

# Electrophysiological and Kinematic Correlates of Communicative Intent in the Planning and Production of Pointing Gestures and Speech

David Peeters<sup>1</sup>, Mingyuan Chu<sup>1,2</sup>, Judith Holler<sup>1</sup>, Peter Hagoort<sup>1,3</sup>,  
and Aslı Özyürek<sup>1,3</sup>

## Abstract

■ In everyday human communication, we often express our communicative intentions by manually pointing out referents in the material world around us to an addressee, often in tight synchronization with referential speech. This study investigated whether and how the kinematic form of index finger pointing gestures is shaped by the gesturer's communicative intentions and how this is modulated by the presence of concurrently produced speech. Furthermore, we explored the neural mechanisms underpinning the planning of communicative pointing gestures and speech. Two experiments were carried out in which participants pointed at referents for an addressee while the informativeness of their gestures and speech was varied.

Kinematic and electrophysiological data were recorded online. It was found that participants prolonged the duration of the stroke and poststroke hold phase of their gesture to be more communicative, in particular when the gesture was carrying the main informational burden in their multimodal utterance. Frontal and P300 effects in the ERPs suggested the importance of intentional and modality-independent attentional mechanisms during the planning phase of informative pointing gestures. These findings contribute to a better understanding of the complex interplay between action, attention, intention, and language in the production of pointing gestures, a communicative act core to human interaction. ■

## INTRODUCTION

Human communication in everyday life is canonically driven by a speaker's communicative intentions to convey meaning to an addressee and dependent on the successful recognition of such intentions by that addressee (Bara, 2010; Sperber & Wilson, 1995; Grice, 1975). Ontogenetically, one of the first ways in which we express our communicative intent is by producing pointing gestures to the things around us (e.g., Carpenter, Nagell, & Tomasello, 1998). Such pointing gestures are a foundational building block of human communication (Kita, 2003) and pave the way for the acquisition of language (Csibra, 2010; Iverson & Goldin-Meadow, 2005; Butterworth, 2003; Moore & D'Entremont, 2001; Bates, Camaioni, & Volterra, 1975). Throughout life, in concert with speech, they allow us to directly connect our communication to the material world around us (Enfield, Kita, & De Ruiter, 2007; Bangerter, 2004; Clark, 2003; Kita, 2003). Cognitive models of speech and gesture production generally acknowledge the role of one's communicative intentions in driving the production of co-speech gesture (Melinger & Levelt, 2004; Kita & Özyürek, 2003;

De Ruiter, 2000; but see Krauss, Chen, & Gottesman, 2000). Here we investigate whether and how the kinematics of pointing gestures are indeed shaped by one's communicative intentions and whether this is modulated by the presence of concurrently produced speech. In addition, we explore the neural and cognitive mechanisms involved in the planning of intentionally communicative pointing gestures and speech. To set the stage for a description of two experiments, we will first discuss previous research on communicative actions and intentions in general and then pointing gestures and speech more specifically.

In everyday life, our hands and arms rarely rest. As humans, we interact with the world around us through manipulating and acting upon objects, and on many occasions, we do not do so just by ourselves but in the context of joint activities involving the presence of others (e.g., Vesper & Richardson, 2014). Crucially, those others have been shown to influence the way we perform instrumental actions (see Becchio, Manera, Sartori, Cavallo, & Castiello, 2012). For example, the movement kinematics of actions such as reaching for and grasping an object have been found to be shaped by the actor's communicative intentions (Sartori, Becchio, Bara, & Castiello, 2009). In turn, observers may derive and anticipate the actor's intentions by attuning to such subtle

<sup>1</sup>Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands, <sup>2</sup>University of Aberdeen, UK, <sup>3</sup>Radboud University Nijmegen, The Netherlands

kinematic parameters in the actor's movement (Sartori, Becchio, & Castiello, 2011). Research in the domain of (representational) co-speech hand gestures (e.g., Özyürek, 2002; see also Holler & Stevens, 2007; Gerwing & Bavelas, 2004) and interpersonal signaling in communicative actions more broadly (Vesper & Richardson, 2014) suggest that the close link between action and intention may not be restricted to instrumental or representational actions (see Pierno et al., 2009). Preliminary indications indeed suggest a relation between the kinematics of pointing gestures and the speaker-gesturer's communicative intent (Cleret de Langavant et al., 2011; Enfield et al., 2007). The field work by Enfield and colleagues (2007), for instance, suggests that the size of a pointing gesture may depend on whether it is intended to carry informationally foregrounded or backgrounded information in the speaker's utterance. However, whether, and if so how, the kinematics of pointing gestures, like instrumental actions, are shaped by context-specific communicative intentions remains largely unclear.

Pointing gestures often come with concurrent deictic speech such as spatial demonstratives (e.g., "this" and "that" in English). Speech and gesture are temporally tightly interconnected in the production of referring expressions (e.g., Chu & Hagoort, 2014; Kendon, 2004; McNeill, 1992; Levelt, Richardson, & La Heij, 1985) and can be used independently or simultaneously to single out a referent (Bangerter, 2004), that is, an object, person, or event on which one wishes to focus the attention of one's addressee by referring to it. Previous work has investigated whether the presence of speech as a second modality changes the kinematics of a corresponding gesture. Chieffi, Secchi, and Gentilucci (2009) found no kinematic difference between a condition in which participants manually pointed to a remote referent and a condition in which they did the same but also concomitantly produced congruent deictic speech ("there"). In contrast, Gonseth, Vilain, and Vilain (2013) found that pointing gestures produced without corresponding speech had a lower velocity and a longer poststroke hold phase compared to when deictic speech was concomitantly produced. This discrepancy in findings asks for further investigation.

The current study also aims to advance our understanding of the neural mechanisms involved in the planning and production of pointing gestures. Both in infants and adults, frontal markers of neuronal activity have been identified as being involved in the production of pointing gestures establishing a joint, interpersonal focus of attention on a referent (Cleret de Langavant et al., 2011; Henderson, Yoder, Yale, & McDuffie, 2002; Mundy, Card, & Fox, 2000). This frontal activation has been interpreted as reflecting the involvement of intention-related "mentalizing" networks (e.g., Brunetti et al., 2014). Using magnetoencephalography, Brunetti et al. (2014) found enhanced activity in dorsal regions of the ACC (in medial frontal cortex) for declarative pointing ("pointing to share

attention to an object—interpersonal" in their manipulation) compared to imperative pointing ("pointing to request an object—instrumental" in their manipulation) and argue that this difference reflects enhanced mentalizing activity. Central to the difference between the two conditions is the explicit assumption that imperative pointing has only an instrumental purpose. This is problematic though, because arguably also in imperative pointing the person gesturing considers her addressee as a mental, intentional agent when requesting an object by pointing (see Southgate, Van Maanen, & Csibra, 2007). Therefore, in the current study, we compare two situations that are both communicative and differ only in the communicative intent of the speaker-gesturer. Furthermore, it is an open question whether also other (e.g., attentional) neuronal mechanisms are involved in the planning and production of communicative pointing and whether (and if so, how) the presence of concomitantly produced speech interacts with possible intentional and attentional mechanisms involved.

In the current study, we adapted a paradigm introduced by Levelt et al. (1985) in which participants produce pointing gestures in an experimental setting (see also Chu & Hagoort, 2014; De Ruiter, 1998). In our manipulation, participants were asked to point with their index finger at one of four circles that lit up on a screen with or without producing concurrent speech. Index finger kinematics, speech, and EEG were continuously recorded. Crucially, as a proxy of the participants' communicative intent in the current study, we manipulated the informativeness of the pointing gestures. The notion of informativeness has been used successfully in previous studies to tap into communicative intentions involved in speech production (e.g., Willems et al., 2010). Everyday pointing gestures canonically occur in a context in which interlocutors share a joint attentional frame in which one person directs the attention of another person toward a location, event, or other entity in the perceptual environment, usually precisely to be informative about these referents (Tomasello, Carpenter, & Liszkowski, 2007). In the current study, in line with findings on communicative actions more broadly (Vesper & Richardson, 2014), participants may alter the kinematic properties of their movements to make them more informative, for instance, by slowing down the movements. Alternatively, different intentions may lead to different patterns of neural activity (see below) but lack behavioral consequences as reflected in the kinematic properties of the pointing movements (cf. Brunetti et al., 2014).

The current approach allows for time-locking ERPs not only to the onset of the gesture but also to the presentation of the stimulus/referent. Several effects can be predicted on the basis of previous work. Potential frontal effects in the current study may reflect participants' communicative intentions in planning their pointing gestures (cf. Brunetti et al., 2014; Cleret de Langavant et al., 2011; Henderson et al., 2002). Furthermore, upon the intention

to produce a more informative gesture, participants may allocate more attentional resources to the task. P3b amplitude may be modulated by task-related cognitive demands that drive attentional resource allocation, such that its amplitude is smaller when a task requires greater amounts of attentional resources (Polich, 2007), in particular when attentional resource allocation is under voluntary control and perceptual quality of the stimuli does not differ across conditions (Kok, 2001), as in our setup. Smaller amplitude of the stimulus-locked P3b in our study may therefore index that participants voluntarily allocate more attentional resources when planning a more informative gesture for their addressee. A final possibility is that the readiness potential (or “Bereitschaftspotential”; Kornhuber & Deecke, 1965) is sensitive to our manipulation of communicative intent, which would be marked directly preceding the onset of the pointing gesture’s execution as measured over contralateral, central electrode sites. Hence, in addition to investigating the effects of communicative intent on pointing gesture production, we also consider specific ERP components during the course of planning informative pointing gestures, including the P3b.

We present two experiments that aim to further our understanding of the basic human communicative act of producing pointing gestures to a visible referent. On the basis of the theoretical considerations outlined above, Experiment 1 investigates (i) whether and how communicative intentions shape the kinematic properties of manual pointing gestures, (ii) whether and how this is modulated by the presence of speech as a second modality, and (iii) the neural mechanisms underlying the communicative intent involved in planning pointing gestures and speech. In everyday multimodal referential communicative acts, the informational burden can be distributed differentially over the spoken and gestural modalities (e.g., Enfield et al., 2007). Therefore, Experiment 2 tests to what extent the kinematic and electrophysiological find-

ings obtained in Experiment 1 are modality-independent, that is, whether they generalize to situations in which speech, rather than gesture, carries the informational burden in identifying a referent for an addressee in a multimodal utterance.

