stimuli were presented centrally on the screen from the perspective of the participant, the bodily orientations of speaker and implicit addressee were not aligned (see Fig. 1), which may be a prerequisite for creating shared space. If this is the case, aligning speaker and participant would lead to a preference for a proximal demonstrative for the object close to the speaker in the lateral orientation of objects. We tested this possibility in Experiment 2, by aligning the speaker in the pictures with the position of the addressee participant.

Furthermore, Coventry et al. (2008) have argued that there may be a parallel between the linguistic encoding of space as evident in the proximal–distal demonstrative distinction and the neural perceptual encoding of space as divided into a peripersonal (within the reach of one's arm) and an extrapersonal (beyond the reach of one's arm) region of space. Proximal demonstratives would be used to refer to objects in peripersonal space, and distal ones for objects in extrapersonal space (Coventry et al., 2008). Although this hypothesis cannot explain why we did not find an ERP difference for demonstratives in the picture stimuli with a lateral orientation of objects, one could argue that, even though in the sagittal pictures one object was relatively more distal from the speaker than the other, they were still both in peripersonal space. This possibility was also tested in Experiment 2, by enlarging the space between the two objects in the sagittally oriented picture stimuli.

3. Experiment 2

3.1. Method

3.1.1. Participants

Twenty-six participants (mean age 20.3, range 18–25, all female) took part in the experiment. They were all right-handed as assessed by a Dutch translation of the

Edinburgh Inventory for hand dominance (Oldfield, 1971). Data from two participants was discarded due to a large number of trials that contained eye blinks and movement artifacts. All participants were Dutch, studying in Nijmegen, and Dutch was their native language. They had normal or corrected-to-normal vision, no language or hearing impairments or history of neurological disease. They provided written informed consent and were paid for participation.

3.1.2. Materials and pretest of materials

There were two changes in Experiment 2 compared to Experiment 1. First, the picture stimuli with a lateral orientation of objects were now presented with the speaker central on the screen, thus aligned with the participant, who was sitting centrally in front of the screen (as in Experiment 1). Second, the pictures with a sagittal orientation of objects were modified using Adobe Photoshop such that the distance between the two objects in the pictures became larger and the distal object was clearly out of reach of the speaker (see Fig. 8).

The 120 target pictures used in Experiment 1 but adapted in the two ways described above were presented in a pretest that was similar to the pretest described for Experiment 1.

The results of this second pretest (9 subjects, 4 male, mean age 23.9, not participating in the main experiment) were similar to the results of the pretest that preceded Experiment 1. Whenever the speaker pointed to the close object, participants used a proximal demonstrative (“this”) in 96.80% of the cases. When the speaker pointed to the remote object, participants used a distal demonstrative (“that”) in 99.17% of the cases. This preference was again irrespective of the orientation (lateral vs. sagittal) of objects in the pictures.

3.1.3. Procedure, EEG recording and analysis

The experimental procedure, and the recording and analysis of the EEG data were identical to the procedure described for Experiment 1. As in Experiment 1, trials containing ocular or muscular artifacts were not taken into

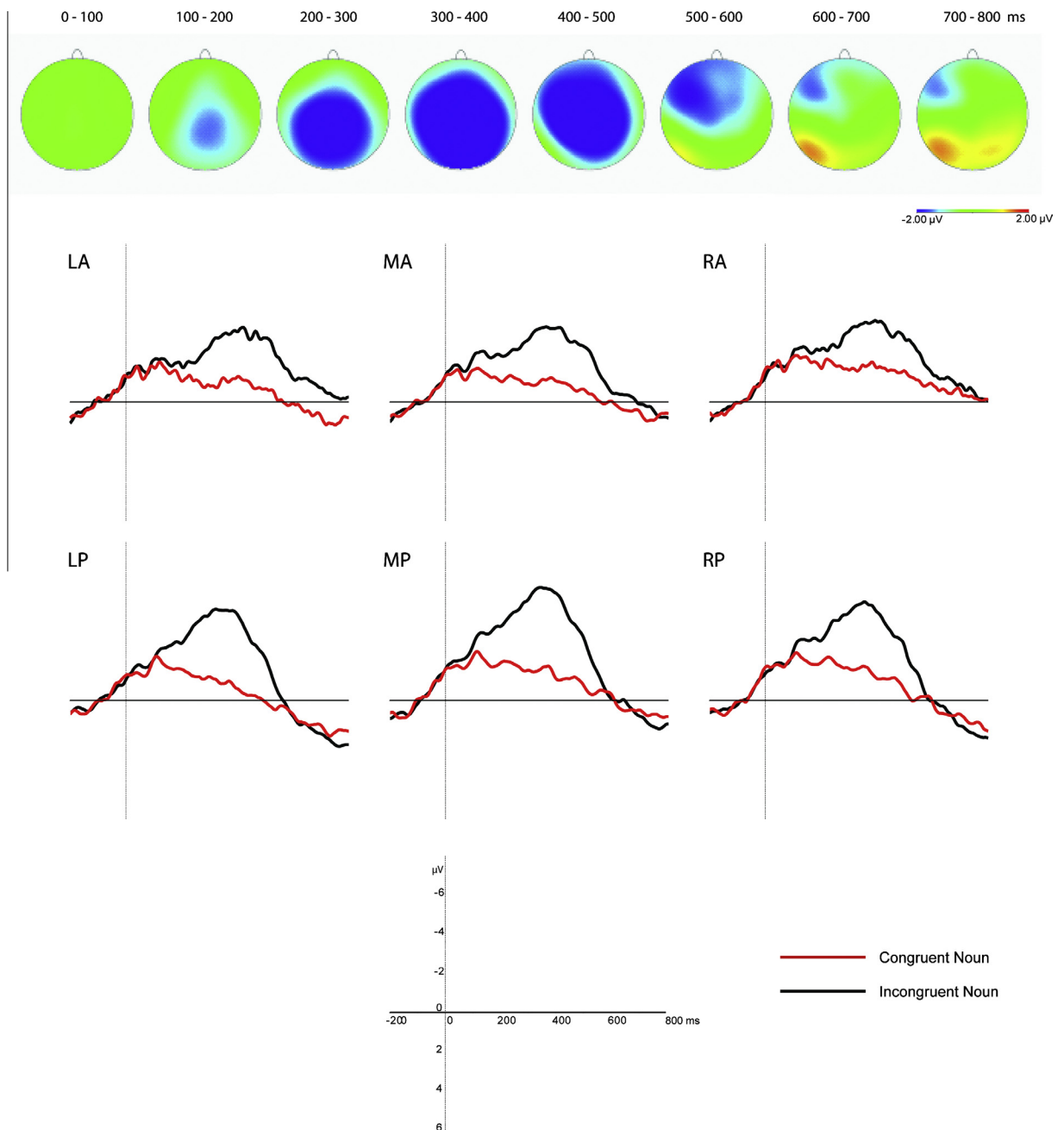


Fig. 7. Grand average waveforms time-locked to the onset of the noun (incongruent vs. congruent) in Experiment 1. The topographic plots show the corresponding voltage differences between the two conditions, for 100 ms time-windows.

consideration in the averaging process (7.83% of all data). Participants correctly identified 94.1% of all catch trials.

