MAX PLANCK INSTITUTE FOR HUMAN COGNITIVE AND BRAIN SCIENCES LEIPZIG

Douglas Weinbrenner: Abstract pointing: ERP and behavioral evidence for its role in reference tracking. Leipzig: Max Planck Institute for Human Cognitive and Brain Sciences, 2017 (MPI Series in Human Cognitive and Brain Sciences; 187)

Abstract pointing

ERP and behavioral evidence for its role in reference tracking

Impressum

Max Planck Institute for Human Cognitive and Brain Sciences, 2017

Druck: Sächsisches Druck- und Verlagshaus Direct World, Dresden

Titelbild: © Douglas Weinbrenner, 2017

ISBN 978-3-941504-73-8

Abstract pointing

ERP and behavioral evidence for its role in reference tracking

Von der Fakultät für Biowissenschaften, Pharmazie und Psychologie
der Universität Leipzig
genehmigte

DISSERTATION

zur Erlangung des akademischen Grades
Doctor rerum naturalium
Dr. rer. nat.

vorgelegt
von
Dipl.-Psych. J. E. Douglas Weinbrenner
geboren am 09. März 1978 in Köln

Dekan:	Prof. Dr. Tilo Pompe
Gutachter:	Prof. Dr. Dr. h.c. Angela D. Friederici
	Professor Sotaro Kita

Tag der Verteidigung: 27. April 2017

«Je sers la science et c'est ma joie.»

Motto of Basile in *Léonard* by Turk & Bob de Groot

Acknowledgements

Professor Friederici gave me the opportunity to conduct my thesis in a most stimulating environment. I thank her and Professor Kita for acting as a reviewer after… well, after what could be called a long inter-trial interval.

Thom Gunter provided me with continued supervision and played an essential role during all phases of this project.

Christian Obermeier has been a great gesture colleague.

Thank you to my wife.

Contents

List of Figures

List of Tables

Part I

General Introduction

1 Brief introduction and outline

"As the tongue speaketh to the ear so the gesture speaketh to the eye." According to Francis Bacon (2005) these words go back to King James IV./I. and, indeed, it is a compelling idea that gestures carry information to the recipient. But is this idea of gestural communication correct?

For some gesture types there is no reason to doubt that. When your supervisor shows you a thumbs-up gesture, you know that you did something right. When you are looking for your keys and your friend tries to support you by pointing to the sideboard and saying, "They're over there", it is the gesture that provides you with the crucial information.

Then there are gestures, however, where the situation is less obvious. Suppose you are having a conversation with a friend on the subject of music. She could utter, "Of course, I listen mostly to CDs, but I still love the sound of LPs" and accompany this sentence with two pointing gestures that are seemingly directed towards empty space. In particular, she could raise her left hand, extend the index finger and point to the left on the word "CDs" and she could do a mirroring movement with her right hand on the word "LPs".[1] All that would happen with no CDs or LPs in the vicinity. Is this kind of gesture – *abstract pointing* – helpful for the recipient?

This was the starting point for the research of the following years with the majority of it being covered in this dissertation. Specifically, it was explored whether abstract pointing can help to keep track of who or what the speaker is talking about via *spatial reference tracking*. During a later part of the conversation, for instance, a further pointing of your friend to the right could support you in realizing that LPs are the momentary subject and not CDs.

Before turning to the specific experiments, the next chapters provide an introduction on various levels. First of all, the method of event-related potentials is explained in more detail, as it is of importance for the remainder of this thesis. Then I give an overview of gestures in general and

[1] When describing gestures, I use the labels "left" and "right" from the observer's point of view throughout the whole thesis.

abstract pointing in particular in order to make the reader acquainted with the basic concepts and specific terms of this field. Furthermore, the research on gestures' communicative value is reviewed. Since there is only a small amount of according literature specifically about abstract pointing, I also present experiments that have been conducted on other gesture types. This will provide an idea of how the communicative value of abstract pointing can be explored.

Subsequently, I present three experiments conducted within the scope of this thesis and concerning the *comprehension* of abstract pointing. The results of the first study suggest that abstract pointing can indeed be utilized in order to infer who or what the speaker talks about. Paradoxically, they also seem to suggest that this usage is not helpful for the recipient. The next experiment qualifies this finding, as the beneficial effect of abstract pointing depends heavily on the reliability of the gestures. The third comprehension experiment covers questions such as whether successful spatial reference tracking is only possible with abstract pointing and whether spatial reference tracking is a general phenomenon. With the final experiment of the dissertation a first step towards the *production* of abstract pointing is undertaken. In particular, it is explored whether the recipient adopts the gesturing order provided by the speaker (e.g. *CD – left and LP – right*) and whether this can affect her or his motor responses.

Finally, a chapter with a general discussion of all experiments and with ideas for potential future research follows suit.

2 Event-related potentials – ERPs

There are various ways to explore whether and how gestures are processed by the recipient with one of them being the method of event-related potentials (ERPs). Over the course of the last decade this technique has become popular in gesture research and since it is used in two of the present experiments as well, I present it in the upcoming sections in a more detailed manner. Note, however, that I cover mainly aspects that are important for experimental design. For a more in-depth reading, also of technical issues, please refer for example to Kappenman and Luck (2012) or Coles and Rugg (1997).

2.1 The method

Neurons communicate electrochemically with each other, in particular ions flow into and out of them. These currents create in essence an electric dipole with positive being at one end and negative being at the other end. When putting two electrodes on the head of a participant, these dipoles can in theory be measured as an *electric potential* or *voltage*. It reflects the electric current's potential to flow from one electrode to the other. Since the communication between neurons is an ongoing process, the potential between two electrodes changes constantly. Every second, millions of neurons are sending and receiving currents. If you plot the variation of voltage over time, you receive an output, which is known as the *electroencephalogram*, the EEG. When the EEG is being time-locked to a specific event like the presentation of a stimulus during an experiment, the according output is called an *event-related potential* or ERP. It is considered to represent the brain's response to the specific stimulus.

An important characteristic of the EEG in general is that it measures only selected activity of a part of the brain. In particular, today's assumption is that the EEG is caused by large so-called *equivalent current dipoles*, which are generated by a sufficiently large number of synchronous excitatory post-synaptic potentials occurring at spatially aligned cortical pyramidal cells. This means, for instance, that the EEG does not pick up the

activity of single cells or inhibitory activity of the brain. Also, activity of subcortical regions like the thalamus is not detected by the EEG (at least not substantially). Thus, numerous processes in the brain cannot be measured with EEG, but this selectivity is not necessarily a disadvantage – for example, it makes the analysis of ERPs probably much easier.

Depending on the orientation of an equivalent current dipole within the brain, some sites on the scalp are better suited to measure its potential than others. As a consequence, the EEG is usually measured with many electrodes, so that electric activity of various orientations can be detected. In order to make comparisons between different studies easier, researchers rely on standardized electrode placements and nomenclatures. An example is the 10-10-system as provided by the American Clinical Neurophysiology Society (2006). A visualization of the system is depicted in Figure 2.1. The electrode sites are named using a combination of letters and numbers. The letters are derived from the underlying brain lobes or other anatomical landmarks (e.g. "P" refers to the parietal lobe). The numbers indicate the location in the lateral plane (odd numbers – left hemisphere, even numbers – right hemisphere, "z" refers to the midline).

As mentioned, voltage reflects the potential of the current to flow from one electrode to another. When obtaining ERPs it is common practice to choose a single reference electrode and to measure all individual electrodes against it. The reference is usually placed at a site, which is electrically rather neutral, so that is does not get affected by the experimental manipulation. Examples for such sites are the ear lobes, the mastoid bones behind the ears or the tip of the nose. Another strategy in order to obtain a reference is to calculate virtual references like the mean of both mastoid electrodes or the mean of all scalp electrodes.