## EXPERIMENT 1

### Methods

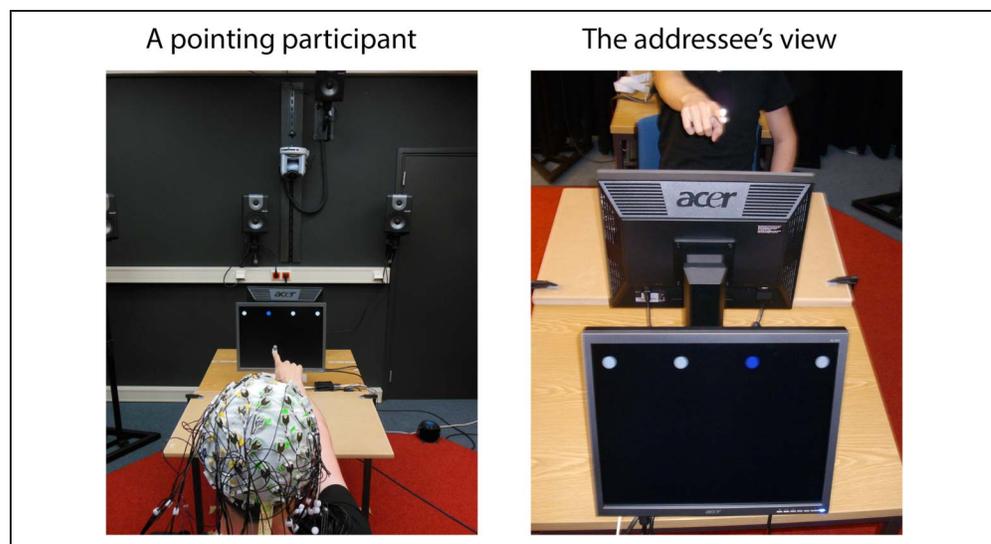
#### Participants

Twenty-four native speakers of Dutch (12 women; mean age = 20.6 years), studying at Radboud University Nijmegen, participated 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 additional participants were discarded due to a large number of trials that contained movement artifacts. Participants had normal or corrected-to-normal vision and no language or hearing impairments or history of neurological disease. They provided written informed consent and were paid €20 for participation.

#### Experimental Design and Setup

Participants were seated at a distance of 100 cm from a computer screen that was placed back-to-back with another computer screen (henceforth: the back screen). Stimuli were four white circles in a horizontal line on the top of the screen the participant viewed, mirroring four circles on the back screen. The circles could light up in either blue or yellow. A second participant (a confederate; henceforth, the addressee) looked at the back screen and the participant’s pointing gesture via a camera. Figure 1 shows the addressee’s view via the camera. On all trials, participants referred to the circle that lit up. The addressee noted on a paper form which of the four circles the participant referred to on each

**Figure 1.** Left: A participant pointing at a circle while EEG, motion tracking kinematics, and speech were continuously recorded. Right: The addressee’s view of the back screen and the pointing participant during a less informative trial.



trial (in speech and/or gesture). To make the deictic act more informative in one case but less informative in the other, the following setup was used. In both conditions, via a camera, the addressee observed the pointing gesture of the participant, as well as the circles on the back screen providing the corresponding view of the four circles the participant was seeing. This way, the addressee saw which of the four circles the participant pointed at. Before the arrival of the addressee, the experimenter showed the participant the computer to be used by the addressee and demonstrated that the addressee could see the participant's pointing gestures referring to circles on the computer screen. In this way, the participant knew that the addressee would look at the participant's gestures and to the circles presented on the back screen.

We manipulated the informativeness of the gesture (more informative vs. less informative) as well as the modality of the deictic act (gesture-only vs. gesture + speech) in a  $2 \times 2$  within-participants design. In the more informative condition, a circle turned blue or yellow only on the participant's screen but not on the back screen. Therefore, the participant's pointing gesture was the only source of information on which the addressee could base his or her decision in selecting the circle referred to by the participant. In the less informative condition, the respective circle would light up on both the participant's and the addressee's screen. Thus, the participant's pointing gesture was less informative, because the addressee saw the respective circle light up on the back screen at the same moment as the participant saw the corresponding circle light up (i.e., even before the onset of the participant's pointing gesture and/or speech). The participant received written instructions on the screen before each block, specifying whether during that block the addressee would or would not also see circles light up during that block. We decided to not have the addressee give feedback to the participant during the experiment and keep the gesturer's head out of the camera's shot to avoid differences in feedback across conditions and participants (cf. Campisi & Özyürek, 2013; Holler & Wilkin, 2011) and control for the deictic function of eye gaze.

The modality factor was manipulated by having participants use either one or two modalities in referring to the circles. In gesture-only blocks (G-only), participants pointed to a circle when it turned blue or yellow without producing speech. In gesture + speech blocks, participants pointed to the circle and said either *die blauwe cirkel* ("that blue circle") or *die gele cirkel* ("that yellow circle"), depending on the color of the circle. Note that, because any of the four circles could turn blue or yellow on any trial, the speech, which only ever referred to color but never to location, was never informative (neither in the more informative nor the less informative blocks) in this experiment. The rationale for this was that we were interested in the possible effect of the mere

presence of speech as a second modality, in addition to the informativeness of the deictic act that was manipulated separately in the gesture. Figure 2 gives an overview of the manipulation.

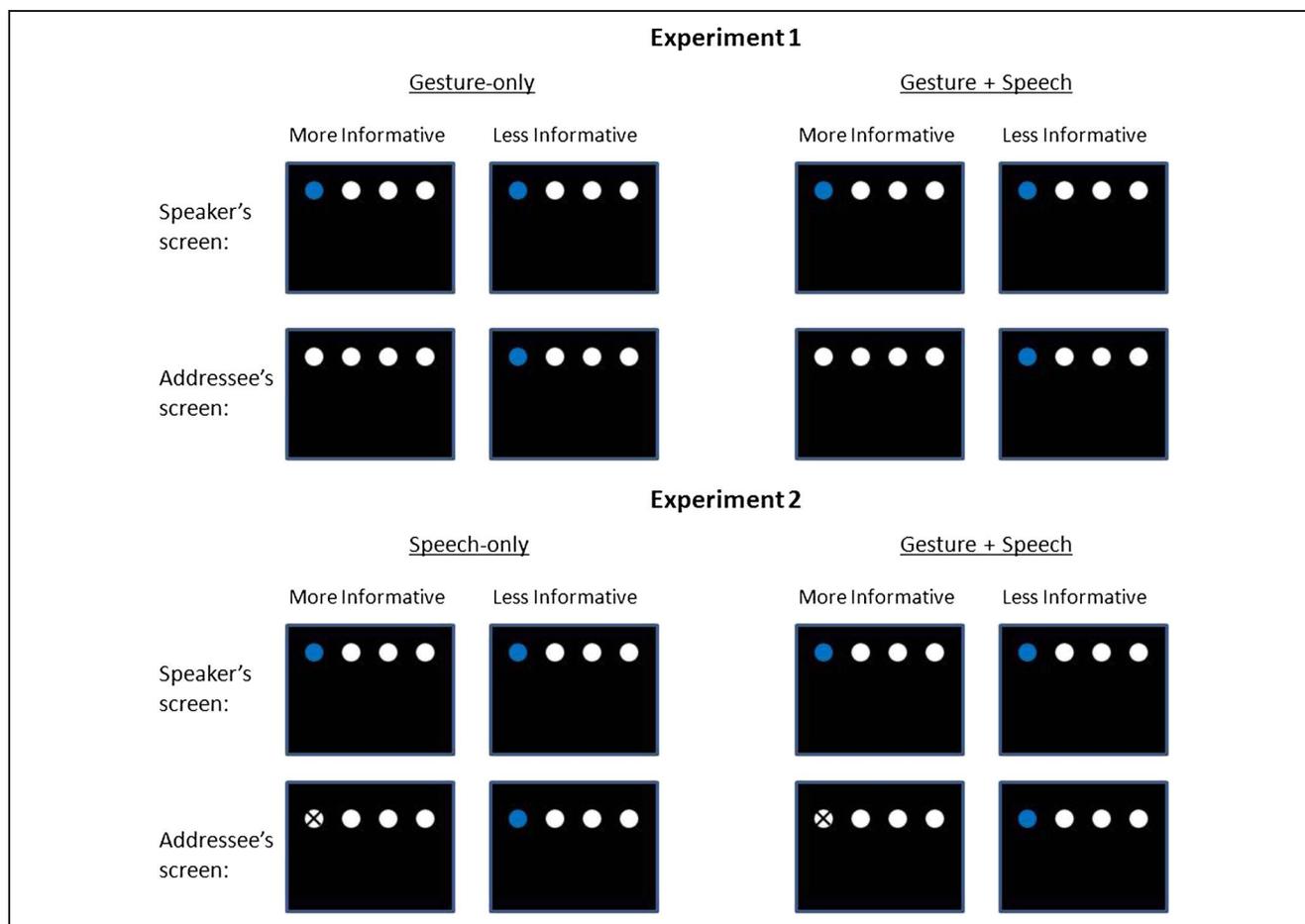
Each trial started with a fixation cross, displayed for 500 msec, followed by the presentation of four white circles. After a jittered period of 500–1000 msec, one of the circles turned yellow or blue. At this point, the participant was allowed to release her finger from a button, pointed to the blue or yellow circle, and (in the gesture + speech blocks), in speech, referred to the color of the circle. The experiment consisted of 16 blocks of 20 trials each. Every condition in the experiment was represented by four blocks. The order of presentation of blocks was counterbalanced across participants. In half of the trials a circle lit up yellow, in the other half it lit up blue. Each block of 20 trials consisted of 10 circles lighting up yellow and 10 lighting up blue, equally distributed over the four circles and the four conditions throughout the experiment in a randomized way.