3.2. Results

Table 4 gives an overview of the results of the time-window analysis of the ERPs time-locked to the onset of the demonstrative in Experiment 2. The overall analysis revealed a significant three-way interaction effect between Demonstrative, Orientation, and Quadrant in time-win-

dows between 100 and 300 ms after demonstrative onset. Follow-up analyses revealed a significant interaction effect of Demonstrative \times Quadrant in these time-windows for the picture stimuli with a sagittal orientation of referents. This effect started earlier and was larger in anterior than in posterior quadrants. Fig. 9 shows this effect, which denotes a more negative wave for the distal demonstrative compared to the proximal demonstrative, having an anterior scalp distribution. No such effect was found for the pictures with a lateral orientation of objects. The follow-up

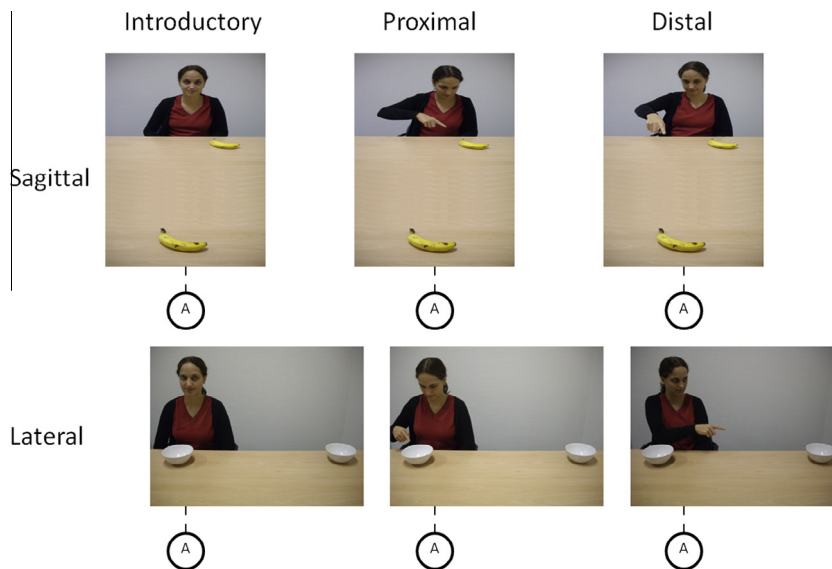


Fig. 8. Subset of stimuli used in Experiment 2. Each row contains an introductory picture, a target picture in which the speaker refers to the proximate object, and a target picture in which she refers to the distal object. The top row shows pictures with a sagittal orientation of objects. The bottom row shows pictures with a lateral orientation of objects. The position of the participant addresser is marked below the pictures.

Table 4

Results of the ERP analyses time-locked to the onset of the demonstrative in Experiment 2.

	100–200	200–300	300–400	400–500	500–600	600–700	700–800
EXPERIMENT 2							
<i>Overall</i>							
Demonstrative × Orientation × Quadrant	7.70**	7.13*	3.06	2.01	1.31	<1	1.50
Demonstrative × Orientation × Distance	<1	<1	<1	6.11*	1.64	2.28	<1
Demonstrative × Distance	<1	2.41	1.02	<1	<1	<1	<1
Demonstrative × Orientation	7.19*	7.74*	6.70	4.04	2.54	1.46	<1
Demonstrative × Quadrant	1.04	2.49	2.41	1.27	<1	<1	1.62
Demonstrative	<1	6.84*	<1	<1	<1	5.39*	1.17
<i>Lateral orientation</i>							
Dem × Quadrant	2.57	1.46	3.09	3.10	<1	<1	<1
Dem × Distance	<1	<1	<1	3.40	1.20	<1	<1
Demonstrative	2.71	<1	4.94*	2.58	3.91	4.84*	1.70
<i>Sagittal orientation</i>							
Dem × Quadrant	7.32**	7.09*	<1	<1	1.47	1.49	2.68
Dem × Distance	<1	3.11	1.76	3.27	<1	4.40*	<1
Demonstrative	2.99	12.07**	3.26	1.14	<1	<1	<1
In LA quadrant	9.00**	22.30**	1.98	<1	1.07	<1	<1
In RA quadrant	4.92*	12.65**	3.19	1.37	1.40	<1	<1
In LP quadrant	<1	6.16*	2.13	<1	<1	<1	<1
In RP quadrant	<1	3.67	3.84	<1	<1	1.05	1.28
In vertical midline	3.59	12.02**	3.78	1.38	<1	<1	<1

* $p < .05$.

** $p < .01$.

*** $p < .001$.

analysis for the picture stimuli with a lateral orientation of objects did yield a significant main effect of Demonstrative in an early and a late time-window. Fig. 10 shows the grand average waveforms and the topographic plots corresponding to the demonstrative contrast for the picture stimuli with a lateral orientation of referents. The small three-way interaction with Distance in the overall analysis did not yield any significant follow-up effects related to the Demonstrative factor. Table 4 and Fig. 11 show that the

Demonstrative × Distance interaction predicted by the egocentric proximity account was again not present in the data.

Table 5 shows the outcome of the ERP analyses time-locked to the onset of the noun in Experiment 2. Similar to Experiment 1, a large N400 effect was found when comparing incongruent to congruent referential nouns, which started out as an effect that was larger in centro-parietal regions in early time-windows and was larger in anterior