A problematic issue about ERPs is their signal-to-noise ratio. The situation is roughly the same as when trying to have a conversation at a rock concert. At the concert, the music (noise) is too loud to understand the dialog partner (signal). With ERPs, the overall brain activity (noise) superimposes a single ERP (signal). A common way to solve this problem is to present a certain type of stimulus not only once during the experiment but several times. Then, the data from all epochs, where the stimulus type was presented, get averaged. There are two assumptions about this averaging process. First, the individual responses to the same type of stimulus are assumed to be identical or at least similar. Second, the background brain activity is assumed to be random. As a consequence, the true response should stay the same while the noise should get smaller

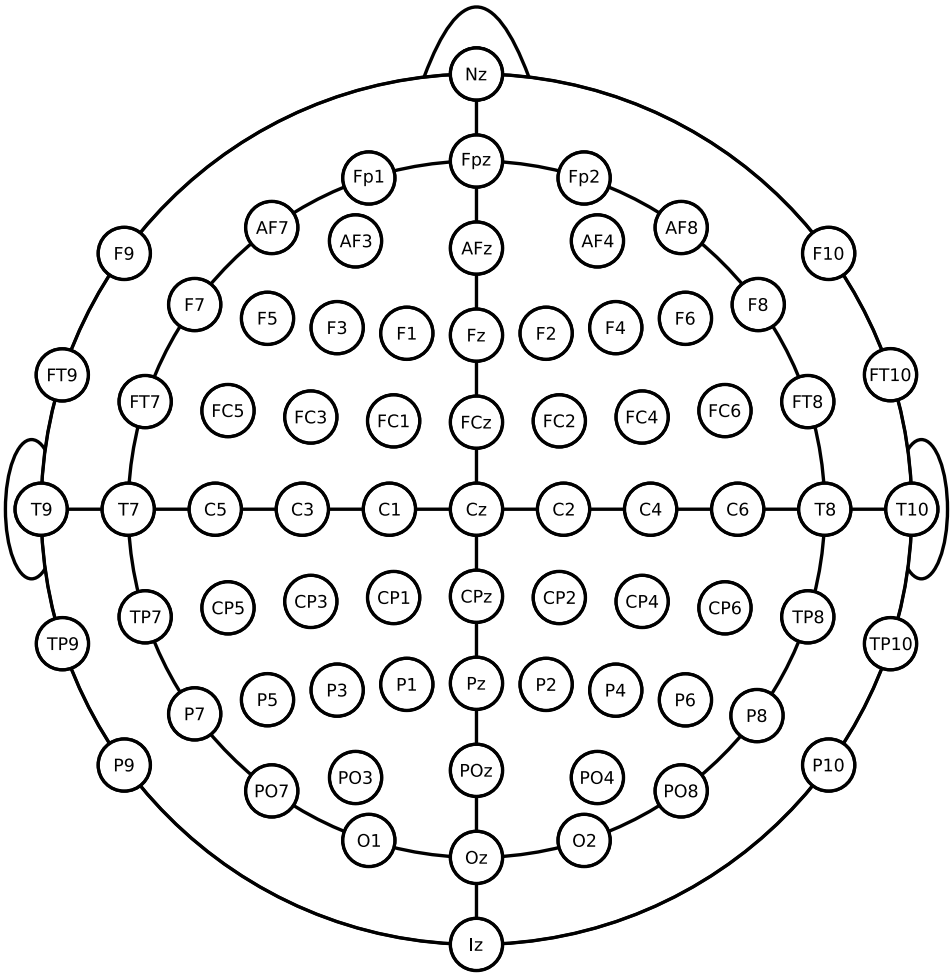

Figure 2.1: Electrode placement and labels according to the 10-10-system.

with every epoch that is added to the averaging process. Usually, the term *ERP* refers not to a single measurement, but to such an averaged ERP.

Another difficulty to deal with in ERP research are artifacts. The objective when using the ERP method is to identify the *brain's* response. Unfortunately, there are other sources than the brain that can influence the measurement – for example eye movements. An eye is from an electrical point of view a dipole. Thus, when the eyes move, voltage shifts can be measured. Two characteristics about these voltage shifts are problematic. For one, they are typically quite large, so that they are easily picked up by electrodes, which are meant to measure brain activity. For two, they can be of systematic nature (e.g. a participant could always blink at the same point in time during the experimental trials). In consequence, such unwanted voltage changes can co-occur with the event of interest and influence the data profoundly. Aside from the eyes, muscle activity and skin potentials can cause artifacts, too. There are two major ways to avoid such unwanted influence. One measure is to instruct the participants to sit still and to avoid unnecessary movements as well as eye blinks. Another measure is to apply afterwards automatic and manual procedures, with which contaminated epochs are identified and rejected from the averaging process. A basic logic behind these procedures is that they look for and reject time windows, where unusually big voltage changes take place that cannot be caused by brain activity. Especially in case of eye movements, the Electrooculogram (EOG) can be recorded in parallel to the EEG, which makes a later identification of contaminated epochs even easier.

A big advantage of ERP data lies in its *temporal resolution*. When an equivalent current dipole is generated in the brain, the according voltage distribution at the scalp can be detected in a virtually instantaneous manner and if the EEG is measured at a sample rate of 500 Hz, for example, the acquired precision lies at 2 ms. So, in approximation it is safe to say that ERPs taken from the scalp reflect what is happening in the brain at the same moment in time. Concerning the *spatial resolution* of ERPs, the following caveat is important to note: The ERP per se provides the spatial information about the voltage distribution on the scalp. It does not provide the information, however, where in the brain the according dipole was situated, i.e. what anatomical part of the brain generated the effect.

In the end, the acquired data is interpreted by trying to isolate ERP components, which were relevant during the stimulus presentation: The

ERPs of at least two experimental conditions are compared and analysed for systematic differences. When deviations are discovered, they are described by their polarity (is the crucial condition more positive or negative than the neutral condition), by their latency (when does the deviation occur), by their distribution across the scalp (where does the deviation occur) and by the experimental manipulation that led to it (c.f. Donchin, Ritter, & McCallum, 1978). Presumably for the sake of brevity, it is common practice to refer to found components only by mentioning their polarity and their latency. Typical labels are *N400* for a negative deflection occurring roughly 400 ms after stimulus onset or *P600* for a positive deflection occurring roughly 600 ms after stimulus onset. Both components will prove to be very important for the ERP experiments presented in this dissertation and, therefore, they are described in more detail.

2.2 The N400 – the meaning component

The N400 effect was first reported in two articles by Kutas and Hillyard (1980a, 1980b). In a visual word-by-word procedure the participants were presented – amongst others – with semantically appropriate sentences as in (1) and with sentences as in (2), where the meaning of the last word was inappropriate with regard to the rest of the sentence.[1]

(1) He shaved off his mustache and *beard*

(2) He shaved off his mustache and *city*

When comparing the ERPs time-locked to the onset of the last word, there was a negative deviation for the semantically inappropriate condition compared to the inconspicuous condition around 400 ms. Crucially, this effect was only demonstrated for semantically unexpected stimuli, but not for physically unexpected stimuli that were written in capital letters. Therefore, the still valid assumption arose that the N400 is connected to the processing of meaning.

Since these seminal studies more than 30 years have passed and more than 1000 published articles have helped to expand the knowledge about the N400 effect and the conditions, under which it is elicited (for a review see Kutas & Federmeier, 2011). For example, it is not only observable in

[1] Within the context of ERP experiments, italicized words as in (1) and (2) will mark the position where the ERPs were taken throughout the whole thesis.

the above applied sentence violation paradigm, where the critical word is inappropriate with regard to the sentence context, but also in priming paradigms, where the context consists usually of a single word: In a typical trial, the presentation of a *prime* word is followed by the presentation of a *target* word and when both words are semantically unrelated as in *table – animal*, the N400 for the target word is more negative than when both words are semantically related as in *doctor – nurse* (Bentin, McCarthy, & Wood, 1985; Holcomb, 1988). At the same time, the N400 effect is also observable at a discourse level, where the semantic conflict exists between a critical word and the preceding story line (van Berkum, Hagoort, & Brown, 1999).