### Procedure

On arrival of the participant, the experimenter explained that a second participant (i.e., the confederate addressee) would perform a behavioral task on the basis of the participant's gestures. The experimenter showed the participant the computer and form to be used by the addressee and demonstrated that the addressee could view the participant's pointing gestures referring to circles on the computer screen.

To keep participants motivated, it was emphasized that they were in a joint activity with the addressee and that the success of this joint activity depended on how well they worked together. The participant was then seated in a comfortable chair in the experiment room. The height of the screen was adjusted to the height of the eyes of the participant. The button used by the participant was placed at the height of the participant's elbow, 23 cm in front of the participant calculated from the vertical axis corresponding to the position of the participant's eyes. Participants were instructed to always rest their finger on this button, except when making the pointing gesture, which allowed calculating the duration and onset of the pointing gesture. An active, wireless sensor was placed on the participant's right index finger nail to allow for motion tracking of the pointing movements. Participants' EEG was recorded continuously throughout the experiment (see below).

After montage of the motion tracking sensor, the experimenter picked up the confederate addressee. The addressee was shown the room in which the participant performed the task, greeted the participant, and was seated in a chair in front of a computer in a room adjacent to the participant's room. Thirty-two test items (eight per condition) preceded the main experiment as a practice set. Participants received specific instructions to point with



**Figure 2.** Overview of the design of the Experiments 1 and 2.

or without speech before each block. In addition, before each block, the participant was instructed whether the addressee could also see the same circles light up at the back screen or not during that block. Participants were asked to only move their hand and arm when pointing. During the experiment, participants were allowed to have a short break after every fourth block. Before and during the experiment, the communication between experimenter and addressee was minimal and fully scripted to be consistent across participants. The addressee provided no feedback to the participant during the experiment. After the experiment, the addressee was thanked for participation and left the room. After filling out a post-test questionnaire, participants were debriefed, financially compensated, and thanked for participation. The results of the post-test questionnaire revealed that all participants thought the confederate addressee was another (naive) participant who performed well on his task.

#### *Kinematic and Speech Recording and Analysis*

Behavioral and kinematic data were acquired throughout the experiment using experimental software (Presentation, Neurobehavioral Systems, Inc., Berkeley, CA) and a 60-Hz motion tracking system and DTrack2 tracking

software (both from Advanced Realtime Tracking, Weilheim, Germany). In line with previous work (Chu & Hagoort, 2014; Levelt et al., 1985), we focused on different kinematic aspects of the pointing movements, including the gesture initiation time, the stroke duration, the apex time, the hold duration, the incremental distance traveled by the pointing finger, and the velocity of the movement. Praat software (version 5.2.46; Boersma & Weenink, 2009) was used to calculate offline the speech duration and the maximum loudness and mean loudness of speech. Table 1 gives an overview of how the kinematic and speech-related dependent variables were defined and calculated (cf. Chu & Hagoort, 2014; Levelt et al., 1985).

#### *Electrophysiological Recording and Analysis*

Throughout the experiment, the participant's EEG was recorded continuously from 59 active electrodes (Brain Products, Munich, Germany) held in place on the scalp by an elastic cap (Neuroscan, Singen, Germany). In addition to the 59 scalp sites, three external electrodes were attached to record participants' 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). 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 $\Omega$ . The continuous EEG was recorded with a sampling rate of 500 Hz, a low cutoff filter of 0.01 Hz and a high cutoff filter of 200 Hz. EEG was filtered offline (high pass at 0.01 Hz and low pass at 40 Hz). All electrode sites were referenced online to the electrode placed over the left mastoid and re-referenced offline to the average of the right and left mastoids.

Markers were sent from the computer presenting the stimuli to the computer recording the EEG, at light onset and at gesture initiation. Using Brain Vision Analyzer software (Brain Products, Munich, Germany), ERPs were time-locked to light onset (i.e., stimulus-locked) and to gesture initiation (i.e., the onset of the pointing gesture;

henceforth called “gesture-locked”). In the stimulus-locked ERPs, the 100-msec prestimulus period was used as a baseline. In the gesture-locked ERPs, the period 700–600 msec before gesture initiation was used as a baseline, because this time window reliably preceded stimulus onset (see gesture initiation time in Table 2), such that the gesture-locked ERP would globally reflect the time between stimulus onset and gesture initiation. Note that in both the stimulus-locked and the gesture-locked ERPs we thus look at the activity preceding the onset of the gesture. Trials containing muscular artifacts were removed from further analysis (5.5% of the total stimulus-locked data set; 13.7% of the total gesture-locked data set). The amount of removed trials was similar across the different levels of the informativeness and modality

**Table 1.** Definition of the (Behavioral) Kinematic and Speech-related Dependent Variables in Experiments 1 and 2, as Calculated for Each Experimental Trial

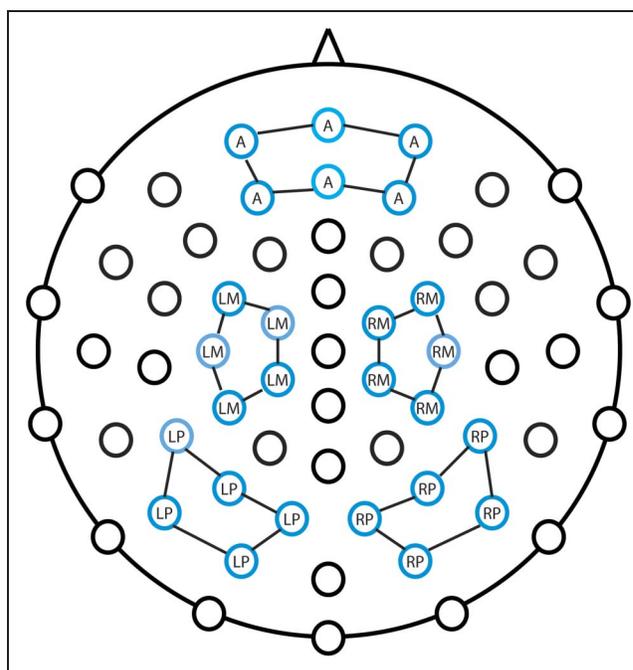
<i>Variable</i>	<i>Definition</i>
<i>Kinematic Dependent Variables</i>	
Gesture initiation time (msec)	Gesture onset – Light onset
Stroke duration (msec)	Gesture apex – Gesture onset
Apex time (msec)	Gesture apex – Light onset
Hold duration (msec)	Retraction time – Gesture apex
Incremental distance (cm)	The amount of distance traveled by the pointing index finger between Gesture onset and Gesture apex
Velocity (cm/sec)	Apex time/Incremental distance
<i>Speech-related Dependent Variables</i>	
Speech duration (msec)	Speech offset – Speech onset
Speech onset time (msec)	Speech onset – Light onset
Synchronization time (msec)	Speech onset time – Apex time
Maximum loudness (db)	The maximum loudness of speech during an utterance
Mean Loudness (db)	The average loudness of speech across an utterance
<i>Other Variables Used in Calculations</i>	
Light onset	The moment in time a circle lit up
Gesture onset	The moment in time the participant’s finger left the button in order to point
Gesture apex	The moment in time where the pointing index finger was at least 7 cm from the button and moved forward only less than 2 mm for two consecutive samples
Speech onset	The moment in time the participant started speaking
Speech offset	The moment in time the participant stopped speaking
Retraction time	The moment in time where the pointing index finger moved back in the direction of the button for at least 2 mm in two consecutive samples

**Table 2.** Overview of the Behavioral Results per Condition in Experiment 1

Condition	<i>GIT</i>	<i>Stroke*</i>	<i>Apex*</i>	<i>Dist</i>	<i>Velocity*</i>	<i>Hold*</i>
<i>More Informative</i>						
Gesture-only	534 (21)	834 (30)	1368 (42)	51 (1)	38.5 (1)	1252 (135)
Gesture + speech	550 (22)	840 (27)	1389 (39)	51 (1)	37.8 (1)	1219 (121)
<i>Less Informative</i>						
Gesture-only	532 (22)	819 (29)	1351 (41)	51 (1)	39.0 (1)	1138 (116)
Gesture + speech	541 (24)	826 (27)	1367 (40)	51 (1)	38.5 (1)	1149 (106)
Condition	<i>SpeechDur</i>	<i>SOT*</i>	<i>Sync</i>	<i>Max_Loudness</i>	<i>Mean_Loudness</i>	
<i>More Informative</i>						
Gesture-only						
Gesture + speech	1167 (35)	1385 (65)	4 (54)	82.0 (1)	70.8 (1)	
<i>Less Informative</i>						
Gesture-only						
Gesture + speech	1155 (36)	1351 (66)	16 (54)	82.2 (1)	70.8 (1)	

Duration in msec is displayed for gesture initiation time (*GIT*), stroke duration (*Stroke*), apex time (*Apex*), hold duration (*Hold*), speech duration (*SpeechDur*), speech onset time (*SOT*), and synchronization time (*Sync*). Furthermore, the incremental distance in cm (*Dist*), velocity in cm/sec (*Velocity*), and the maximum and mean loudness of the speech (*Max\_Loudness* and *Mean\_Loudness*) in db are provided. The *SEM* is indicated in parentheses. An asterisk next to a variable's name indicates a significant main effect of informativeness in the analysis.

factors. Subsequently, independent component analysis (ICA) was used to correct for ocular artifacts (extended infomax procedure; cf. Lee, Girolami, & Sejnowski, 1999). The mean amplitudes of the ERP waveforms for each condition per subject were entered into repeated-measures ANOVAs in a time window analysis of 100-msec time windows after stimulus onset (0–400 msec) or before gesture initiation (–600 msec until gesture onset), respectively. A subset of five ROIs was selected for the analyses (see Figure 3) based on previous, related work outlined in the Introduction. An anterior ROI was selected on the basis of the findings in Henderson et al. (2002). A potential modulation of the readiness potential as a function of our informativeness manipulation would be reflected in an effect over left central but not right central electrode sites, because all participants were right-handed and pointing with their right index finger. Therefore, a left middle ROI and a right middle ROI were selected. Finally, a possible P300 (P3b) effect would occur in posterior electrode sites, possibly right-lateralized (Polich, 2007), which led to the selection of a left posterior ROI and a right posterior ROI. In summary, the ERP analyses thus contained the independent variables informativeness (more informative vs. less informative), modality (gesture-only vs. gesture + speech), and ROI (anterior, left middle, right middle, left posterior, right posterior). The Greenhouse and Geisser (1959) correction was applied when appropriate. Corrected degrees of freedom are reported.