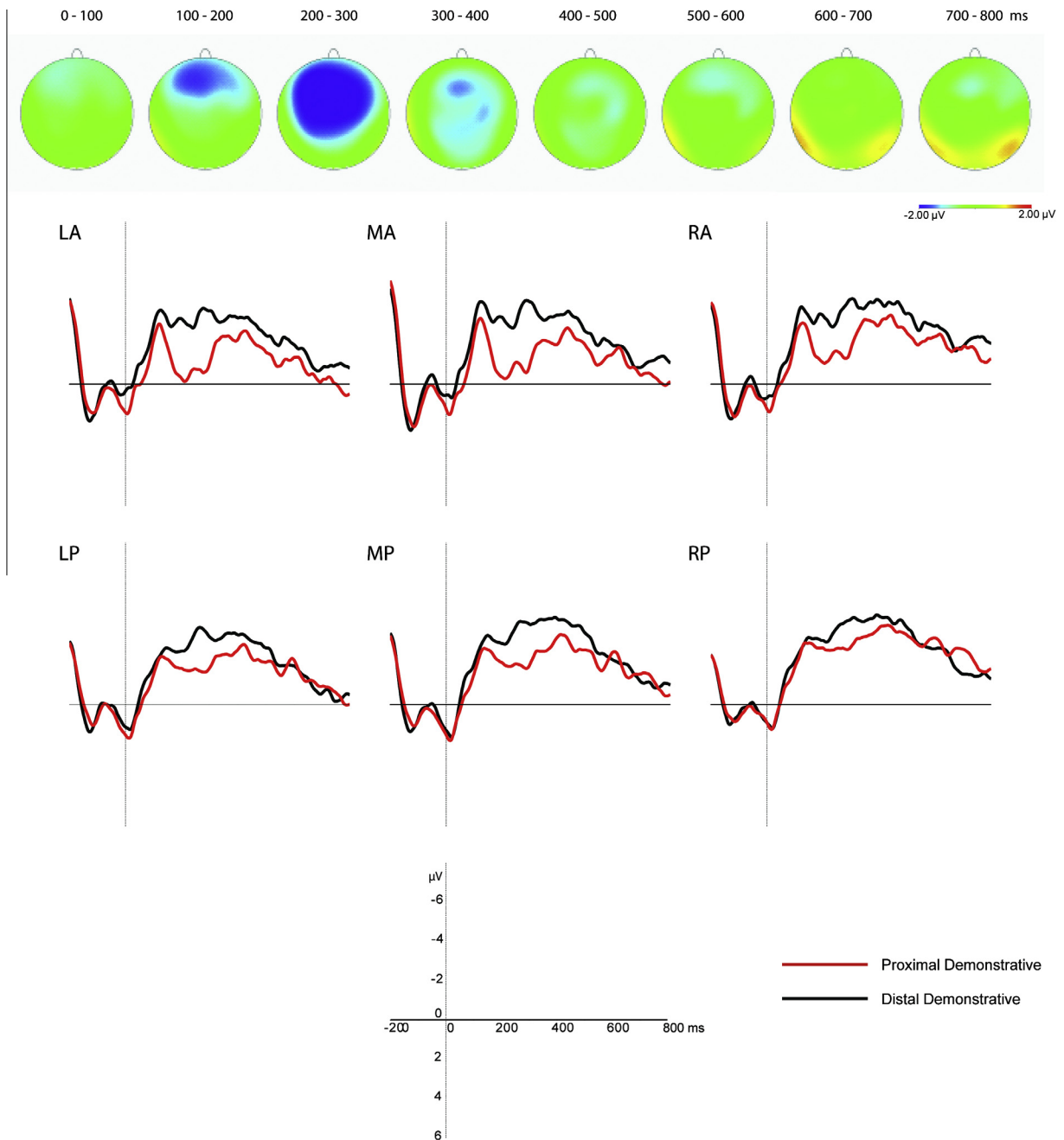


Fig. 9. Grand average waveforms time-locked to the onset of the demonstrative (proximal vs. distal) for the picture stimuli with a sagittal orientation of referents in Experiment 2. The topographic plots show the corresponding voltage differences between the two conditions, for 100 ms time-windows.

regions in late time-windows. The higher-order interaction effects denote that the effect of Noun was larger posteriorly in early time-windows and larger anteriorly in later time-windows, that it started slightly later when the referent was proximate compared to when it was distal to the speaker, and that it started slightly earlier and lasted longer when the referents were in sagittal orientation compared to when they were laterally oriented. Fig. 12 shows the overall grand average waveforms and the topographic plots.

Identical to Experiment 1, we directly compared the effect found time-locked to the demonstrative (as presented in Fig. 9) to the congruity effect found on the noun (as presented in Fig. 12). This analysis yielded significant Word Class \times Quadrant interaction effects in the 200–300 ms time-window, $F(2,50) = 3.96$, $p = .022$, the 500–600 ms time-window, $F(2,44) = 5.45$, $p = .008$, and the 600–700 ms time-window, $F(2,44) = 4.76$, $p = .015$. The early effect (200–300 ms) reflected trends towards a significant main effect of Word Class in both the left and