On the way to ERP studies on gestures, two other expansions of the classic effect are notable. First, the N400 is not restricted to purely linguistic material, but also triggered by other meaningful items like pictures (McPherson & Holcomb, 1999; West & Holcomb, 2002). Moreover, it can even be elicited when switching between stimulus types. For instance, Ganis, Kutas, and Sereno (1996) presented sentences on the screen, but the last word was substituted for a line drawing. Instead of the word "pipe", for example, the participants saw a drawing of a pipe. Regardless of this switching between stimulus types an N400 effect could be observed. A second important expansion is that the N400 is not only shown with highly standardized material, where one item at a time is presented on the screen, but also with more dynamic and natural stimuli. In the linguistic domain this is demonstrated by experiments, where spoken utterances are presented to the participants instead of printed ones (McCallum, Farmer, & Pocock, 1984; Holcomb & Neville, 1990). In the non-linguistic domain, this was demonstrated by Sitnikova, Kuperberg, and Holcomb (2003), for example, who presented short videos, during which objects were either appropriately applied or not. For instance, a man, who had prepared for shaving, stroked either a razor across his face or a rolling pin and this manipulation led again to an N400 effect.

Despite the vast amount of research on the N400, the precise function behind this component is still a matter of debate (Federmeier & Kutas, 1999; Kutas & Federmeier, 2000; Lau, Phillips, & Poeppel, 2008; Friederici, 2011; Kutas & Federmeier, 2011). One of the current views suggests that the N400 is correlated to the process called *lexical access* (Brouwer, Fitz, & Hoeks, 2012; van Berkum, 2009): Whenever you encounter an item like a word or a picture of an object, you have to retrieve information about the meaning of this item from long-term memory. In particular, it

is assumed that features that are connected to the item get activated. For example, when encountering the word "Christmas" this could lead to the activation of features like tree, Santa Claus, presents, Jesus Christ, winter and so on. The lexical access view of the N400 suggests that the easier this information retrieval or feature activation is, the less negative the N400. In this line, the attenuated N400 for the word "beard" in (1) could be explained by the fact that the sentence context had already preactivated relevant features – like a beard, a mustache is made of hair, can be found in a men's face and can be cut with a razor. The feature overlap between the sentence context and the word "city" in (2), however, is rather low and, therefore, lexical access is more effortful and an increased N400 is found.

Another very prominent view suggests that the N400 reflects how easy it is to integrate the meaning of a new item into an existing context. According to this hypothesis, the word "beard" elicits a less negative N400, because its meaning can be easily combined with the meaning of the preceding sentence part (as opposed to the meaning of "city"). As already alluded to, there is so far no agreement on which functional interpretation of the N400 is to prefer. A simple reason is that in many experiments both views are suited to explain the results as shown for the above beard-city example. In order to avoid confusion when reading this dissertation, the N400 will be uniformely regarded as a marker for lexical access.

2.3 The P600 – the representation component

The P600 – which is usually most prominent over centro-parietal sites – was first discovered in experiments that used syntactic anomalies. For instance, Hagoort, Brown, and Groothusen (1993) compared the processing of syntactically correct sentences as in (1) with the processing of syntactically incorrect sentences as in (2), where a number violation between the subject and the verb exists. When analysing the ERPs time-locked to the onset of the verb, the authors found a P600 effect, i.e. the syntactically incorrect condition was more positive than the correct condition around 600 ms.

(1) The spoilt child *throws* the toy

(2) The spoilt child *throw* the toy

Another type of syntactic anomaly resulting in an increased P600 are so called *garden-path sentences*. The initial part of such sentences misleads the (majority of) participants to assume a syntactic structure, which turns out to be wrong in the end. For example, when a sentence starts with "The tailor hired ...", most people would assume that the verb is applied in active voice as in "The tailor hired somebody". In a typical garden-path sentence like "The tailor hired to repair the clothes", however, it becomes clear at the word "to" that the active voice interpretation cannot be maintained and an increased P600 is the result (Osterhout & Holcomb, 1992). Findings such as these suggested two things: First, the P600 is observed when the critical item cannot be integrated into the assumed syntactic structure. Second, the P600 effect is correlated to a reanalysis process of the stimulus material, for instance in order to revise the initially assumed structure (see also Friederici, Hahne, & Mecklinger, 1996).

During the last couple of years, an increasing number of experiments has challenged the syntax specificity of the P600, as they showed a more positive P600 for sentences, which are syntactically fine, but from a semantic point of view surprising (e.g. Kuperberg, Sitnikova, Caplan, & Holcomb, 2003; Hoeks, Stowe, & Doedens, 2004; Kim & Osterhout, 2005; Sanford, Leuthold, Bohan, & Sanford, 2011; Regel, Gunter, & Friederici, 2011). For instance, van Herten, Kolk, and Chwilla (2005) found a larger P600 for sentences like (4) compared to sentences like (3) at the position of the second noun.

(3) The poacher that hunted the *fox* [...].

(4) The fox that hunted the *poacher* [...].

Various frameworks have evolved trying to explain such semantic P600 effects (e.g. Kuperberg, 2007; Bornkessel-Schlesewsky & Schlesewsky, 2008; van de Meerendonk, Kolk, Chwilla, & Vissers, 2009; Friederici, 2011). They represent insofar an own class of P600 theories, as a main assumption of all of them is that language is analysed by more than one process and whenever a conflict arises regarding the interpretation of the input a more positive P600 is the result. According to the *monitoring hypothesis* by the group around Kolk, Chwilla and Vissers, for example, language perception is monitored for errors and a P600 effect is supposed to be observed, when the perceived information gets in conflict with a strong expectation about what should be perceived. In (4), presumably, a purely linguistic process leads to the conclusion that it is indeed the fox that

hunts the poacher. A second, heuristic process, however, leads to the conclusion that it should be the poacher that hunts the fox, simply because this is most likely given the words "fox", "hunted" and "poacher". The conflict between both processes indicates a potential error in language perception and in order to solve this issue a reanalysis is started, which is reflected by the more positive P600 (see also Kolk, Chwilla, van Herten, & Oor, 2003; van Herten, Chwilla, & Kolk, 2006; van de Meerendonk, Kolk, Vissers, & Chwilla, 2010).

As if a syntactic and a semantic sensitivity would not be enough, the P600 effect has also been demonstrated in cases of completely well-formed sentences (Kaan, Harris, Gibson, & Holcomb, 2000; Fiebach, Schlesewsky, & Friederici, 2002; Phillips, Kazanina, & Abada, 2005; Gouvea, Phillips, Kazanina, & Poeppel, 2010). Most relevant for this dissertation, Burkhardt (2006) could show a modulation of the P600 depending on the processing of discourse referents. She presented the participants with three conditions: In the given referent condition (5), an already known referent is repeated within a short amount of time. In the bridged referent condition (6), the referent relates to information that was provided earlier in the sentence. In the new referent condition (7), a specific referent is introduced for the first time.

(5) Tobias visited a conductor in Berlin. He said that *the conductor* was very impressive.

(6) Tobias visited a concert in Berlin. He said that *the conductor* was very impressive.

(7) Tobias talked to Nina. He said that *the conductor* was very impressive.

For conditions (6) and (7) the data revealed a more positive P600 compared to (5). Additionally, a modulation of the N400 was found with (7) showing the most pronounced N400, (6) a middle one and (5) the least pronounced one. The N400 effect is rather easy to explain as an increased difficulty in lexical access: In the given referent condition (5), all relevant features about "the conductor" are already preactivated due to the initial sentence. In the bridged referent condition (6), it can be assumed that the word "concert" primed at least some of the relevant features. For the new referent condition (7), however, it is feasible to assume that no features were preactivated when the critical item arrived resulting in the most negative N400.