**Figure 3.** Electrode montage. Five ROIs were used in the analysis of the electrophysiological data: anterior (A), left middle (LM), right middle (RM), left posterior (LP), and right posterior (RP).

## Results

### Behavioral Results

Trials in which the gesture initiation time was below 100 msec or above 2000 msec were considered errors and excluded from all analyses (0.7% of total data set). In addition, trials containing hesitations or errors in the participant's speech were removed from further analyses (0.2% of all data). Separate analyses of variance were performed for each dependent variable with Informativeness (more informative vs. less informative) and Modality (gesture-only vs. gesture + speech) as within-subject factors. The analyses performed on the gesture initiation time and the incremental distance did not yield any significant main or interaction effects (all  $p$ s > .05).

The analysis of the stroke duration yielded a significant main effect of Informativeness,  $F(1, 23) = 10.97, p = .003, \eta_p^2 = .32$ . This effect reflected that the duration of the stroke was significantly longer in the more informative condition ( $M = 837$  msec) than in the less informative condition ( $M = 823$  msec). No significant main effect of Modality was found. There was no significant interaction between the two factors.

The analysis of the apex time showed a significant main effect of Informativeness,  $F(1, 23) = 8.15, p = .009, \eta_p^2 = .26$ . This effect denoted that the apex was reached significantly later in the more informative condition ( $M = 1379$  msec) than in the less informative condition ( $M = 1359$  msec). No significant main effect of Modality was found. There was no significant interaction between the two factors.

The analysis on the mean velocity yielded a significant main effect of Informativeness,  $F(1, 23) = 5.75, p = .025, \eta_p^2 = .20$ . The velocity of the pointing gesture was significantly lower in the more informative condition ( $M = 38.2$  cm/s) than in the less informative condition ( $M = 38.7$  cm/s). Again, no significant main effect of Modality or interaction between the two factors was found.

The analysis performed on the hold duration yielded a significant main effect of Informativeness,  $F(1, 23) = 10.17, p = .004, \eta_p^2 = .31$ . The hold duration was significantly longer in the more informative condition ( $M = 1235$  msec) compared with the less informative condition ( $M = 1143$  msec). No significant main effect of Modality was found and there was no significant interaction between the two factors.

An analysis on the speech onset time (gesture + speech conditions only) revealed a significant main effect of Informativeness,  $F(1, 23) = 6.79, p = .016, \eta_p^2 = .23$ . This effect reflected that the speech onset on average took place significantly later in the more informative condition ( $M = 1385$  msec) than in the less informative condition ( $M = 1351$  msec).

Finally, an analysis on the synchronization time in the gesture + speech conditions did not show a significant main effect of Informativeness ( $p = .16$ ), indicating that the onset of the speech and the apex of the gesture were aligned similarly across conditions and independently

from the informativeness of the gesture. The maximum loudness and mean loudness of the speech did not differ significantly across informativeness, nor did the speech duration (all  $p$ s > .05).

In summary, participants prolonged the duration of the stroke and the hold phase of their pointing gesture to be more informative, which led to a gesture with a lower velocity and delayed the moment at which the apex was reached. Table 2 summarizes all behavioral results from Experiment 1.

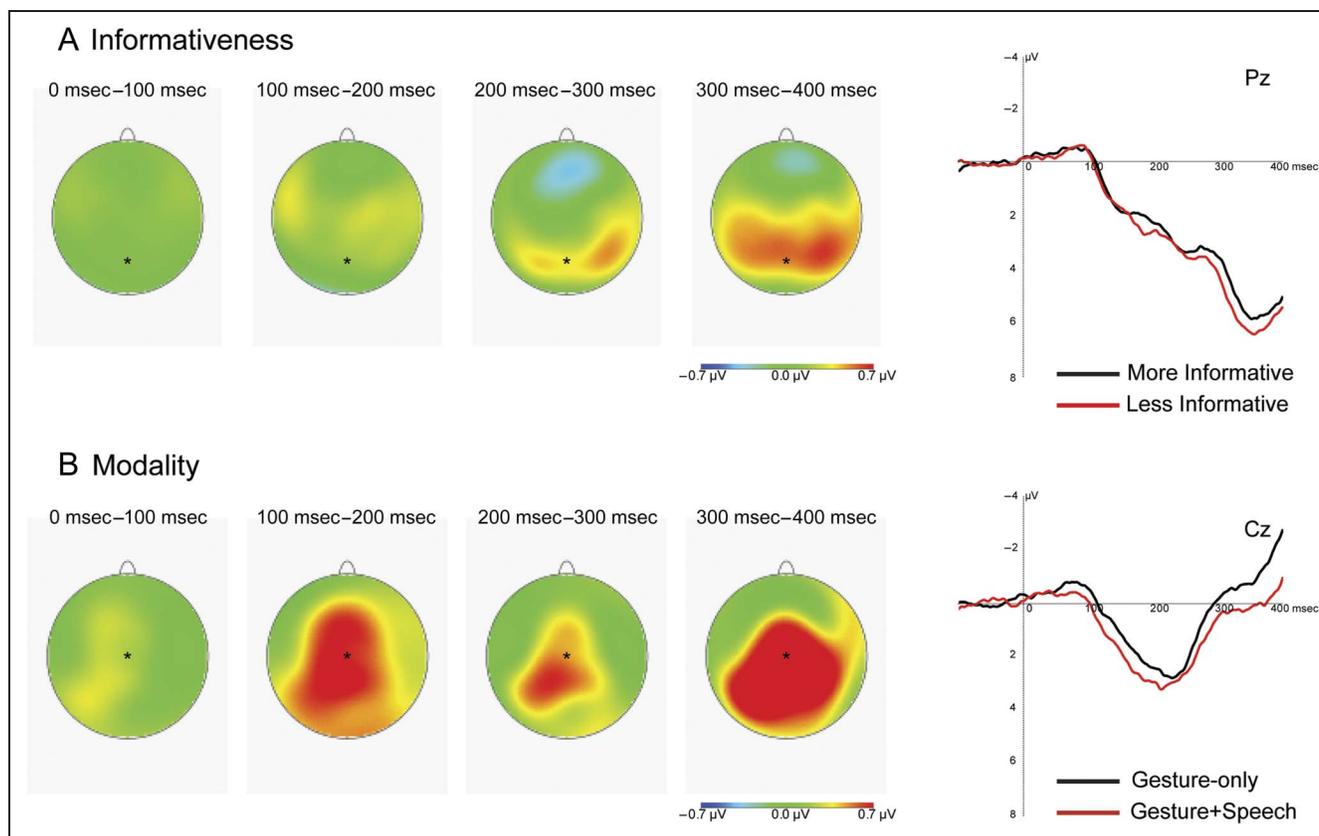
### Electrophysiological Results

Trials defined as errors or outliers in the behavioral analyses were also excluded from the ERP analyses. Separate ERPs were computed for each condition in the experiment. By-participant analyses (both stimulus-locked and gesture-locked) were performed with Informativeness, Modality, and ROI as independent variables. Only significant effects at the 5% level are reported, unless explicitly stated otherwise.

**Stimulus-locked analysis.** The omnibus stimulus-locked analysis firstly revealed a significant Informativeness  $\times$  ROI interaction effect in time windows 200–300 msec,  $F(2, 52) = 4.58, p = .012, \eta_p^2 = .17$ , and 300–400 msec,  $F(2, 45) = 3.39, p = .044, \eta_p^2 = .13$  after stimulus onset. Follow-up analyses showed a significant main effect of Informativeness in the 300–400 msec time window in the right posterior ROI only,  $F(1, 23) = 5.53, p = .028, \eta_p^2 = .19$ . This effect reflected a significantly more positive ERP wave for the more informative condition compared with the less informative condition. We will refer to this effect as a P300 or P3b effect (cf. Polich, 2007). There was a trend toward a similar effect of Informativeness in the 200–300 msec time window in the right posterior ROI,  $F(1, 23) = 3.14, p = .090, \eta_p^2 = .12$ .

Second, the omnibus analysis revealed a significant main effect of Modality in the 100–200 msec time window,  $F(1, 23) = 6.27, p = .020, \eta_p^2 = .21$ , the 200–300 msec time window,  $F(1, 23) = 4.77, p = .039, \eta_p^2 = .17$ , and the 300–400 msec time window,  $F(1, 23) = 11.17, p = .003, \eta_p^2 = .33$ . These main effects of Modality reflected a significantly more positive ERP wave for the gesture + speech condition compared with the gesture-only condition. No Informativeness  $\times$  Modality interaction effect was found in any time window (all  $F$ s  $\leq 1$ ). Figure 4 graphically presents the stimulus-locked ERP results.

**Gesture-locked analysis.** The omnibus analysis locked to the onset of the gesture revealed a significant Informativeness  $\times$  ROI interaction effect in the  $-100$  to  $0$  msec time window, that is, directly preceding gesture initiation,  $F(3, 66) = 6.09, p = .001, \eta_p^2 = .21$ . Follow-up analyses yielded a significant main effect of Informativeness in this time window in the anterior ROI only,  $F(1, 23) = 5.03, p = .035, \eta_p^2 = .18$ . This effect reflected a significantly more negative ERP wave for the less informative condition compared with the



**Figure 4.** Grand-averaged waveforms and topographic plots corresponding to the voltage difference between conditions in subsequent time windows in the stimulus-locked ERP analysis in Experiment 1 for (A) the main effect of Informativeness (collapsed over modality) and (B) the main effect of Modality (collapsed over informativeness). The electrode site used for the waveforms is indicated in the corresponding topographic plots.

more informative condition (see Figure 5). We will refer to this effect as a frontal marker of informativeness/communicative intent. No such effect was found in any other ROI (all  $F_s \leq 1$ ). No main effect of Modality or Informativeness  $\times$  Modality interaction effects were found (all  $F_s < 1$ ).