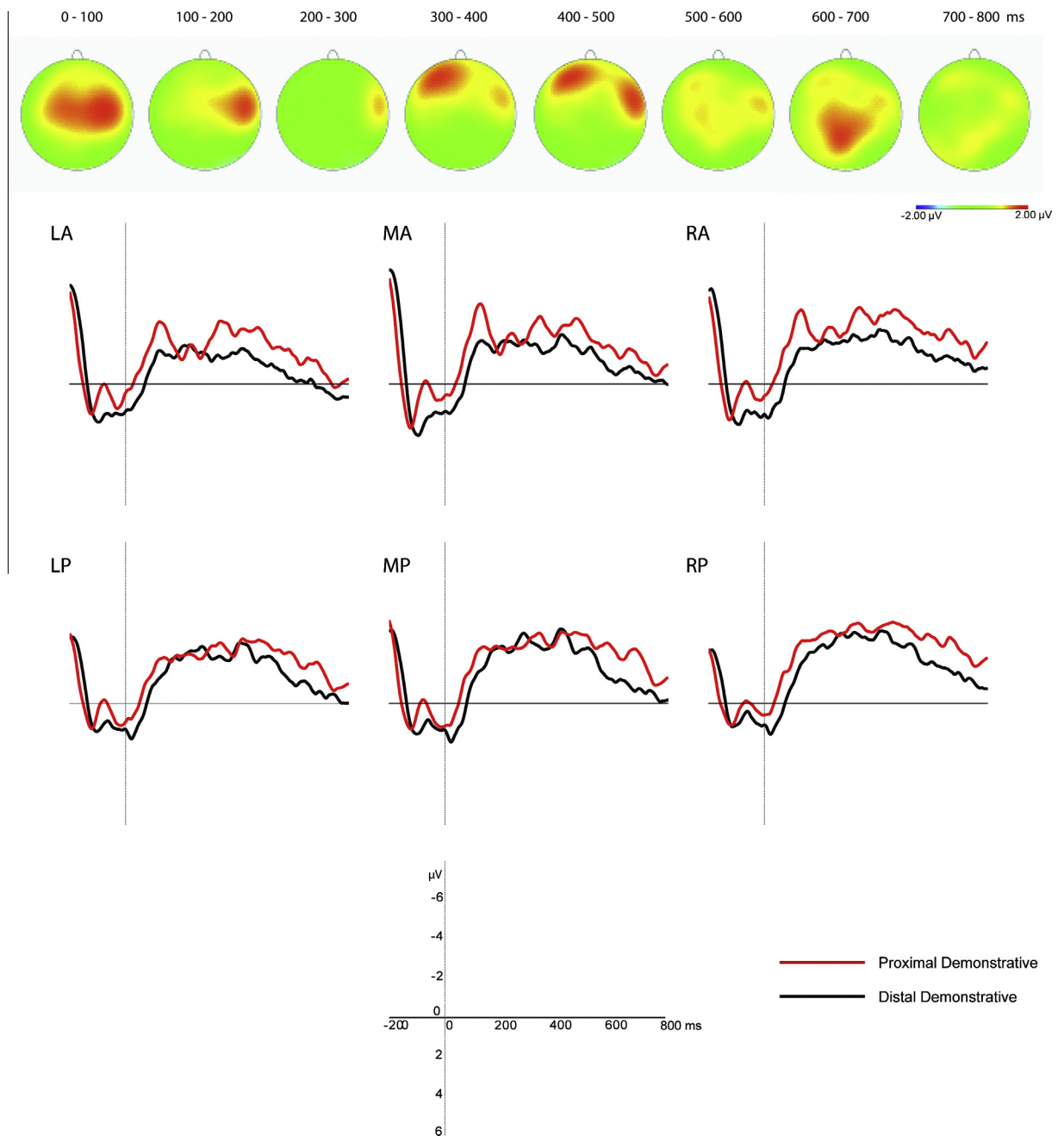


Fig. 10. Grand average waveforms time-locked to the onset of the demonstrative (proximal vs. distal) for the picture stimuli with a lateral orientation of referents in Experiment 2. The topographic plots show the corresponding voltage differences between the two conditions, for 100 ms time-windows.

right posterior quadrants ($p = .054$ and $.062$ respectively) but not in the other quadrants. The late effects reflected significant main effects of Word Class in the left anterior quadrant ($p = .001$ and $.003$ in the two subsequent late time-windows), the right anterior quadrant ($p = .011$ and $.010$ respectively) and over the vertical midline channels ($p = .033$ and $.036$ respectively). These findings reflect that the demonstrative effect was slightly larger in anterior regions in an early time-window but smaller in later

time-windows compared to the congruity effect on the noun.

3.3. Discussion

Experiment 2 further tested the dyad-oriented account of demonstrative reference. In Experiment 1, it was found that participants did not prefer one demonstrative over another when a speaker referred to one of two objects that

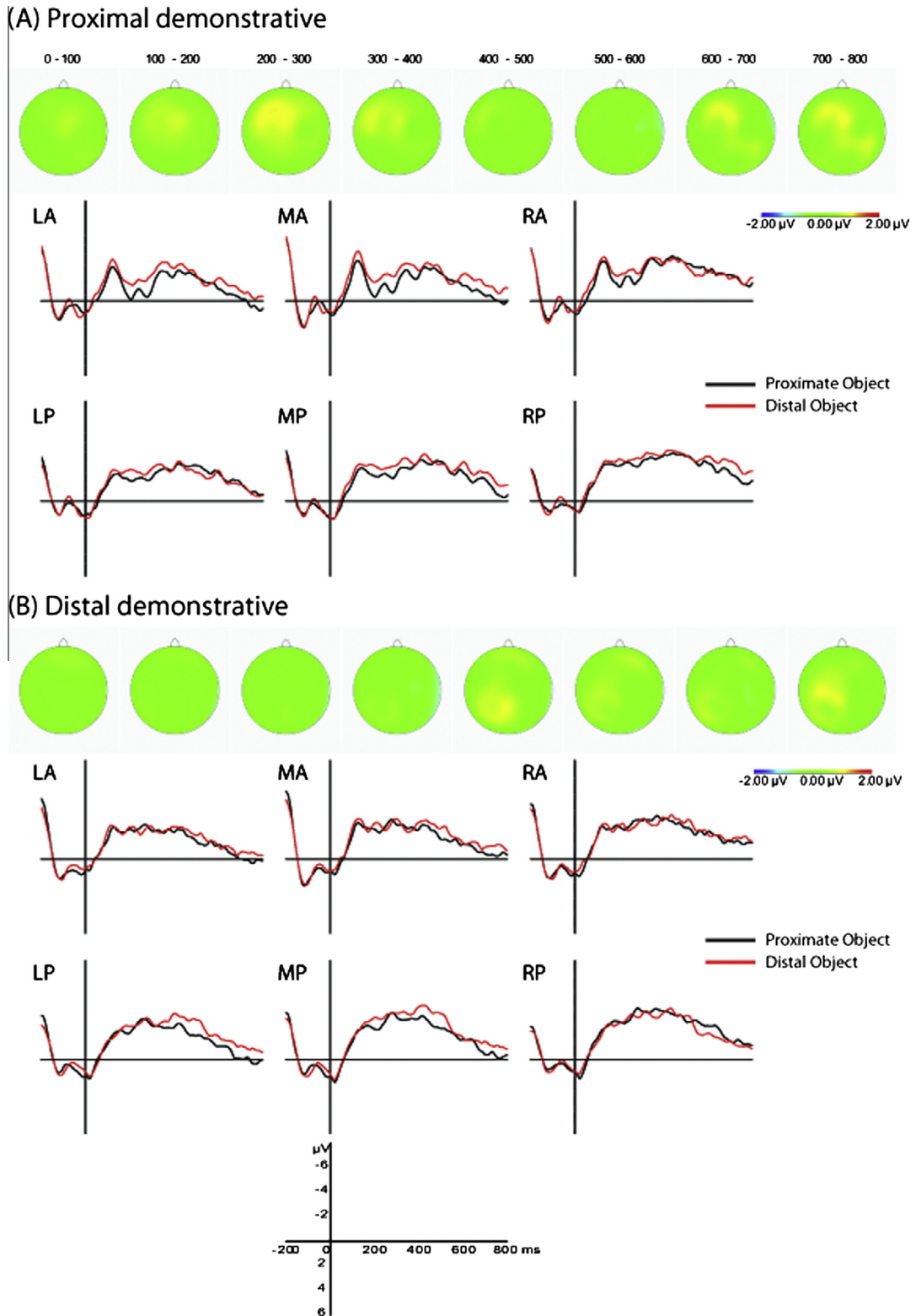


Fig. 11. Grand average waveforms time-locked to the onset of the demonstrative (panel A: proximal demonstrative; panel B: distal demonstrative) in Experiment 2 for the picture stimuli in which the referent was close to the speaker (black line) or relatively distal (red line), collapsed across Orientation. The topographic plots show the corresponding voltage differences between the two conditions, for 100 ms time-windows. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5
Results of the ERP analyses time-locked to the onset of the noun in Experiment 2.