The P600 effects are more difficult to explain. In the bridged referent condition (6), for instance, there is no syntactic anomaly apparent and it is also hard to imagine a conflict between two processing streams. Brouwer et al. (2012) presented a framework, which specifically tries to account for such findings. A core element of their theory is the idea that the recipient builds a *M*ental *R*epresentation of what is being *C*ommunicated to him – hence the short label *MRC hypothesis* (cf. Zwaan & Radvansky, 1998). The authors suggest that the mental representation is constantly refined based on the incoming input. The amplitude of the P600 is supposed to reflect the difficulty of this process. If a new word requires a more considerable modification of the MRC, a more positive P600 will be the result. In this line, the increased P600 in (6) is supposed to be due to the fact that a specific referent has to be added to the MRC. This is not necessary in (5), as the referent is already part of the representation. An analogous explanation is applicable for the P600 effect between (7) and (5).

On the way from the syntactic view to multi-stream models such as the monitoring hypothesis to the MRC hypothesis one idea of the P600 has remained unchanged – that its amplitude reflects in a general sense processing costs: the more processing necessary, the more positive the P600. What has changed are the assumptions about what kind of incidences can cause the increased processing. The syntactic reanalysis approach as well as the multi-stream models relate the P600 effect to some extraordinary event like an outright syntactic violation. In contrast to that, the MRC hypothesis assumes that the P600 is the brain's ordinary response to an item that provides new information. This item may not render the sentence agrammatical or semantically surprising – as long as it requires to adapt the mental representation more considerably than the control condition, a more positive P600 will be the result. Under this asssumption, the P600 in (2) is not elicited, because an error per se is present, but because the construction of a coherent MRC is more difficult, when subject and verb do not agree on how many actors there are. In (4), the more positive P600 is not elicited because of a conflict between a linguistic and a heuristic analysis, but because it requires more processing to create the mental representation of a poacher hunting fox in a world of fox hunting poachers.

Today, it is still a matter of debate, which of the functional views of the P600 is most suitable. Within this dissertation, however, data will be provided, which favor the MRC hypothesis.

Irrespective of the assumptions about what processes it reflects, another interesting characteristic of the P600 is that it is not language specific. For example, it is also elicited by harmonic anomalies in music (Patel, Gibson, Ratner, Besson, & Holcomb, 1998), by violations in abstract sequences (Lelekov, Dominey, & Garcia-Larrea, 2000) or by violations in numeric sequences (Nùñez-Peña & Honrubia-Serrano, 2004). The P600 amplitude can even indicate an interaction of linguistic and extra-linguistic information. For instance, Lattner and Friederici (2003) could show that a statement like "I like to wear lipstick" yields a more positive P600 when uttered by a male voice than when uttered by a female voice, i.e. the gender of a speaker as inferred by the sound of the voice can modulate how an utterance is processed. Within the framework of the MRC hypothesis this result could be explained in the following way: When the attributes of a speaker do not fit with what the speaker says, it is more difficult to create an according mental representation (for a comparable P600 effect due to incompatibility of speech content and speaker's gender see van Berkum, van den Brink, Tesink, Kos, & Hagoort, 2008).

2.4 ERPs in gesture research – example of use

At least because of three characteristics ERPs are well suited for the research on gestures: First, ERPs are a direct measure of the underlying brain activity and can indicate whether gestures are processed or not irrespective of any obvious behavior. Second, due to their excellent temporal resolution they are invaluable to study gesture speech integration, which occurs very quickly. Third, ERPs can hint at the nature of the underlying neurocognitive processes by comparing the results with established ERP components such as the N400 or P600.

An example of how ERPs can be applied in gesture research is given by a study on *emblems* (Gunter & Bach, 2004). These gestures are insofar special, as they have a predefined meaning. The thumbs-up gesture is a typical emblem, for instance, and most people will without further explanation identify it as a sign of approval. Gunter and Bach (2004) wanted to know whether emblems trigger a similar kind of semantic processing as words. The participants in the experiment watched 396 photographs of hand postures and had to judge their meaningfulness. Half of the hand postures were meaningful emblems and half of them were random shaped. The results showed a difference in the ERP waves between the

conditions. On a basic level this indicates that emblematic gestures are not processed as ordinary hand shapes. On a more specific level, an N400 effect was demonstrated with a less negative N400 for meaningful emblems. This result suggests that gestures with a predefined meaning are processed by a similar semantic system as words (for an experiment on sign language with a similar finding, see Neville et al., 1997).

Aside from emblems, a lot of different gesture types exist. Some of them do not only have a meaning, but they can also reflect syntactic information. Then again, there are gesture types whose purpose is not clear-cut. The next chapter provides an according overview and also an introduction to the research on gestures' communicative value.

3 Theoretical and empirical background

3.1 Gestures in general

When a friend asks me for the topic of my dissertation, I do not say that it is about abstract pointing. Nobody has a clue what abstract pointing is. Instead I say, "It's about gestures." This usually opens a door: "I see, I have once read this book about body language." "Are you talking about sign language?" "The Italians, they gesture a lot, don't they?"

So, let's start with entering the world of gestures in general.

3.1.1 An attempt to define gestures

Adam Kendon proposed that gestures are all those body movements that show "manifest deliberate expressiveness" (Kendon, 2004, p. 15). The term *expressiveness* refers to the fact that the relevant body movements are considered to have a communicative intention. When extending the thumb in the air, while flexing all other fingers into a fist-like shape, this is regarded as a gesture. A major reason is that this thumbs-up gesture is expressive, it conveys information to the recipient, in this case that something was well done. The feature *manifest* means that the relevant movements are readily perceived as expressive. For instance, the thumbs-up is recognized as having a communicative intention and easily discriminated against other non-expressive body movements like scratching one's nose. The feature *deliberate* means that most people consider gestures as being conducted on a voluntary basis. Accordingly, the thumbs-up movement is assumed to be under the control of the speaker and this puts it into contrast to other movements like yawning, which is considered inadvertent.

All three gesture characteristics provided by Kendon can be objected. Some might doubt that all gestures are communicative. Some might find it recursive that a gesture is supposed to be that body movement, which

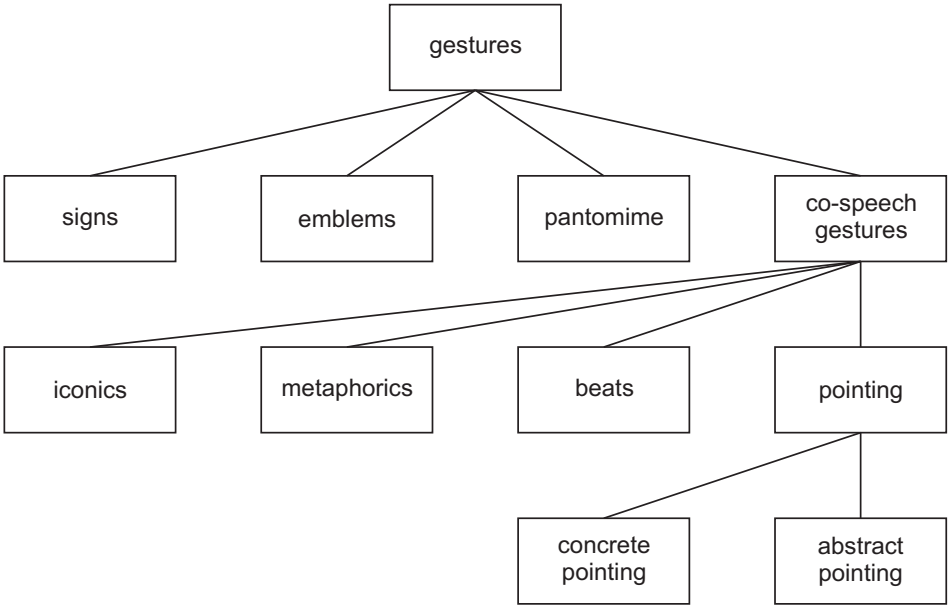

Figure 3.1: A categorization scheme of gestures.

is recognized as a gesture. And when observing some people talking it is tempting to judge that they have absolutely no control over their gesturing. Nevertheless, Kendon's definition gives a good first impression and, admittedly, it is a difficult task to formulate a gesture definition accepted by everybody. A reason for this is that there are a lot of different gesture types and sometimes they show characteristics, which are diametrically opposed. Figure 3.1 provides an overview of some of these gesture types and is based on a categorization scheme by McNeill (1992, 2005). Already with a few basic questions it is possible to show the distinctiveness of the gesture types on the second level (signs, emblems, etc.). Such questions include: Is the gesture type accompanied by speech or not? Does the gesture type have linguistic properties – for example, is there something like a lexicon (cf. McNeill, 1992, 2005; Kendon, 2004)?