## Discussion

Experiment 1 revealed behavioral and electrophysiological correlates of communicative intent in the planning and production of index finger pointing gestures.

Behaviorally, participants prolonged the duration of the stroke of their pointing gesture to be more informative, which led to a gesture with a lower velocity and delayed the moment at which the apex was reached. In addition, the poststroke hold phase of the gesture was maintained for longer. The kinematic properties of participants' pointing gestures were not affected by the concurrent production of speech (in line with Chieffi et al., 2009), and similar kinematic effects of communicative intent were found in situations where people only used gesture to communicate compared to situations where speech and gesture were simultaneously produced. In addition, participants temporally aligned the onset of their deictic linguistic expression with the moment the pointing gesture reached its apex, regardless of whether

the gesture was more or less informative. No effect of participants' communicative intentions was found in the quality of the (largely informationally redundant) speech itself. We will discuss the theoretical implications of these findings in the General Discussion.

Neurophysiologically, the stimulus-locked ERPs showed a (parietal) P3b effect with smaller amplitude for the more informative condition, independent of modality. As outlined in the Introduction, P3b amplitude may be modulated by task-related cognitive demands that drive attentional resource allocation, such that its amplitude is smaller when a task requires greater amounts of attentional resources (Polich, 2007). Smaller amplitude of the stimulus-locked P3b in the more informative condition may therefore reflect that participants voluntarily allocated more attentional resources when planning a more informative gesture for their addressee. Furthermore, the gesture-locked waveforms showed a frontal marker of communicative intent directly preceding the onset of the pointing movement. As shown in Figure 5, they resembled the readiness potential (Kornhuber & Deecke, 1965) but had a clearly different distribution over the scalp (i.e., more anterior and less lateralized). The frontal locus of this effect is reminiscent of the locus of electrophysiological findings in infant studies tapping into developing joint attentional mechanisms related to pointing in infancy (e.g., Henderson et al.,

2002). More generally, in the planning and production of pointing gestures, frontal effects have been interpreted as reflecting the involvement of intention-related “mentalizing” networks (e.g., Brunetti et al., 2014). Our effect modulating the readiness potential (see Figure 5) thus suggests an interaction between planning a motor program and activation of the mentalizing network (Amodio & Frith, 2006).

It is an open question to what extent the kinematic and electrophysiological findings obtained in Experiment 1 are specific to situations in which the gesture is carrying the main informational burden in a multimodal speech act. It is possible that whenever speech itself is informative enough to single out a referent, people no longer design the kinematics of their concomitant gesture to be maximally informative (Cooperrider, 2011; Bangerter, 2004; see also Enfield et al., 2007), although they may still have a similar communicative intention. We tested this possibility in a second experiment, presented below, in which participants crucially had to refer to the color of the circle that lit up, and the addressee’s task was to note down the color of the circle. Because the color and not the location of the circle was now the important aspect of the stimulus, in this case the speech modality and not the gestural modality carried the informational burden. This manipulation thus allowed us to investigate whether the extent to which people modify the kinematic characteristics of their pointing gesture on the basis of their communicative intentions is dependent on

how they distribute the informational burden over two modalities (speech and gesture). In Experiment 2, therefore, the informativeness of the deictic act was thus manipulated in the speech modality, which was either paired with a redundant pointing gesture (bimodal condition) or not (unimodal condition).

Furthermore, Experiment 2 will show whether the intentional and attentional neurophysiological markers that we found in Experiment 1 are specific to cases where pointing gestures carry the main informational burden or whether they are modality-independent instead. A frontal speech-locked ERP effect of informativeness may suggest a common intention-related mechanism in the planning of both referential gesture and speech. Moreover, if the stimulus-locked P3b effect indeed reflects (voluntary) attentional resource allocation, it will be independent of whether participants’ task is to refer to the spatial location (as in Experiment 1) or color (as in Experiment 2) of the entity they point at and attend to.

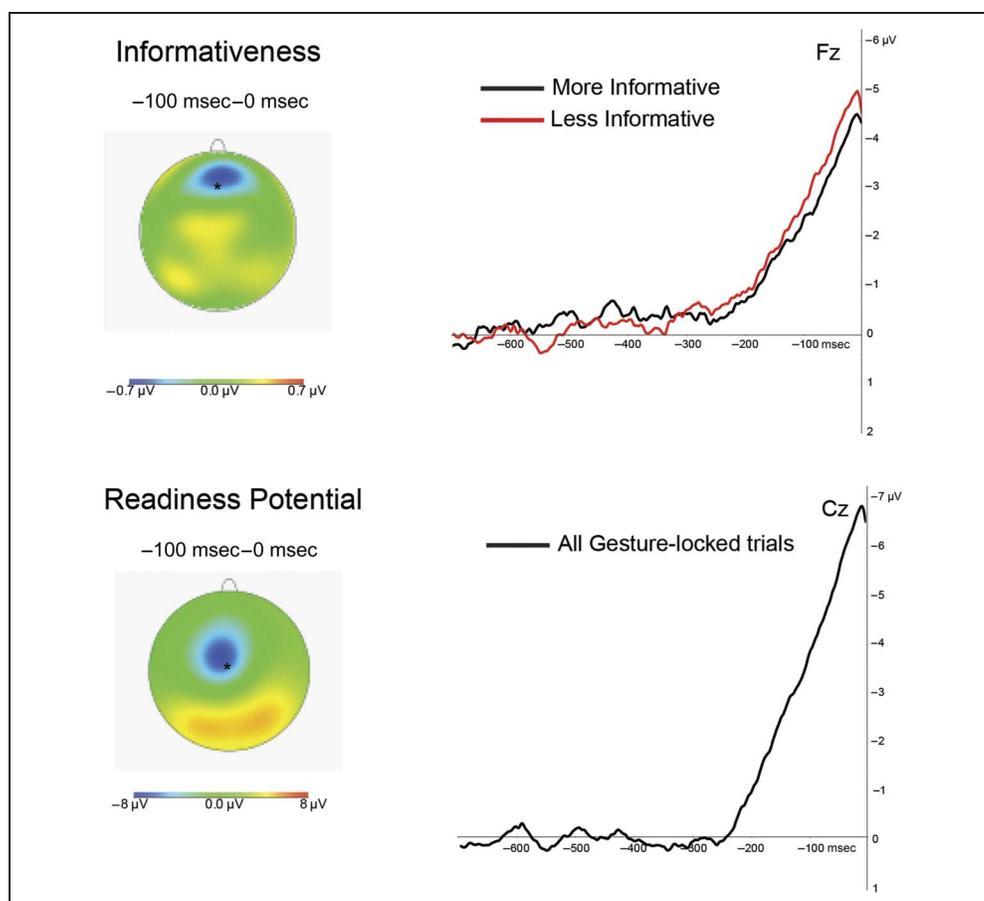
## EXPERIMENT 2

### Method

#### Participants

Twenty-four new participants (12 women; mean age = 21.7 years) matching the criteria from Experiment 1 took

**Figure 5.** Top: Grand-averaged waveforms and topographic plot corresponding to the voltage difference between conditions for the main effect of Informativeness (collapsed over modality) in the gesture-locked ERP analysis in Experiment 1 in the time window directly preceding gesture initiation in the anterior ROI. Bottom: The readiness potential and its locus over the scalp across all gesture-locked trials. The electrode site used for the waveforms is indicated in the topographic plots.



part in Experiment 2. Data from six additional participants were obtained but had to be discarded due to technical failure during the experiment or due to the presence of a large number of trials that contained movement artifacts. All participants provided written informed consent and were paid €20 for participation.

### Experimental Design and Setup

Similar to Experiment 1, stimuli were four white circles in a horizontal line on the top of the screen, mirroring four circles on the back screen. Each circle could light up in blue or yellow. Again, the addressee (the same confederate as in Experiment 1) looked at the back screen (providing the corresponding view of the four circles the participant was seeing) and the actual participant via a camera. On all trials, participants referred to the circle that lit up. In contrast with Experiment 1, the addressee noted on a paper form the color of the circle that lit up (and not the location). In addition, the addressee listened to the participant's speech via speakers in the addressee's room.

The informativeness of the speech (more informative vs. less informative) as well as the modality of the deictic act (speech-only vs. gesture + speech) were manipulated in a  $2 \times 2$  within-participants design. In the more informative condition, a circle turned blue or yellow only on the participants' screen but not on the back screen. To make the pointing gesture in the speech + gesture condition redundant, the location of the circle that lit up was marked by a cross in the more informative condition on the back screen only (see Figure 2). The participant's speech was the only source of information on which the addressee had to base his decision in determining the color of the circle referred to by the participant. In the less informative condition, the corresponding circles would light up on both the participant's and the addressee's screen. This rendered the participant's speech less informative in this condition, because the addressee saw the respective circle light up in either blue or yellow on the back screen at the same moment as the participant saw the corresponding, same color circle light up.

The modality factor was manipulated by having participants use either one or two modalities in referring to the circles. In speech-only blocks, when a circle lit up participants said *de blauwe cirkel* ("the blue circle") or *de gele cirkel* ("the yellow circle"), depending on the color of the circle, without producing a pointing gesture. In gesture + speech blocks, participants uttered the same phrase but now also produced an index finger pointing gesture toward the location of the circle that lit up. Note that, because on all trials the location of the circle was known by the addressee and because the location of the circle was irrelevant for the task performed by the addressee in Experiment 2, the gesture was never informative (neither in the more informative nor in the less informative blocks). The rationale for this was that we were interested in the possible effect of the mere pres-

ence of gesture as a second modality, independent from the informativeness of the deictic act that was manipulated separately in the speech modality. The trial structure was the same as in Experiment 1.

### Procedure

The experimental procedure was the same as in Experiment 1. Again, the results of the posttest questionnaire revealed that all participants thought the confederate addressee was another (naive) participant who performed well on their task.