	100–200	200–300	300–400	400–500	500–600	600–700	700–800
EXPERIMENT 2							
<i>Overall</i>							
Noun × Distance × Quadrant	2.81	5.44*	2.54	2.47	1.93	<1	1.23
Noun × Quadrant	<1	<1	<1	3.78*	20.56***	23.17***	18.54***
Noun × Orientation	3.50	9.78**	8.60**	1.91	<1	1.23	1.51
Noun × Distance	<1	<1	<1	<1	1.79	4.82*	1.92
Noun main effect	15.55**	38.37***	62.96***	61.98***	21.84***	6.49*	<1
<i>Follow-up</i>							
Noun main effect							
In LA quadrant	5.03*	19.72***	43.57***	58.14***	65.77***	45.54***	6.56*
In RA quadrant	6.94*	21.79***	42.13***	54.99***	48.37***	25.27***	5.91*
In LP quadrant	12.48**	35.62***	55.45***	35.65**	1.12	<1	3.64
In RP quadrant	18.93***	45.24***	64.50***	42.30**	1.64	<1	1.85
In vertical midline	14.99**	37.32***	54.41***	51.41***	19.05***	6.54*	<1
For proximate objects	3.66	18.35***	29.93**	42.58***	20.05***	9.27**	1.68
For distal objects	8.83**	19.42**	60.71***	38.77**	6.69*	<1	<1
In lateral orientation	1.29	8.70**	13.93**	15.20**	4.13	<1	<1
In sagittal orientation	14.54**	38.34***	65.02***	41.64***	17.77***	6.18*	1.54

* $p < .05$.

** $p < .01$.

*** $p < .001$.

were oriented laterally in front of her, as indicated by an absence of a significant ERP difference across the auditory presentation of different demonstratives. This raised the question whether the bodily orientations of speaker and (implicit) addressee need to be aligned in order to create a notion of shared space. An alternative possibility was that the dyad-oriented account only holds for situations in which all possible referents are located within the shared dyad space in between speaker and addressee. The findings of Experiment 2 were in line with this latter possibility. When the bodily orientations of speaker and participant were aligned, and the speaker referred to the object close to her while the other object was remote in a lateral plane, no proximal demonstrative preference was detected. If anything, the small negative-going main effects suggest a preference for a distal demonstrative for both objects in the lateral plane. Thus, these findings further specify the dyad-oriented account in underlining that people do indeed prefer a proximal demonstrative for referents in the shared space between speaker and addressee, but only if all possible referents are located within that space. In other words, the presence of another possible referent outside of the dyad, which introduces a lateral axis and as such a possible boundary between speaker and addressee, eliminates their experience of sharing the extra-linguistic space in between them, by creating a different division of that space. When space is no longer shared, addressees may prefer a distal demonstrative for all potential referents.

4. General discussion

In two ERP experiments, participants watched images of a speaker referring to one of two objects while they listened to her referential speech. In order to contrast and test two theoretical accounts on demonstrative reference,

participants' concurrently recorded electroencephalograms were analyzed to find out whether they had a preference for one demonstrative over another as a function of the proximity of the object-referent to the speaker and the orientation of objects in space. It was found that participants preferred a proximal demonstrative to all objects located in the shared space between speaker and addressee, but only when there was no other possible referent outside of this space. Whenever there was a potential referent outside of the shared space, participants either had no preference (as in Experiment 1) or preferred a distal demonstrative (as in Experiment 2) both for referents close to the speaker as well as for referents distal from the speaker. The theoretical importance of these findings will now be discussed in the light of the theoretical views on demonstrative reference presented in the introduction.

The first theoretical view on demonstratives put forward was termed the egocentric proximity account. This view (e.g., Anderson & Keenan, 1985; Clark & Sengul, 1978; Diessel, 1999; Fillmore, 1982; Halliday & Hasan, 1977; Lakoff, 1974; Lyons, 1977; Rauh, 1983; Russell, 1940) predicted that addressees would prefer a proximal demonstrative for referents physically close to the speaker and a distal demonstrative for referents physically remote from the speaker. Our ERP data strongly falsify this view on demonstrative reference and cannot be interpreted in line with it in any way, because the data show that participants simply did not base their online demonstrative comprehension on the referent's relative proximity to the speaker. In line with the egocentric proximity account, Coventry et al. (2008) proposed that speakers would prefer a proximal demonstrative for referents in peripersonal space and a distal demonstrative for referents in extrapersonal space. Experiment 2 falsifies this proposal, in showing that participants may prefer a proximal demonstrative for referents in the speaker's extrapersonal space. There are other theoretical and empirical arguments against the suggested