3.1.2 Gesture categories

Signs are the building blocks of sign languages such as German or American Sign Language. Their goal is to work without any accompanying speech. They represent own language systems and show typical linguis-

Figure 3.2: Left: Churchill using an emblem, right: an iconic.

tic properties: There is a standard of well-formedness, i.e. a sign has to be shaped according to a convention in order to be interpretable. There is a lexicon, which assigns a specific meaning to a specific sign. There are also syntactic rules, which allow to combine single signs into phrases and sentences. Due to these characteristics it is possible to express everything you could orally talk about with signs as well.

Emblems such as the thumbs-up or the victory sign as used by Winston Churchill (Figure 3.2, first picture) could be called the signs for non-signers. Like signs they can be understood without additional speech and they also show some language-like properties. For instance, there are conventions about their shape. The thumbs-up has to be conducted with an extended thumb and not an extended little finger, otherwise the recipient will not understand the message. Emblems do also have predefined meanings, i.e. there is a lexicon. Unlike signs, however, the vocabulary is rather small, as the amount of emblems is very limited. This aspect alone makes it virtually impossible to tell a coherent story with emblems. Consequently, there is also no need for specified syntactic rules.

Pantomime functions likewise without speech. A difference to signs and emblems is that the absence of speech is even mandatory. Additionally, pantomime lacks linguistic properties and rules. Mimes are completely free in how they express a certain concept as long as they get the intended message across. Professional mimes have in this regard a high competence and can tell full stories without a single word.

Co-speech gestures, finally, are probably the type of gesture that is encountered most often in everyday life. They are those hand and arm

movements, which are spontaneously produced while talking. For instance, when a speaker says, "... and then I ran over the street", and simultaneously moves her right hand – palm down – from the right to the left while imitating a walking movement by wiggling her index and middle finger, a typical co-speech gesture is carried out (Figure 3.2, second picture). Co-speech gestures are insofar different from the other gesture types, as they need to be accompanied by speech, otherwise it is very hard to infer their exact meaning. For example, it would not be clear whether she is referring to herself with this gesture or a third person. This lack of clear-cut interpretability is due to the fact that co-speech gestures do not have linguistic properties like a lexicon or a standard of well-formedness. When talking about running it is up to the speaker whether she uses wiggling fingers, imitates a runner with her whole body or makes a swift move with an extended index finger from left to right.

Co-speech gestures themselves can be further divided into various subcategories (see Figure 3.1). *Iconics* are described as pictorial gestures that try to convey an image to the recipient. Importantly they illustrate a concrete object or action, so something that could be observed in the real world. The above given wiggling finger example accompanying "... and then I ran over the street" is a typical iconic. The gesture vividly depicts the action of running.

Metaphorics are also very pictorial gestures. In contrast to iconics, however, the picture is not used to describe a concrete object or action, but an abstract idea. A typical metaphoric would be to bring one hand in a cup shape on chest height and move it from yourself to the interlocutor while uttering, "I have a question for you." A question is nothing that can be carried from one position to another position. Nevertheless the gesture invokes this idea.

Beats received their name, because they look as if somebody was beating music. Usually, they are simple flicks of the hand going up and down or back and forth. Presumably, beats are used when the speaker wants to stress a certain part of the speech stream or introduce something newsworthy. For instance, a beat may accompany the first mention of a specific character or a politician could conduct a beat on the last word in the sentence, "I want to make that clear", in order to emphasize his intention.

Finally, there is a gesture type called deictic gesture or *pointing*. It represents a quite robust gesture type, because aside from McNeill (1992, 2005) it has been mentioned by various other authors, who developed a way to categorize gestures: Quintilian (1995) in the first century AD,

Johann Jacob Engel (1968) in the 18th century, Gilbert Austin (1966) at the beginning of the 19th century and Wilhem Wundt (1911), David Efron (1941) and Ekman and Friesen (1969) over the course of the 20th century (for a comparison see Kendon, 2004).

3.2 Pointing in particular

When talking about pointing most people probably visualize somebody raising the arm and extending the index finger while the other fingers are flexed. There are, however, a lot of variations to this typical movement. Enfield, Kita, and de Ruiter (2007), for example, make a difference between B(ig)-points and S(mall)-points. B-points are described as "gestures in which the whole arm is used as articulator, outstretched, with elbow fully raised" and S-points as "gestures in which the hand is the main articulator, the arm is not fully straightened, typically with faster and more casual articulation" (Enfield et al., 2007, p. 1725). The shape of the hand can be modified, too, and this might even be correlated to situational factors (Kendon & Versante, 2003; Wilkins, 2003). For instance, in many cultures it is considered rude to point at a person with the index finger. What can be done instead is to use an open hand palm up. A further variation of the typical pointing gesture can be achieved by using a body part other than the hand. Suppose you want to make the interlocutor aware of something in a discrete way. Probably you would use a simple head nod or an eye pointing instead of the classic gesture. Furthermore, it is imaginable to point with an elbow or the whole torso and there is even evidence for lip-pointing (Enfield, 2001; Wilkins, 2003).

Regardless of how exactly it is executed, there are three components to each pointing gesture: a point of origin, a target and a trajectory linking the origin and the target. Again, like gestures and co-speech gestures, pointing can be subcategorized, as not all pointings are one of a kind (Figure 3.1). One way to do so is by looking at the target (McNeill, Cassell, & Levy, 1993; McNeill, 2003): In *concrete pointing* the target is present whereas it is absent in *abstract pointing*.

3.2.1 Concrete pointing

If concrete and abstract pointing were twins, concrete pointing would be the popular sibling. It is the gesture type everybody thinks of when hearing "pointing". Examples are an infant pointing to a toy, a witness

pointing out the culprit during a trial or Uncle Sam on the famous poster by James M. Flagg, where he is looking for new recruits ("I want YOU for U.S. Army"). Importantly, the target in all these gestures is physically present at the end of the trajectory. There is a toy, a culprit and a viewer of the poster. Concrete pointing is insofar a typical co-speech gesture, as it depends heavily on accompanying speech. If I point to a telephone without saying anything, you would be clueless about my intention. Maybe it rings very quietly and I want you to pick it up? Or maybe I am baffled by its very strange color and want to share my surprise with you? Without speech this is not clear (cf. Tomasello, Carpenter, & Liszkowski, 2007). However, there are situations, where concrete pointing can even work without speech. Imagine that we both become witness of how a third person gets a heart attack, I look at you and I point to the phone – it would be probably clear to you that you are supposed to call an ambulance.

Concrete pointing is one of the first gestures to be seen in infants. On average, they start to conduct this gesture type shortly before their first birthday and, thus, in a prelinguistic phase (Leung & Rheingold, 1981; Butterworth & Morisette, 1996; Carpenter, Nagell, & Tomasello, 1998; Melinder, Konijnenberg, Hermansen, Daum, & Gredebäck, 2015). This aspect has gained considerable attention, because in their pointing infants depict already behavior, which is thought to be fundamental for language usage. For instance, before they do so through speech, infants create joint attention with the recipient via concrete pointing (Carpenter et al., 1998). Furthermore, infants try to infer the knowledge of the recipient and use this information in order to adapt their pointing behavior. For example, in light of a searching adult, infants point more often to an object about whose location the adult is unaware than to an object about whose location the adult is aware (Behne, Carpenter, & Tomasello, 2005; Liszkowski, Carpenter, & Tomasello, 2008). Given that typical language competences are shown in pointing, it fits well that there is a significant longitudinal relation between both domains: the better the comprehension and production of concrete pointing, the better the later language proficiency (Colonnesi, Stams, Koster, & Noom, 2010). In sum, the development of concrete pointing is regarded as a key transition phase on the way to language (Butterworth, 2003; Iverson & Goldin-Meadow, 2005; Tomasello et al., 2007).