### Kinematic, Electrophysiological, and Speech Recordings

The kinematic, electrophysiological, and speech recordings were done similarly to Experiment 1. EEG was recorded continuously, and ERPs were time-locked separately to light onset (i.e., stimulus-locked), gesture initiation (in the gesture + speech blocks; "gesture-locked"), and voice onset (in the speech-only blocks; "speech-locked"). The stimulus-locked preprocessing and ERP analyses were the same as in Experiment 1. The gesture-locked analysis was also the same as in Experiment 1 except for the absence of the modality factor due to gesture being produced only in gesture + speech blocks in this experiment. An additional analysis in 100 msec time windows preceding the speech was carried out on ERPs time-locked to speech onset during speech-only blocks. Separate analyses were carried out for 100 msec time windows during the 900-msec preceding speech onset. The 1000–900 msec time window preceding speech onset was used as a baseline period, because this time window reliably preceded speech onset time in the speech-only blocks (regardless of informativeness). Trials containing muscular artifacts were removed from further analysis (7.4% of the total stimulus-locked data set; 17.1% of the total gesture-locked data set; 8.2% of the total speech-locked data set). The amount of removed trials was similar across the different levels of informativeness and modality. Inspection of the EEG data confirmed that it was not feasible to further look into speech-locked ERPs in the gesture + speech blocks due to the concurrent pointing gesture creating movement artifacts before speech onset (gesture onset systematically preceded voice onset).

## Results

### Behavioral Results

Trials on which the gesture initiation time or the speech onset time was below 100 msec or above 2000 msec were considered errors and excluded from all analyses (0.5% of total data set). In addition, trials containing hesitations or errors in the participant's speech were removed from further analysis (0.3% of all data).

First, separate analyses of variance were performed on the speech duration and the speech onset time with Informativeness (more informative vs. less informative) and Modality (speech-only or gesture + speech) as within-subject factors. The analysis of the speech onset time revealed a significant main effect of Modality,  $F(1, 23) = 87.49, p = .001, \eta_p^2 = .79$ , with the speech onset being significantly later in the gesture + speech condition ( $M = 976$  msec) compared with the speech-only condition ( $M = 706$  msec). The analysis of the speech duration yielded a significant main effect of Modality,  $F(1, 23) = 5.74, p = .025, \eta_p^2 = .20$ , driven by the speech duration being significantly longer in the gesture + speech condition ( $M = 1111$  msec) compared with the speech-only condition ( $M = 1095$  msec).

Both the analysis of the maximum loudness of the speech and the analysis of the mean loudness of the speech showed a significant main effect of Modality,  $F(1, 23) = 16.55, p = .001, \eta_p^2 = .42$  and  $F(1, 23) = 8.73, p = .007, \eta_p^2 = .28$ , respectively. This indicated that participants spoke more loudly in the bimodal compared with the unimodal conditions. In all these analyses, no significant main effect of Informativeness was found, and there was no significant interaction between the two factors.

In the gesture + speech conditions, participants manually pointed at the circle on the screen while linguistically referring to it. Repeated-measures analyses of variance with Informativeness as the single within-subject factor were carried out on the same dependent variables as in Experiment 1.

The analysis of the stroke duration showed a significant main effect of Informativeness,  $F(1, 23) = 5.42, p = .029, \eta_p^2 = .19$ . This effect denoted that the duration of the stroke was significantly longer in the more informative condition ( $M = 707$  msec) than in the less informative condition ( $M = 698$  msec). Analyses of gesture initiation time, apex time, incremental distance, velocity, hold duration, and synchronization time did not yield any significant effect (all  $ps > .05$ ). Table 3 summarizes all behavioral results from Experiment 2.

### Electrophysiological Results

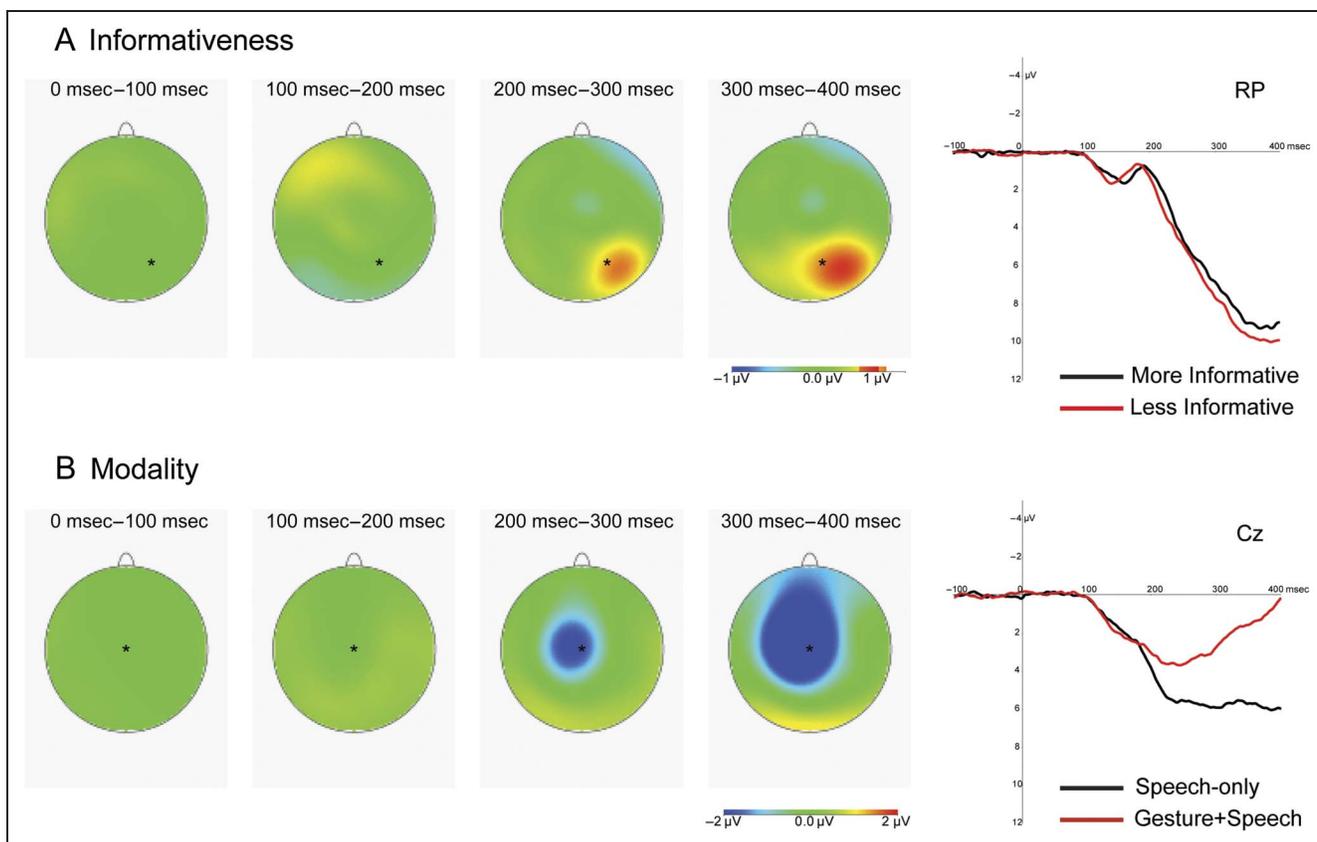
Trials defined as errors or outliers in the behavioral analyses were also excluded from the ERP analyses. Separate ERPs were computed for each condition in the experiment. By-participant analyses were performed with Informativeness and ROI as independent variables. In the stimulus-locked analysis, Modality was added as a factor. Only significant effects at the 5% level are reported, unless explicitly stated otherwise.

**Stimulus-locked analysis.** The omnibus stimulus-locked analysis firstly revealed a significant Informativeness  $\times$  ROI interaction effect in the 200–300 msec time window,  $F(2, 42) = 3.32, p = .049, \eta_p^2 = .13$ , and in the 300–400 msec time window,  $F(2, 40) = 4.37, p = .024, \eta_p^2 = .16$ . Follow-up analyses revealed that these interactions reflected a

**Table 3.** Overview of the Behavioral Results per Condition in Experiment 2

Condition	GIT	Stroke*	Apex	Dist	Velocity	Hold
<i>More Informative</i>						
Speech-only						
Gesture + speech	552 (27)	707 (24)	1259 (39)	42 (1)	34.3 (1)	592 (78)
<i>Less Informative</i>						
Speech-only						
Gesture + speech	548 (26)	698 (25)	1247 (38)	42 (1)	34.5 (1)	576 (76)
Condition	SpeechDur	SOT	Sync	Max_Loudness	Mean_Loudness	
<i>More Informative</i>						
Speech-only	1095 (39)	711 (31)		78.6 (1)	66.1 (1)	
Gesture + speech	1114 (43)	977 (46)	28 (3)	79.5 (1)	66.5 (1)	
<i>Less Informative</i>						
Speech-only	1095 (42)	702 (32)		78.5 (1)	65.8 (1)	
Gesture + speech	1108 (44)	976 (48)	27 (3)	79.2 (1)	66.5 (1)	

Duration in msec is displayed for gesture initiation time (GIT), stroke duration (Stroke), apex time (Apex), hold duration (Hold), speech duration (SpeechDur), speech onset time (SOT), and synchronization time (Sync). Furthermore, the incremental distance in cm (Dist), velocity in cm/sec (Velocity), and the maximum and mean loudness of the speech (Max\_Loudness and Mean\_Loudness) in db are provided. The SEM is indicated between parentheses. An asterisk next to a variable's name indicates a significant main effect of Informativeness in the analysis.