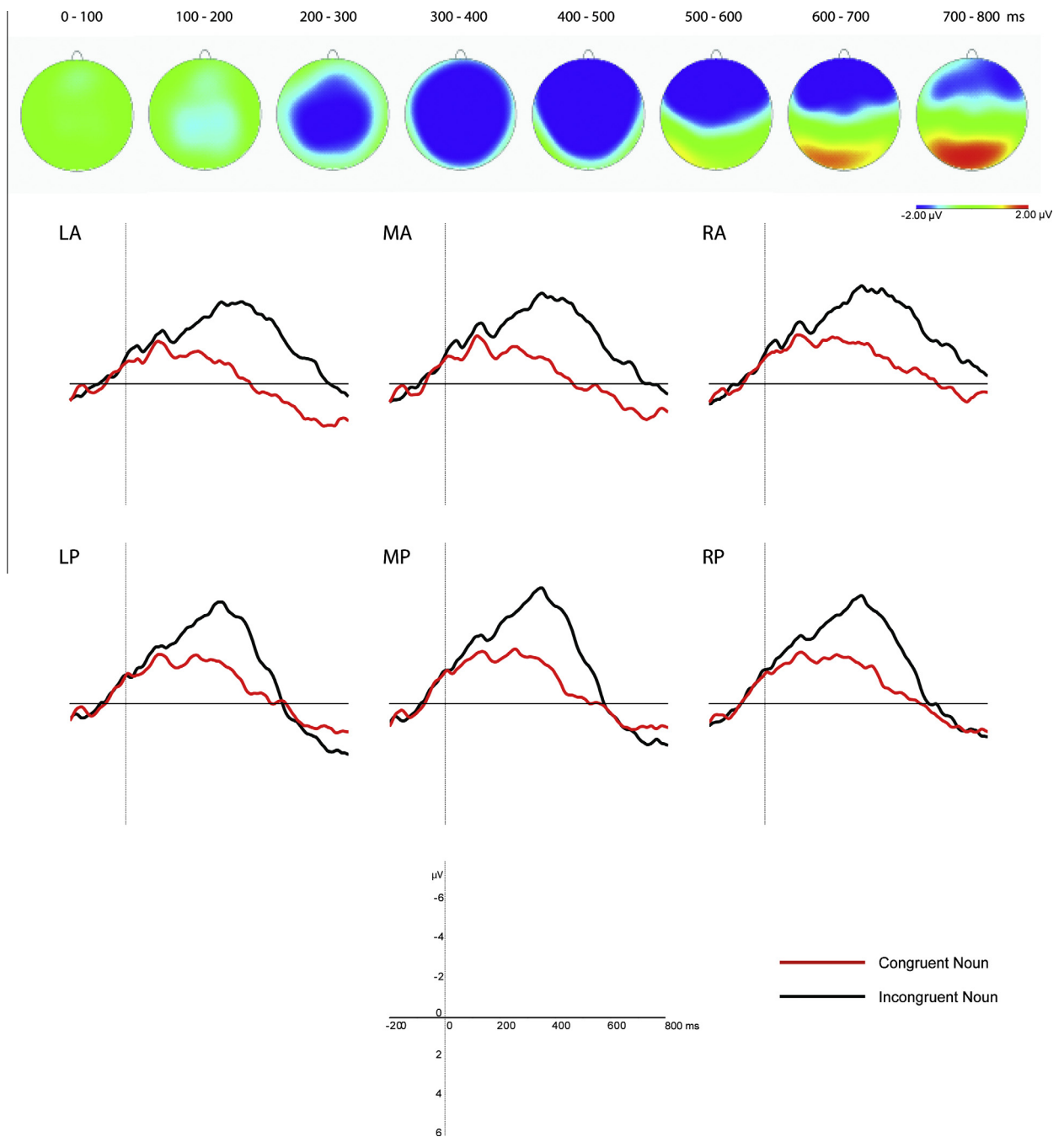


Fig. 12. Grand average waveforms time-locked to the onset of the noun (incongruent vs. congruent) in Experiment 2. The topographic plots show the corresponding voltage differences between the two conditions, for 100 ms time-windows.

parallel between perceptual and linguistic encoding of space (see e.g., Bonfiglioli et al., 2009; Kemmerer, 1999; Kirsner, 1993). As argued by Kemmerer (1999, p. 47), for instance, the mere existence of many languages in which there are three or more types of demonstrative shows that demonstrative systems need not correspond to the near-far perceptual contrast.

Nevertheless, the egocentric proximity view was in line with native speakers' linguistic intuitions on demonstra-

tives, as evident in the pre-tests of the current study and in linguistic judgments made by participants in previous studies (e.g., Stevens & Zhang, 2013). Assuming that reference grammars often partly or fully rely on linguistic intuitions, this may explain why the egocentric proximity account is omnipresent in such grammars (Diessel, 2005). However, observational (Enfield, 2003; Hanks, 1990; Jungbluth, 2003; Piwek et al., 2008) and online processing data (this study) strongly suggest that such linguistic

intuitions are not reliable (see Clark & Bangerter, 2004, for a similar argument). Although there may be instances in which speakers use demonstratives on the basis of the physical distance of referents, possibly when speaker and addressee sit side-by-side (Jungbluth, 2003), this does not imply that they are egocentric in only considering the distance of referents from their own physical location without taking their addressee's location and perspective into account. Below we will argue in favor of a sociocentric view on reference to replace this egocentric view (cf. Clark & Bangerter, 2004; Jungbluth, 2003; Weinrich, 1988).

The current ERP results are not in line with a previous EEG study on the comprehension of demonstratives either. Stevens and Zhang (2013) only found an ERP effect for situations in which a speaker referred to an object close to the addressee and speaker and addressee shared gaze at this object. They interpreted this effect as denoting a preference for a distal demonstrative in this situation, in line with participants' linguistic intuitions. In our study we found a preference for a proximal demonstrative for the referent close to the addressee. One possible explanation for this discrepancy across studies may be that overhearing referential speech, as in Stevens and Zhang (2013), may require different processing mechanisms compared to being the (implicit) addressee of referential speech and gesture (see e.g., Schober & Clark, 1989) as in the current study. An alternative explanation for the finding in Stevens and Zhang (2013), and the absence of ERP effects for any other contrast in their study, may be the fact that their participants on every trial made congruency judgments on the relation between the auditorily presented demonstrative and the visually presented image. Participants' use of linguistic intuitions and the fact that they were not naïve to the goal of the experiment may have prevented a natural processing of the presented demonstratives.

Our data are most in line with the dyad-oriented account of demonstrative reference (Jungbluth, 2003; see also Weinrich, 1988, and Laury, 1996). For the stimuli with a sagittal orientation of objects, the dyad-oriented account exactly predicted the pattern of results we found. Indeed, irrespective of the proximity of the referent to the speaker, the ERP results indicate that addressees preferred a proximal demonstrative for the referents within the dyad. However, the dyad-oriented account cannot explain why participants did not prefer a proximal demonstrative for the object close to the speaker in the lateral orientation of objects, in both experiments. This object was located in the space in between speaker and participant, and did not yield a preference for a proximal demonstrative, even when the bodily orientations of speaker and addressee were aligned as in Experiment 2. We here propose a shared-space account of demonstrative reference, which can explain these findings.

The shared-space account underlines that in human interaction physical space is transformed into meaningful space (Enfield, 2003, p. 88; Hanks, 1990; Kendon, 1977, 1990, 1992; Scheflen & Ashcraft, 1976). In the case of demonstrative reference, this means that at the time of a certain referential utterance during a conversation, interlocutors may experience some part of the extra-linguistic

space as being shared (somewhat similar to Enfield's *here-space*), which would elicit the use of proximal demonstratives for referents within the shared space. Being in a face-to-face situation with an alignment of bodily orientations, as in a conversational dyad, the space within the dyad is likely to be experienced as shared (Jungbluth, 2003). However, the presence of another possible referent in the same physical (lateral) axis as the actual referent may create a boundary that separates speaker and addressee and as such eliminates the experience of sharing space, consequently leading to a preference for a distal demonstrative or to no particular preference at all. This lateral barrier may have been experienced as more salient in Experiment 2 compared to Experiment 1, because it was orthogonal with the sagittal axis in the former, therefore leading to a stronger preference for a distal demonstrative for objects located on the lateral axis in Experiment 2. Physical boundaries may also play a role in eliminating the construal of shared space. In retrospect, this explains why participants in a building task reported by Clark and Krych (2004) used significantly more distal than proximal demonstratives in referring to their building blocks when there was a physical barrier in between of the two partners. Indeed, this preference decreased drastically when the physical barrier was removed. The shared-space account does not claim that the space interlocutors share can uniquely be in between of their aligned bodies. Instead, we propose that interlocutors in a conversation build up shared space throughout the course of a conversation, taking into account previous interactions and mutually shared knowledge.