Aside from developmental psychology, concrete pointing is also discussed within the field of comparative psychology, as it is produced by non-human animals such as apes and monkeys (Hare & Tomasello, 2005;

Miklósi & Soproni, 2006; Tomasello & Call, 2007; Mulcahy & Hedge, 2012). Admittedly, there are a lot of restrictions to this statement such as the fact that pointing is almost exclusively conducted by apes that live in captivity and interact with humans, but not by apes that live in their natural habitat (Leavens, Hopkins, & Bard, 2005). Also, apes point apparently only *imperatively*, i.e. when they want humans to do something for them (Tomasello & Call, 2007; Bullinger, Zimmermann, Kaminski, & Tomasello, 2011). In contrast to that, already human infants can point *declaratively* in order to share an emotion and *informatively* in order to help the interlocutor (Bates, Camaioni, & Volterra, 1975; Tomasello et al., 2007; Liszkowski et al., 2008). Despite the mentioned restrictions the bottom line is that concrete pointing can be found in non-human animals. Given its presence in ontogeny and phylogeny it comes as no surprise that concrete pointing is speculated to play an important role in the evolution of language (Tomasello, 2005; Tomasello & Call, 2007).

Last but not least, concrete pointing is a gesture type, which has without a doubt a communicative value. Imagine that you are in desperate need for a restroom. You enter a restaurant, ask for its whereabout and the waiter answers, "You can find it over there." If this response is not accompanied by a concrete pointing, it is of no use for you. The communicative value of concrete pointing is that it provides you with the information, where the object of interest can be found.

3.2.2 Abstract pointing

Abstract pointing is different. It starts with the fact that it is not directed at a physically present target, but instead at an empty location. When your friend utters, "Of course, I listen mostly to CDs, but I still love the sound of LPs" and accompanies the word "CDs" with a pointing to the left and the word "LPs" with a pointing to the right, this is abstract pointing, because there are no CDs or LPs at the end of the trajectories (McNeill et al., 1993; McNeill, 2003). This still leaves room for variants of abstract pointing, some of them require to comply with geometrical constraints. For instance, when I describe my new apartment to you and point out the location of the living room, the bed room and so on, these entities might not be present, but I have to stick to the actual floor plan. Comparable to this kind of *topographical pointing* is a phenomenon that could be called *pointing at absent referents* (cf. Liszkowski, Carpenter, & Tomasello, 2007; Liszkowski, Schäfer, Carpenter, & Tomasello, 2009). When your col-

league comes back into the office after a vacation and points to an empty spot on the desk while asking, "Where is my plant?", he can also not point after his fancy, but has to point at the plant's former location in order to get the message across. Then, there are gestures, which do not only set up referents, but whole scenes. This might be done in order to represent a real-life event and, crucially, these scenes often include further cues regarding the associations between referent and gesture location. Examples are that referents are permanently presented and not only for a short moment or that additional iconic gestures depict interaction, which makes it clear which referent belongs to which location (Cassell, McNeill, & McCullough, 1999; Sekine & Kita, 2015). In sum, it would be worth a discussion on its own what kinds of gesturing are covered by the term *abstract pointing*. For the present dissertation, however, it shall suffice to say that the pure, geometrically unrestrained type was explored. Your music loving friend can choose whether she points *CDs – left and LPs – right* or vice versa. Also, no further cues hint at where a specific referent is positioned.

In contrast to concrete pointing, abstract pointing is one of the last gestures to emerge in child development. Its usage does not start before the age of eight years (McNeill, 1992, p. 322) and it takes until the age of twelve, before it reaches a considerable level (McNeill, 2005, p. 40). Looking for further research, one gets the impression that abstract pointing is not the popular but the inconspicuous twin. For instance, abstract pointing has not been discussed yet within the fields of language development, animal behavior or language evolution. A reason for this negligence could be that it is by no means clear whether abstract pointing is useful. "Of course, I listen mostly to CDs, but I still love the sound of LPs" – it appears that this statement is perfectly understandable without the pointing gestures. Hence, does abstract pointing have a value for communication and, if so, what is its nature? This question represents the starting point for the research on this peculiar gesture type.

One possiblity is that abstract pointing could fulfill a contrasting function. Maybe it serves as a visual cue for the recipient that the speaker does not talk about a homogeneous mass, but about two different things and that this distinction should be noticed. To translate this idea into pure words: It is a difference whether your friend talks about "CDs and LPs", where both music media are treated as one of a kind, or about "CDs on the one hand and LPs on the other hand", where it is clear that the media are put in contrast to each other. Abstract pointing could have

the function to visually highlight a contrast and replace such figures of speech. A second possibility goes even one step further: Perhaps abstract pointing cannot only inform the recipient that two things are contrasted, but also about the identity of the discourse referents.

Reference tracking is an essential task in each conversation: The recipient has to identify, who or what the speaker is momentarily talking about and he has to update this information continuously. Does the speaker still talk about CDs? Did she switch to LPs? Or is she talking now about tapes from the 80s? Abstract pointing could support this task. In particular, it could be that the recipient associates distinct locations in gesture space with the referents that are mentioned in speech. I.e. the left location in gesture space would get associated with the referent CDs and the right location with the referent LPs. Now imagine that your friend continues, "Whenever I have enough time to consciously enjoy music, I listen to LPs", and accompanies the last word with a further abstract pointing to the right. Based on the established gesturing order, this gesture could be – aside from speech – an indication that the momentary referent is the LP.

The idea of such *spatial reference tracking* is not a new one. After a short excursus, existing knowledge about this topic will be presented.

3.2.3 Excursus: Is abstract pointing pointing?

When considering that abstract pointing serves as referent indicator, a side question about gesture categorization evolves: Does abstract pointing truly belong to the gesture category *pointing*? Let's reconsider the statement, "Whenever I have enough time to consciously enjoy music, then I listen to LPs.", where an abstract pointing to the right accompanies the last word. This gesture has certainly a pointing component, because it indicates a specific location in gesture space. However, it has also a metaphoric component, because with this gesture the speaker refers not to this empty location but to something completely different, i.e. LPs. When looking at more formal aspects, the situation becomes not easier: Sometimes the shape of abstract pointing resembles not the classic pointing gesture, but a platform or container carrying something, which is typical for metaphorics as in the cup shaped hand carrying a question to the interlocutor.

In fact, this is a general problem of gesture categorization and not solely of abstract pointing. Very often it is difficult to assign a gesture unambiguously to a specific category. Imagine that I ask you to hand me something

over and simultaneously I stretch out my arm towards the desired object while rhythmically flexing and extending my fingers. This gesture has a concrete pointing component, because I indicate the location of the object. However, there is also an iconic component, because I allude to the action of moving something towards me through the movement of my fingers. These difficulties in categorization are known and it has been acknowledged that it would be more appropriate to speak of *dimensions* rather than categories (c.f. Kendon, 2004; McNeill, 2005). Then it would be possible to say that the gesture on "LPs", for instance, scores to a certain degree on the pointing dimension and to another degree on the metaphoricity dimension. These scores could even be adjusted for each single gesture: When carried out with the hand in container form an abstract pointing could get a higher metaphoricity score than when carried out in classic pointing form.

Due to reasons of brevity and simplicity, however, I will continue to use the categorizational approach in this thesis. Mostly because of personal preference I refer to the studied gesture type as abstract pointing and not, for example, as metaphorics.