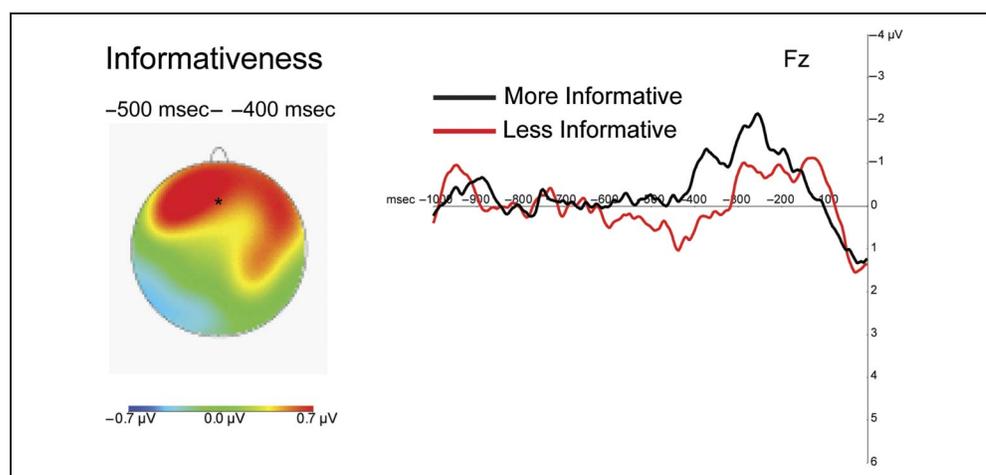


**Figure 6.** Grand-averaged waveforms and topographic plots corresponding to the voltage difference between conditions in subsequent time windows in the stimulus-locked ERP analysis in Experiment 2 for (A) the main effect of Informativeness (collapsed over modality) and (B) the main effect of Modality (collapsed over informativeness). The electrode site used for the waveforms is indicated in the corresponding topographic plots.

significant main effect of Informativeness in the right posterior ROI in the 200–300 msec time window,  $F(1, 23) = 8.87, p = .007, \eta_p^2 = .28$ , and in the 300–400 msec time window,  $F(1, 23) = 7.19, p = .013, \eta_p^2 = .24$ . We will again refer to this effect as a P300 or P3b effect (cf. Polich, 2007). A trend toward such a main effect of informativeness was found in the left posterior ROI in the 300–400 msec time window,  $F(1, 23) = 3.68, p = .068, \eta_p^2 = .14$ . No main effects of Informativeness were found in the other ROIs (all  $F$ s < 1).

Secondly, the omnibus analysis revealed a significant Modality  $\times$  ROI interaction effect in the 100–200 msec time window,  $F(3, 59) = 3.11, p = .040, \eta_p^2 = .12$ , in the 200–300 msec time window,  $F(2, 56) = 12.88, p = .001, \eta_p^2 = .36$ , and in the 300–400 msec time window,  $F(3, 66) = 24.73, p = .001, \eta_p^2 = .52$ . Follow-up analyses revealed a significant main effect of Modality that reached significance in the 200–300 msec time window in the left middle ROI only ( $p < .001$ ) before becoming significant

**Figure 7.** Grand-averaged waveforms and topographic plot corresponding to the voltage difference between conditions in the speech-locked ERP analysis in Experiment 2 for the main effect of Informativeness. The electrode site used for the waveforms is indicated in the corresponding topographic plot.



in the middle ROIs in the 300–400 msec time window as well (all  $p$ s < .01), but not in the two posterior ROIs (both  $F$ s < 1). Figure 6 shows the effects of informativeness and modality in the stimulus-locked analysis.

**Gesture-locked analysis.** The omnibus gesture-locked analysis showed a significant Informativeness  $\times$  ROI interaction effect in the time window 200–100 msec preceding gesture initiation,  $F(2, 52) = 4.15, p = .017, \eta_p^2 = .15$ . However, this effect did not reflect a significant main effect of Informativeness in any of the ROIs separately (all  $p$ s > .14).

**Speech-locked analysis.** The only significant effect in the omnibus analysis locked to speech onset was a significant Informativeness  $\times$  ROI interaction effect in the time window 500–400 msec preceding speech onset,  $F(2, 48) = 3.18, p = .049, \eta_p^2 = .12$ . This effect reflected a trend toward a main effect of Informativeness in the anterior ROI in this time window,  $F(1, 23) = 3.78, p = .064, \eta_p^2 = .14$ , which was absent in other ROIs (all  $F$ s < 2.4). The anterior finding reflected a more negative ERP wave for the more informative condition compared with the less informative condition (see Figure 7).

## Discussion

Experiment 2 revealed that the kinematic effects obtained in Experiment 1 were largely specific to situations in which gesture carried the main informational burden. Unlike in Experiment 1, in Experiment 2 no effects of informativeness were found in the time in which apex was reached or in the duration of the poststroke hold phase. However, a small effect of informativeness was found in the duration of the stroke of the gesture, with a longer stroke in case of more informative speech. Similar to Experiment 1, no effects of informativeness were found in the speech that participants produced. In comparison with the speech-only condition, the concurrent production of a gesture delayed the onset of speech, prolonged the speech duration, and enhanced its loudness. The stimulus-locked ERP data replicated the P300 effect obtained in Experiment 1, hence suggesting that this effect is modality-independent.<sup>1</sup> A trend toward a frontal ERP effect of informativeness was found preceding the onset of speech. We will discuss the theoretical implications of the findings of both experiments in General Discussion.

## GENERAL DISCUSSION

Two experiments were carried out to further our understanding of how our intentions shape our actions in the specific case of the planning and production of pointing gestures and speech to single out a visible referent, a core everyday human communicative act (Tomasello, 2008; Kita, 2003). Specifically, we investigated whether, and if so how, the kinematics of pointing gestures are shaped by one's communicative intentions and whether

this is modulated by the presence of concurrent speech. In addition, we explored the neural and cognitive mechanisms involved in the planning of communicative pointing gestures and speech.

Behaviorally, the first experiment showed that the kinematics of a pointing gesture vary as a function of the speaker–gesturer's communicative intent. Specifically, the duration of the stroke (and as such its velocity and the moment the apex is reached) was used to be informative. Presumably, this was done to be as precise as possible in pointing at a target, which could be achieved by pointing more slowly. An additional benefit would then be that the addressee would have more time to identify toward which referent the gesture was heading. In addition, participants prolonged the poststroke hold phase of their pointing gesture presumably to assure that the addressee had enough time to identify which referent she pointed to. The fact that people slow down their movement to be more informative generalizes to instrumental actions such as reach-to-grasp movements (Becchio et al., 2012) and communicative manual actions more broadly (Vesper & Richardson, 2014). Presumably, the duration of different subcomponents of the pointing gesture is not the only parameter people may use to communicate effectively, as previous work suggests that also the endpoint location and trajectory of a pointing gesture may be varied in relation to the location of the addressee (Cleret de Langavant et al., 2011).

In line with a previous study (Chieffi et al., 2009), in Experiment 1, the presence of speech as a second modality did not influence pointing gestures' kinematics. Other studies did find effects of the presence of speech on the kinematics of concurrently produced gestures. Gonseth et al. (2013) reported a slower gesture and a longer poststroke hold phase in cases where a pointing gesture was produced without speech compared to when it was produced with concurrent speech. Bernardis and Gentilucci (2006) found that participants shortened various movement phases of their symbolic gestures (e.g., a hand with protruding index finger moving from left to right meaning “NO”) when the gesture was produced with meaningful speech compared to when it was produced in isolation. An explanation for the absence of an influence of speech production on gesture kinematics in our study is that speech was purposefully kept very simple, noninformative, and repetitive across our first experiment. Increasing variation in speech (as in Gonseth et al., 2013) or adding a stronger symbolic component to it (as in Bernardis & Gentilucci, 2006) may lead to a stronger influence of speech on gesture kinematics.

The second experiment further specified that the influence of one's communicative intentions on the kinematics of one's pointing gestures is reduced in situations in which one's speech is carrying the informational burden in a multimodal referential act. For instance, participants did not use the duration of the hold phase of their gesture to be more informative in Experiment 2. Thus, when

speech suffices in transmitting the required information in a certain context, one does not need to exploit the kinematics of one's gesture to the same extent as when the gesture carries the informational burden. Nevertheless, a small modulation of the duration of the gesture's stroke as a function of participants' communicative intentions was found in both experiments (i.e., a longer stroke duration to be more informative). This confirms that speech and gesture are two highly intertwined modalities in the exophoric use of referential expressions and suggests that, even when speech is carrying the most relevant information in a multimodal referential act, one's more global communicative intentions also "flow" into the gestural modality. In contrast, participants neither exploited the loudness and duration of their speech to be more informative in Experiment 2 (see Willems et al., 2010, for a similar finding). One possible explanation is that in the current task the speech content itself was informative enough such that there was no need to change any acoustic or durational parameters to be more informative.

In both our experiments, participants temporally aligned the onset of their deictic linguistic expression with the moment the pointing gesture reached its apex, regardless of whether the gesture was more or less informative. This finding is in line with previous studies showing such temporal alignment of pointing and speech (e.g., Chu & Hagoort, 2014; McNeill, 1992; Levelt et al., 1985) and with models of speech and gesture production that underline the tight synchronization of speech and gesture (e.g., De Ruiter, 2000; Krauss et al., 2000). Previous experimental studies used artificial exogenous factors (such as the application of a load to a cord attached to a participant's wrist during the execution of a gesture; Levelt et al., 1985) to investigate its effects on speech–gesture synchronization. Here, we also show that when characteristics of the gesture vary for endogenous reasons (i.e., communicative intentions), the temporal synchrony between speech and pointing gestures is maintained.

In general, our results fit well with models of speech and gesture production that allow for a role of the speaker–gesturer's communicative intent in modulating the exact form of a gesture, such as the Sketch model (De Ruiter, 2000) and the Interface model (Kita & Özyürek, 2003). However, these models do not specify the exact sub-components of pointing gestures that people may vary on the basis of their communicative intentions. Our results suggest that duration (and as such the velocity of the stroke and the moment apex is reached) is a free parameter that people use in the execution of their pointing gestures and further specify in which specific components (i.e., stroke or poststroke hold) of the gesture duration is indeed varied. Even when speech is carrying the informational burden in a multimodal referential act, people's communicative intentions may lead to such use of the gesture's movement duration, as evidenced in Experiment 2. Our data cannot be explained by models of speech and gesture production that question whether the speaker's

communicative intent plays a role in shaping the form of a gesture (e.g., Krauss et al., 2000).