The shared-space account moves away from egocentric explanations of deixis and pleads for a sociocentric approach instead (cf. Clark & Bangerter, 2004; Clark & Wilkes-Gibbs, 1986; Jungbluth, 2003; Weinrich, 1988). It underlines that referring is a collaborative and cooperative process (Clark, 1996; Clark & Bangerter, 2004; Clark & Wilkes-Gibbs, 1986) in which speakers are not blind to their addressees. Indeed, speakers actively take into account the location and bodily orientation of their addressee in interaction (Kendon, 1990). Kendon (1990, 1992) underlines that interlocutors often use their bodies to separate their shared space of engagement from the outside world, creating a so-called joint transactional segment. Özyürek (2000, 2002) showed that speakers also take into account the location and orientation in space of their addressees, and as such the shared space between speaker and addressee, when designing their co-speech representational gestures. Liddell (1995, 2000) showed that signers of American Sign Language take into account the location of their addressee in determining the orientation and direction of pronouns and indicating verb signs. Here we show that also in demonstrative reference the bodily orientation of speaker and addressee plays an important role and that mutual awareness of this is likely. As such, we follow Enfield (2003, p. 115) in arguing that all demonstratives are both speaker-anchored and addressee-anchored. This also fits well with a large body of work indicating that in general, speakers take into account characteristics of their specific recipient or the specific communicative situation when designing their communicatively intended message

as transmitted via speech (e.g., Brennan & Hanna, 2009; Clark, 1996), co-speech iconic gestures (e.g., Alibali, Heath, & Myers, 2001), and/or pointing gestures (e.g., Peeters, Chu, Holler, Özyürek, & Hagoort, 2013).

Demonstrative systems are often said to contain or consist of proximal and distal demonstratives (e.g., Diessel, 2005). However, if physical proximity or distance is not the driving force behind a speaker's choice for a particular demonstrative, this terminology becomes confusing. Nevertheless, the terms *proximal* and *distal* hold when interpreted as referring to psychologically proximal and distal objects, instead of referring to physically proximal and distal objects. Referents within the shared space would then be psychologically proximal and referents outside of the shared space psychologically distal. Demonstratives may indeed fit well within a construal-level theory of psychological distance (Bar-Anan, Liberman, Trope, & Algom, 2007; Liberman & Trope, 2008; Trope & Liberman, 2010) provided that such a theory adopts a sociocentric approach to replace its egocentric starting point.

The current study investigated demonstrative determiners in adnominal use. Future studies are needed to investigate whether the shared-space account also holds for demonstrative adverbs (*here* and *there* in English), in cases where demonstratives are used as pronouns and in different functions (see e.g., Kirsner, 1993; Levinson, 2004), and whether it plays a role in the choice of referring expression from the continuum between definite descriptions and zero anaphora or dropped pronouns.

Finally, the present study confirmed a previous finding by Friedrich and Friederici (2004), who showed that hearing an incorrect label to a visually presented object elicits an N400 effect. Here we extend this finding to triadic referential situations in which speakers may correctly or incorrectly refer to objects in the extra-linguistic space surrounding them. The present study allowed for a descriptive comparison of the timing, directionality and topography of ERP effects on the demonstrative and on the noun to the same visual materials. Demonstrative effects were found to take place in similar time-windows as the 'canonical' N400 effects on the noun, although the latter effects were larger and lasted longer. Arguably this is due to the demonstrative violations being more subtle and more reversible in that listeners can still relatively easily derive speaker meaning independently of the type of demonstrative used. In addition, both the effect on the noun and the effect on the demonstrative had a similar negative directionality, which led us to interpret a more negative wave as reflecting a higher processing cost (cf. Stevens & Zhang, 2013). Finally, the effects of the demonstrative had a slightly more anterior topographical distribution in an early time-window, which is not uncommon for N400 effects to picture stimuli (Kutas & Federmeier, 2011), compared to the widespread N400 effects of the noun. Alternatively, differences in word class (demonstrative determiner vs. noun) may underlie this finding. A comparison of the underlying neural circuitry for demonstratives and co-referential nouns by using imaging techniques with a higher spatial resolution than EEG is an exciting avenue for future research.

To conclude, our study contrasted and tested two theoretical views on demonstrative reference. A dyad-oriented account was most but not perfectly in line with our data. We therefore proposed a shared-space account, which embodies a sociocentric approach to deixis and underlines that, in demonstrative reference, the psychological proximity of a referent may be more important than its physical proximity.

Acknowledgments

We thank José Kivits and Flora Vanlangendonck for help in stimulus creation, Ronald Fischer for technical assistance, and Gwilym Lockwood and one anonymous reviewer for helpful comments.

References

- Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*, 169–188.
- Anderson, S. R., & Keenan, E. L. (1985). Deixis. In T. Shopen (Ed.), *Language typology and syntactic description* (pp. 259–308). Cambridge: Cambridge University Press.
- Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother infant and peer-infant interaction. *Child Development*, *55*, 1278–1289.
- Bar-Anan, Y., Liberman, N., Trope, Y., & Algom, D. (2007). Automatic processing of psychological distance: Evidence from a Stroop task. *Journal of Experimental Psychology: General*, *136*, 610–622.
- Bonfiglioli, C., Finocchiaro, C., Gesierich, B., Rositani, F., & Vescovi, M. (2009). A kinematic approach to the conceptual representations of *this* and *that*. *Cognition*, *111*, 270–274.
- Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science*, *1*(2), 274–291.
- Bühler, K. (1934). *Sprachtheorie*. Jena: Fischer.
- Burenhult, N. (2003). Attention, accessibility, and the addressee: The case of the Jahai demonstrative ton. *Pragmatics*, *13*, 363–379.
- Butterworth, G. (2003). Pointing is the royal road to language for babies. In S. Kita (Ed.), *Pointing. Where language, culture, and cognition meet*. Mahwah, NJ: Lawrence Erlbaum.
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *255*, 63, 1–174.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Clark, H. H., & Bangerter, A. (2004). Changing ideas about reference. In I. A. Noveck & D. Sperber (Eds.), *Experimental pragmatics* (pp. 25–49). Basingstoke: Palgrave Macmillan.
- Clark, H. H. (2003). Pointing and placing. In S. Kita (Ed.), *Pointing. Where language, culture, and cognition meet* (pp. 243–268). Hillsdale, NJ: Erlbaum.
- Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*, 62–81.
- Clark, E. V., & Sengul, C. J. (1978). Strategies in the acquisition of deixis. *Journal of Child Language*, *5*(3), 457–475.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.
- Coventry, K. R., Valdés, B., Castillo, A., & Guijarro-Fuentes, P. (2008). Language within your reach: Near–far perceptual space and spatial demonstratives. *Cognition*, *108*, 889–895.
- Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind and Language*, *25*(2), 141–168.
- Da Milano, F. (2007). Demonstratives in the languages of Europe. In P. Ramat & E. Roma (Eds.), *Europe and the mediterranean as linguistic areas* (pp. 25–48). Amsterdam: John Benjamins.
- Diessel, H. (1999). *Demonstratives. Form, function, and grammaticalization*. Amsterdam: John Benjamins.
- Diessel, H. (2006). Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics*, *17*(4), 463–489. <http://dx.doi.org/10.1515/COG.2006.015>.