3.3 Research about spatial reference tracking

3.3.1 In sign language

Spatial reference tracking in itself is not a newly discovered phenomenon. For instance, it is well known that it is a widely applied communicational strategy in sign language. Signers routinely place referents in the space in front of them and the interlocutors interpret further signs to these locations as referent indications (e.g. Poizner, Klima, & Bellugi, 1987; Emmorey, Corina, & Bellugi, 1995; Neidle, Kegl, MacLaughlin, Bahan, & Lee, 2000). Sometimes this referential usage of space is carried out in a pure manner, where the locations for the referents can be chosen arbitrarily. Sometimes it is carried out in a topographical manner, where the referents are set up according to a scheme that meets geographical relations in the real world. This distinction mirrors pretty much the earlier mentioned separation of pure abstract pointing from other types of pointing (see Section 3.2.2). The usage of referential space is even a codified part of sign language, i.e. signers have to comply with rules when applying it. For instance, in American Sign Language it is possible to conduct the explicit sign for the referent with one of the hands while simultaneously

assigning it a location in gesture space with the other hand. In contrast to that, German signers have to execute these actions in a sequential manner (Keller, 1998, pp. 120–121).

It is difficult to estimate what the existence of spatial reference tracking in sign languages means for spoken languages. On the one hand, it could be assumed that this is a specific of sign language and not applied in spoken language. The most intuitive argument for this hypothesis is probably that sign language is inherently spatial while spoken language is not. In sign language, every bit of provided information is located somewhere in space. In contrast to that, speech, which represents the dominant part of spoken language, cannot be placed or positioned at a confined location. On the other hand, speech is not the only part of face-to-face communication. There is also the gesture channel, through which spatial reference tracking could in theory be accomplished and if this was the case, this would add to a considerable number of characteristics that are shared by both language systems. For instance, sign and spoken language have many syntactic characteristics in common (Klima & Bellugi, 1979; Liddell, 1980) and they share developmental characteristics in that children acquire them over the course of various developmental stages (Bellugi, van Hoek, Lillo-Martin, & O'Grady, 1988). Furthermore, they share to a very large degree the underlying neural systems (MacSweeney, Capek, Campbell, & Woll, 2008). In other words, there seem to be universal principles, which are valid in both language systems and it could be that spatial reference tracking is one of them.

A few studies have already investigated whether spatial reference tracking is applied by non-signers. They focussed on the speaker and, thus, on the production of gestures. Importantly, they were not constrained to abstract pointing, but explored co-speech gestures in general.

3.3.2 In co-speech gestures

In order to make spatial reference tracking possible at all, it is essential that gesture space is used consistently. If the speaker gestures at one time to the left and at another time to the right when talking about LPs, it will be difficult to infer the referent from a third gesture based on its location. Some studies already investigated this issue and their experimental design was roughly the same: The participants were put in a situation, where they had to describe a video scene or narrate a short story line. Afterwards, the gestures produced by the participants were checked for spa-

tial consistency. So, Kita, and Goldin-Meadow (2009), for example, found that 35% of all gestures were produced at locations that had already been associated with a specific referent. Gullberg (2006) studied a group of L2-learners and explored *second occurrence gestures*, i.e. gestures that accompany a speech referent, which has already been accompanied by a gesture at least once. The results showed that 42% of these gestures were conducted at the same location as the preceding gesture(s). Alamillo, Colletta, and Kunene (2010), finally, report developmental changes: Whereas only 20% of six year old children showed consistent usage of gesture space, 68% of the adults did. In sum, it seems that speakers are capable of using gesture space consistently, but that this behavior is not shown permanently and not by all speakers.

So, Coppola, Licciardello, and Goldin-Meadow (2005) investigated this matter further by applying an extended experimental design. Again, the participants had to describe short video scenes and, again, they could use speech and gestures in one of the conditions. In another condition, the only communicational means they were allowed to use was the gesture channel. In the speech-and-gesture condition roughly 50% of the second occurrence gestures were conducted consistently, which fits roughly with the results of the above mentioned studies. In the gesture-only condition, however, the rate increased to 90%. Thus, if forced to do so, non-signers can rapidly and without further instructions behave like true signers regarding the spatially consistent production of gestures.

The just described research on the production side of spatial reference tracking is invaluable, because it gives a good impression of how gestures are used in a rather naturalistic setting (aside from the last study, the participants were not instructed to conduct gestures). The downside is that the results are in some regards difficult to interpret. For example, a valid hypothesis would be that the rate of spatially consistent second occurrence gestures should be significantly above chance. However, what is the chance level? If gesture space is only divided in a left and a right section, then it would be at 50%, but it could also be that there is a lower left, an upper left and so on. In consequence, the problem of testing against chance would be that the amount of distinct locations in space is basically up to the speaker and difficult to estimate for the experimenter. To make it worse, interindividual and intraindividual differences are possible. Interindividual differences, because a physically circumscribed part of gesture space might contain a single referential location for one speaker, but two for another speaker. Intraindividual differences, because by di-

viding a spatial location, which contained only one referent during the preceding discourse, a speaker can easily change the amount of distinct locations.

One way to evade such problems is to let independent raters simply judge whether a second occurrence gesture was placed consistently or not. Then, the research hypothesis could be that the rate of spatially consistent second occurrence gestures should be near 100% or at least significantly above the rate of spatially inconsistent second occurrence gestures. This hypothesis is also problematic, especially when studying co-speech gestures in general, as many of the gestures do not only carry the information where they are produced. An iconic, for example, transports also pictorial information and if a speaker wants to stress it, he might neglect spatial consistency. This would lower the rate of spatially consistent second occurrence gestures, but it would hardly imply that speakers are not capable of spatial reference tracking. Instead this would suggest that speakers stress certain aspects of gesture information according to their needs.

Finally, the mentioned studies explored spatial reference tracking in general including the topographical as well as the purely referential type. One might speculate, however, about dissociations between both types. For example, it is imaginable that a speaker uses gesture space more consistently when gesturing topographically, because he recalls the real world situation during his narration.

3.3.3 Observations on abstract pointing

The actual piece of research that got me intrigued with abstract pointing and, thus, indirectly with spatial reference tracking was an observational study by McNeill (2003), where he analysed a conversation between two students and suggested that the course of this conversation was only possible, precisely because abstract pointing was used to indicate and infer the referent (for another observational study including abstract pointing but also co-speech gestures in general, see Stec & Huiskes, 2014). Before turning to the details of this analysis, it is important to familiarize with certain aspects, as the dialog included regional, cultural and personal specifics. The conversation took place in Chicago, which is home to the *University of Chicago* and the *Loyola University Chicago*. The latter one was founded by a Jesuit order. At the time of the conversation, both students, Mr. A and Mr. B, were enrolled at the University of Chicago. Prior to that,

Table 3.1: Selection of a conversation studied by McNeill (2003).

	Mr. A	Mr. B
(1)		So, I came back ...
(2)	Oh, uh-huh.	
(3)		... kind of.
(4)	And you went to undergraduate here or ...	
(5)		In Chicago, at, uh, Loyola.

Mr. B had already received education at Loyola University and in Iowa, in this chronological order. While he was willing to give information about his time in Iowa, Mr. B appeared to be quite reluctant to mention his time at Loyola, maybe in order to avoid this "religious" topic. Table 3.1 shows how the conversation went on.

Surprisingly, Mr. B revealed in the end that he had been a student at Loyola. McNeill argues that the speech content alone cannot explain this twist of the conversation. From (1) to (4), it is never explicitly mentioned that the speakers refer specifically to the *University* of Chicago. Instead, the statements can also be interpreted as referring to the *city* of Chicago. Mr. B could have seized this ambiguity and simply answered with "Yes" in (5) (implicitly referring to the city). This way, he would not have lied to Mr. A, but still have occluded his time at Loyola. Despite this opportunity, Mr. B disclosed his past, as if Mr. A had specifically asked him in (4), "And you went to undergraduate here at the University of Chicago?".

According to McNeill abstract pointing can shed light on Mr. B's surprising "confession". Table 3.2 provides exactly the same information as the preceding table and in addition the information of when and where abstract pointing was conducted. In particular, abstract pointing was directed towards two distinct locations – the gesture space labelled "shared" is used by both speakers and the gesture space labelled "right" only by Mr. B.

McNeill suggests the following development of the conversation: The gesture in (6) to the shared location gives similar to speech no information whether Mr. B refers with his statement to the city or to the University. Importantly, however, he conducts a further gesture in (8), which is directed at the right location. This way, Mr. B creates a spatial contrast and reveals that there are two Chicagos for him. Mr. A detects the distinction made by Mr. B; he cannot know the specific meanings, which Mr. B as-

Table 3.2: Selection of a conversation – abstract pointing included.