Neurophysiologically, we observed in both experiments a stimulus-locked P3b effect preceding the production of gesture and/or speech. We argued that P3b amplitude may be modulated by task-related cognitive demands that drive attentional resource allocation, such that its amplitude is smaller when a task requires greater amounts of attentional resources (cf. Polich, 2007). Smaller amplitude of the stimulus-locked P3b in the more informative conditions in Experiment 1 may therefore reflect that participants voluntarily (Kok, 2001) allocated more attentional resources to the task when planning a more informative gesture for their addressee, independent of whether they concurrently produced speech. Experiment 2 clarified that this effect is not specific to the planning of pointing gestures and also generalizes to situations in which referential speech is planned to describe a referent for one's addressee. This finding also confirms that the effect does not index differential visual attention paid to the spatial location or physical properties (e.g., color) of a referent, but rather in our study reflects the allocation of domain-general attentional resources that may be used to successfully plan an action on the basis of one's (communicative) intentions.

The gesture-locked frontal ERP marker of communicative intent directly preceding the onset of the pointing gesture was specific to the case where the gesture carried the informational burden (i.e., Experiment 1). The locus of this effect modulating the readiness potential is reminiscent of a previous study investigating pointing by infants, which also identified a frontocentral marker of communicative intent measured using EEG (Henderson et al., 2002). Several other studies have also linked frontal effects in ERPs to "mentalizing" or theory-of-mind-related activations (e.g., Liu, Sabbagh, Gehring, & Wellman, 2004; Sabbagh, 2004) and recent neuroimaging studies relate activity in neuronal structures in frontal cortex to the mentalizing involved in the production and comprehension of communicative pointing (e.g., Brunetti et al., 2014). The fact that our effect reflects a modulation of the readiness potential (Kornhuber & Deecke, 1965) over frontocentral areas suggests an interaction between planning the execution of a motor program and activation of the mentalizing network (Willems et al., 2010; Van Overwalle & Baetens, 2009; Amodio & Frith, 2006). In summary, these findings underline that both intentional and modality-independent attentional mechanisms are active when one plans the execution of a communicative, referential pointing gesture for an addressee.

Finally, a speech-locked trend toward an effect of participants' communicative intentions was found 500–400 msec preceding the onset of their speech. Interestingly, it had an opposite directionality compared to the frontal gesture-locked effect of participants' communicative intentions. Future research is needed to verify whether the speech finding is robust. Note that, on the

basis of models of speech production, the timing of the effect is where one would expect an influence of one's intentions in the speech production process (e.g., Indefrey & Levelt, 2004). The current study shows that it is worthwhile and feasible to investigate the intentions behind speech (and gesture) production, a crucial component of the speech production process.

To conclude, we have shown that people shape the exact kinematics of their pointing gesture as a function of their specific communicative intentions, in tight temporal alignment with their speech, and particularly when the gestural modality carries the informational burden. Furthermore we have shown that both intentional and modality-independent attentional neural mechanisms are active in planning the execution of a communicative pointing gesture. These findings contribute to a better understanding of the complex interplay between action, attention, intention, and language in the core human communicative act of planning and producing referential utterances using speech and gesture.

## Acknowledgments

We thank Albert Russel for assistance in setting up the experiments, and Charlotte Poulisse for help in data collection.

Reprint requests should be sent to David Peeters, MPI, P.O. Box 310, NL-6500 AH Nijmegen, The Netherlands, or via e-mail: david.peeters@mpi.nl.

## Note

1. An additional repeated-measures ANOVA with the between-subject factor Experiment (2: Experiment 1, Experiment 2) and the within-subject factors ROI (2: left posterior, right posterior) and Time Window (2: 200–300 msec, 300–400 msec) was performed on the P3b effect (less informative average ERP minus more informative average ERP), calculated for each subject in both time windows. This analysis did not show any significant main or interaction effect of Experiment, indicating that the size of the P3b effect did not differ statistically across the two experiments.

## REFERENCES

- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268–277.
- Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science*, 15, 415–419.
- Bara, B. G. (2010). *Cognitive pragmatics. The mental process of communication*. Cambridge, MA: The MIT Press.
- Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly of Behavior and Development*, 21, 205–226.
- Becchio, C., Manera, V., Sartori, L., Cavallo, A., & Castiello, U. (2012). Grasping intentions: From thought experiments to empirical evidence. *Frontiers in Human Neuroscience*, 6, 117.
- Bernardis, P., & Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologia*, 44, 178–190.
- Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 5.1.05) [Computer program].
- Brunetti, M., Zappasodi, F., Marzetti, L., Perrucci, M. G., Cirillo, S., Romani, G. L., et al. (2014). Do you know what I mean? Brain oscillations and the understanding of communicative intentions. *Frontiers in Human Neuroscience*, 8, 36.
- Butterworth, G. (2003). Pointing is the royal road to language for babies. In S. Kita (Ed.), *Pointing. Where language, culture, and cognition meet* (pp. 9–33). Hillsdale, NJ: Erlbaum.
- Campisi, E., & Özyürek, A. (2013). Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. *Journal of Pragmatics*, 47, 14–27.
- 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*, 63, 1–174.
- Chieffi, S., Secchi, C., & Gentilucci, M. (2009). Deictic word and gesture production: Their interaction. *Behavioural Brain Research*, 203, 200–206.
- Chu, M., & Hagoort, P. (2014). Synchronization of speech and gesture: Evidence for interaction in action. *Journal of Experimental Psychology: General*, 143, 1726–1741.
- Clark, H. H. (2003). Pointing and placing. In S. Kita (Ed.), *Pointing. Where language, culture, and cognition meet* (pp. 243–268). Hillsdale, NJ: Erlbaum.
- Cleret de Langavant, L., Remy, P., Trinkler, I., McIntyre, J., Dupoux, E., Berthoz, A., et al. (2011). Behavioral and neural correlates of communication via pointing. *PLoS One*, 6, e17719.
- Cooperrider, K. (2011). *Reference in action: Links between pointing and language*. Doctoral dissertation, University of California, San Diego.
- Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, 25, 141–168.
- De Ruiter, J. P. (1998). *Gesture and speech production*. Doctoral dissertation, University of Nijmegen, The Netherlands.
- De Ruiter, J. P. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and gesture* (pp. 284–311). Cambridge: Cambridge University Press.
- Enfield, N. J., Kita, S., & De Ruiter, J. P. (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics*, 39, 1722–1741.
- Gerwing, J., & Bavelas, J. (2004). Linguistic influences on gesture's form. *Gesture*, 4, 157–195.
- Gonseth, C., Vilain, A., & Vilain, C. (2013). An experimental study of speech/gesture interactions and distance encoding. *Speech Communication*, 55, 553–571.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95–112.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics: Speech Acts* (Vol. 3, pp. 41–58). New York, NY: Academic Press.
- Henderson, L. M., Yoder, P. J., Yale, M. E., & McDuffie, A. (2002). Getting the point: Electrophysiological correlates of protodeclarative pointing. *International Journal of Developmental Neuroscience*, 20, 449–458.
- Holler, J., & Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*, 26, 4–27.
- Holler, J., & Wilkin, K. (2011). An experimental investigation of how addressee feedback affects co-speech gestures accompanying speakers' responses. *Journal of Pragmatics*, 43, 3522–3536.

- Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*, 101–144.
- Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*, 367–371.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kita, S. (2003). *Pointing. Where language, culture, and cognition meet*. Hillsdale, NJ: Erlbaum.
- Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*, 16–32.
- Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, *38*, 557–577.
- Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passive Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, *284*, 1–17.
- Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261–283). Cambridge, UK: Cambridge University Press.
- Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, *11*, 417–441.
- Levelt, W. J., Richardson, G., & La Heij, W. (1985). Pointing and voicing in deictic expressions. *Journal of Memory and Language*, *24*, 133–164.
- Liu, D., Sabbagh, M. A., Gehring, W. J., & Wellman, H. M. (2004). Decoupling beliefs from reality in the brain: An ERP study of theory of mind. *NeuroReport*, *15*, 991–995.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- Melinger, A., & Levelt, W. J. (2004). Gesture and the communicative intention of the speaker. *Gesture*, *4*, 119–141.
- Moore, C., & D'Entremont, B. (2001). Developmental changes in pointing as a function of attentional focus. *Journal of Cognition and Development*, *2*, 109–129.
- Mundy, P., Card, J., & Fox, N. (2000). EEG correlates of the development of infant joint attention skills. *Developmental Psychobiology*, *36*, 325.
- 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.
- Pierno, A. C., Tubaldi, F., Turella, L., Grossi, P., Barachino, L., Gallo, P., et al. (2009). Neurofunctional modulation of brain regions by the observation of pointing and grasping actions. *Cerebral Cortex*, *19*, 367–374.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*, 2128–2148.
- Sabbagh, M. A. (2004). Understanding orbitofrontal contributions to theory-of-mind reasoning: Implications for autism. *Brain and Cognition*, *55*, 209–219.
- Sartori, L., Becchio, C., Bara, B. G., & Castiello, U. (2009). Does the intention to communicate affect action kinematics? *Consciousness and Cognition*, *18*, 766–772.
- Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition*, *119*, 242–252.
- Southgate, V., Van Maanen, C., & Csibra, G. (2007). Infant pointing: Communication to cooperate or communication to learn? *Child Development*, *78*, 735–740.
- Sperber, D., & Wilson, D. (1995). *Relevance. Communication & Cognition* (2nd ed.). Oxford, UK: Blackwell.
- Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA: MIT Press.
- Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, *78*, 705–722.
- Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *Neuroimage*, *48*, 564–584.
- Vesper, C., & Richardson, M. (2014). Strategic communication and behavioral coupling in asymmetric joint action. *Experimental Brain Research*, *232*, 2945–2956.
- Willems, R. M., de Boer, M., de Ruiter, J. P., Noordzij, M. L., Hagoort, P., & Toni, I. (2010). A dissociation between linguistic and communicative abilities in the human brain. *Psychological Science*, *21*, 8–14.