- Diessel, H. (2005). Distance contrasts in demonstratives. In M. Haspelmat, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures* (pp. 170–173). Oxford: Oxford University Press.
- Enfield, N. J. (2003). Demonstratives in space and interaction: Data from Lao speakers and implications for semantic analysis. *Language*, 82, 117.
- Fillmore, C. J. (1982). Towards a descriptive framework for spatial deixis. In R. J. Jarvella & W. Klein (Eds.), *Speech, place, and action. Studies in deixis and related topics* (pp. 31–59). Chichester: John Wiley & Sons, Ltd.
- Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience*, 16, 1465–1477.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24(2), 95–112.
- Halliday, M. A. K., & Hasan, R. (1977). *Cohesion in English*. London: Longman Group, Ltd.
- Hanks, W. F. (1990). *Referential practice: Language and lived space among the Maya*. Chicago: University of Chicago Press.
- Hanks, W. F. (2005). Explorations in the deictic field. *Current Anthropology*, 46(2), 191–220.
- Himmelmann, N. (1996). Demonstratives in narrative discourse: A taxonomy of universal uses. In B. Fox (Ed.), *Studies in anaphora* (pp. 205–254). Amsterdam: John Benjamins.
- Hottenroth, P.-M. (1982). The system of local deixis in Spanish. In J. Weissenborn & W. Klein (Eds.), *Here and there: Cross-linguistic studies on deixis and demonstration* (pp. 133–153). Amsterdam: John Benjamins.
- Jakobson, R. (1971). *Selected writings* (Vol. 2). The Hague: Mouton.
- Jespersen, O. (1922). *Language. Its nature, development, and origin*. London: George Allen & Unwin, Ltd.
- Jungbluth, K. (2003). Deictics in the conversational dyad: Findings in Spanish and some cross-linguistic outlines. In F. Lenz (Ed.), *Deictic conceptualisation of space, time and person* (pp. 13–40). Amsterdam: John Benjamins.
- Kemmerer, D. (1999). “Near” and “far” in language and perception. *Cognition*, 73, 35–63.
- Kendon, A. (1990). *Conducting interaction. Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.
- Kendon, A. (1992). The negotiation of context in face-to-face interaction. In A. Duranti & C. Goodwin (Eds.), *Rethinking context: Language as an interactive phenomenon* (pp. 323–334). Cambridge: Cambridge University Press.
- Kendon, A. (1977). Spatial organization in social encounters: The F-formation system. In A. Kendon (Ed.), *Studies in the behavior of social interaction* (pp. 179–208). Lisse: Peter de Ridder Press.
- Kirsner, R. S. (1993). From meaning to message in two theories: Cognitive and Saussurean views of the Modern Dutch demonstratives. In R. A. Geiger & B. Rudzka-Ostyn (Eds.), *Conceptualizations and mental processing in language* (pp. 81–114). Berlin: Mouton de Gruyter.
- Küntay, A., & Özyürek, A. (2006). Learning to use demonstratives in conversation: What do language specific strategies in Turkish reveal? *Journal of Child Language*, 33, 303–320.
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203–205.
- Lakoff, R. (1974). Remarks on this and that. In R. A. Fox, M. W. La Galy, & A. Bruck (Eds.), *Papers from the tenth regional meeting* (pp. 345–356). Chicago: Chicago Linguistic Society.
- Laury, R. (1996). Conversational use and basic meaning of Finnish demonstratives. In A. E. Goldberg (Ed.), *Conceptual structure, discourse and language* (pp. 303–319). Stanford: CSLI Publications.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, S. C. (2004). Deixis. In L. Horn (Ed.), *The handbook of pragmatics* (pp. 97–121). Oxford: Blackwell.
- Liberman, N., & Trope, Y. (2008). The psychology of transcending the here and now. *Science*, 322, 1201–1205.
- Liddell, S. (1995). Real, surrogate and token space: Grammatical consequences in ASL. In K. Emmorey & J. Reilly (Eds.), *Language, gesture, and space* (pp. 19–43). Hillsdale: Erlbaum.
- Liddell, S. (2000). Blended spaces and deixis in sign language discourse. In D. McNeill (Ed.), *Language and gesture* (pp. 331–357). Cambridge: Cambridge University Press.
- Lyons, J. (1977). *Semantics* (Vol. 2). Cambridge: Cambridge University Press.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Özyürek, A. (2002). Do speakers design their cospeech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language*, 46, 688–704.
- Özyürek, A. (2000). The influence of addressee location on spatial language and representational gestures of direction. In D. McNeill (Ed.), *Language and gesture* (pp. 64–83). Cambridge: Cambridge University Press.
- Peeters, D., Chu, M., Holler, J., Özyürek, A., & Hagoort, P. (2013). Getting to the point: The influence of communicative intent on the kinematics of pointing gestures. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual meeting of the cognitive science society* (pp. 1127–1132). Austin, TX: Cognitive Science Society.
- Peirce, C. S. (1931). *The collected writings of Charles Sanders Peirce*. Cambridge, MA: Harvard University Press.
- 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.
- Rauh, G. (1983). Aspects of deixis. In G. Rauh (Ed.), *Essays on deixis* (pp. 9–60). Tübingen: Narr.
- Rommetveit, R. (1968). *Words, meanings, and messages: Theory and experiments in psycholinguistics*. New York: Academic Press.
- Russell, B. (1940). *An inquiry into meaning and truth*. London: George Allen & Unwin, Ltd.
- Schefflen, A. E., & Ashcraft, N. (1976). *Human territories. How we behave in space-time*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21, 211–232.
- 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.
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, 78, 705–722.
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117, 440–463.
- Weinrich, H. (1988). Über Sprache, Leib, und Gedächtnis. In H.-U. Gumbrecht (Ed.), *Materialität der Kommunikation* (pp. 80–93). Frankfurt am Main: Suhrkamp.
- Weissenborn, J., & Klein, W. (Eds.). (1982). *Here and there: Cross-linguistic studies on deixis and demonstration*. Amsterdam: John Benjamins.