	Mr. A	Shared	Right	Mr. B
(6)		X		So, I *came back* ...
(7)	Oh, uh-huh.			
(8)			X	... *kind of.*
(9)	And *you went to undergraduate here or* ...	X		
(10)		X		*In Chicago*, at, uh, Loyola.

Note. The labels "shared" and "right" refer to different locations in gesture space and are adopted from McNeill (2003). "X" indicates that the speaker conducted an abstract pointing at the specific location. Italics indicate when the gesture occurred.

sociates with the locations, but he can infer that there is is an ambiguity. In order to clarify, he himself conducts an abstract pointing in (9) to the shared location for a prolonged while. "Mr. A's use is unambiguous: The space means for him the University [...]." (McNeill, 2003, p. 229). In other words, with the help of abstract pointing, Mr. A asks Mr. B whether he did his undergraduate studies specifically at the University of Chicago. Under this condition, Mr. B cannot answer with "Yes", as this would be a lie. Consequently, he reveals that he went to Loyola.

Looking critically at the data, however, there are alternative explanations for the course of the conversation. For example, it could be that Mr. A did not get suspicious because of the gestures, but because of the words "kind of" in (8). They could be regarded as an indication that Mr. B is aware of the fact that his utterance in (6) could be misinterpreted as coming back to the University of Chicago and not the city. In addition, imagine that Mr. A stressed the word "here" in (4) – this could have sufficed to make it clear that he refers specifically to the University of Chicago. In sum, McNeill's analysis of the conversation is certainly suggestive, but it is also just an indication that abstract pointing is utilized for spatial reference tracking.

3.3.4 Summary

Spatial reference tracking is a part of sign language. Due to the close relationship between sign and spoken language, it might be that non-signers also make use of spatial reference tracking – only not via signs, but via co-speech gestures. This idea is supported by the fact that speakers can

produce co-speech gestures in a spatially consistent manner, which is a prerequisite for spatial reference tracking. Aside from that, observational single case studies suggest that speakers use gesture space consistently and that the recipient can infer the referent based on such consistent gesturing.

While these are notable indications that spatial reference tracking might take place in spoken language, there is still ample room to doubt that. For instance, there is no experimental evidence that the recipient can make use of it when the speaker produces spatially consistent gesturing. Seemingly supportive observations such as the one by McNeill (2003) are very helpful, but they also lack clear-cut interpretability and are difficult to generalise. In the end, it is by no means clear whether spatial reference tracking is an actually applied communicational strategy in spoken language. Thus, it is also not possible to decide whether this is the function of abstract pointing.

Therefore, when beginning the work on abstract pointing, a primary interest was to further explore the idea whether it is used for the purpose of reference tracking. A specific goal of this dissertation was to systematically test the hypothesis that the recipient is capable to comprehend spatially consistent abstract pointing, i.e. that he can infer the referent from it. Looking back at the *Chicago conversation*: Did Mr. B truly understand based on the gesture location that Mr. A was specifically referring the University of Chicago? The first detail question was whether the recipient of abstract pointing is sensitive at all for the presumably built-up associations between specific gesture locations and speech referents. To my knowledge, no experimental research had been carried out on that matter before. In contrast to that, a lot of research had accumulated on the comprehension of other gesture types over the last decades, in particular on iconics. This research gives an excellent impression about how experiments on abstract pointing can be designed and is, therefore, presented in more detail.

3.4 The research on iconics as an examplar

The research questions concerning the communicative value of abstract pointing and iconics have things in common. Consider the abstract pointing gesture to the right on the last word in, “Whenever I have enough time to consciously enjoy music, I listen to LPs.” I want to know whether the

recipient extracts a meaning from this gesture, i.e. the object the speaker is referring to, and whether this is beneficial for the speaker. Regarding iconics similar questions had been posed. In the example, "And then I ran over the street", where the speaker moves her hand from left to right with a wiggling index and middle finger, you can also ask whether the recipient extracts a specific meaning from this gesture, in this case that somebody is running, and whether this is helpful for communication.

Before presenting exemplar research on iconics please note for the sake of completeness that there are various other ways, how gestures can support communication (cf. Bavelas, Kenwood, Johnson, & Phillips, 2002; Hostetter, 2011). For instance, they can influence how the speaker is perceived by the recipient, in particular they might let the speaker appear in a more positive way (Maricchiolo, Gnisci, Bonaiuto, & Ficca, 2009; Kelly & Goldsmith, 2004). Furthermore, gestures cannot only have an effect on the recipient, but also on the producer. For example, Rauscher, Krauss, and Chen (1996) could demonstrate that speakers talk more fluently about spatial content, when they can use gestures (see also Frick-Horburry & Guttentag, 1998).

3.4.1 Behavioral research on iconic gestures

Coming back to whether iconics specifically carry a meaning to the recipient, according evidence has been provided by different lines of behavioral research. A first indication comes surprisingly from experiments, which spotlight the producer. The generalized finding is that speakers adapt their gesturing behavior depending on the recipient's situation. For instance, several studies showed that speakers increase their gesturing rate in a face-to-face condition compared to a situation, where the recipient is absent (e.g. Cohen & Harrison, 1973; Krauss, Dushay, Chen, & Rauscher, 1995; Alibali, Heath, & Myers, 2001). Furthermore, it seems that speakers even account for the recipient's cognitive state, as Jacobs and Garnham (2007) could demonstrate that speakers conduct more gestures, when the recipient appears to be attentive. Based on such findings it is tempting to assume that gestures are produced for the recipient's sake and, in turn, that gestures are beneficial for the recipient. This is further supported through a study by Holler and Stevens (2007), which showed that speakers do not only adapt the gesturing rate, but also the semantic content of the gestures. In their experiment, the participants included size information about objects more often in their gestures, when they knew that

the recipient lacks this information. Nevertheless, such findings are only indirect evidence for iconics' communicative value for the recipient. In theory, it could be that gestures are intended for the recipient, but that he cannot utilize the provided information. Moreover, there are also studies, where the proposed effect of visibility on the gesturing rate is either very small or cannot be found (e.g. Rimé, 1982; Bavelas, Gerwing, Sutton, & Prevost, 2008).

The research shifts from the producer to the recipient, when exploring how well a message is comprehended depending on gestures. Broadly speaking an utterance can be divided into the visual or video stream, which contains the gestures, and the audio stream, which contains the speech. If the whole message is presented, recipients have usually no problem to comprehend it. As can be expected, comprehension is decreased in a video-only condition, where speech is missing. Interestingly, however, comprehension is also decreased in an audio-only condition, where the full speech stream is accessible, but not the gestures. Except for one contradicting study (Krauss et al., 1995), this effect has been found by various authors by now suggesting that iconics are indeed beneficial for comprehension (e.g. Graham & Argyle, 1975; Riseborough, 1981; Beattie & Shovelton, 1999, 2005). The question remains whether this facilitatory effect is at least partly due to iconics providing a meaning for the recipient.

According evidence has been accumulated by research on *non-redundant iconics*. These gestures depict an action, object or spatial relationship, which is uniquely specified via the gesture, but not in speech. For example, Kelly, Barr, Church, and Lynch (1999) presented videos of a speaker, who uttered sentences like "The cook stepped outside for a minute", and simultaneously placed one of her hands next to her mouth as if smoking a cigarette. Thus, the information about what exactly the cook did outside was only provided by the gesture. After the presentation of all stimuli, the participants had to write down the *exact words* the woman had said. Despite this specific request, the additional information from the gesture was included in 23% of the responses; for example, a participant wrote, "She said that he went out for a smoke". This effect cannot be due to random errors, because when the participants were only presented with the audio stream, 0% of the responses included the relevant information. This intrusion phenomen caused by non-redundant gestures has also been studied by other authors and the findings suggest in general that semantic information, which is conveyed by iconics, gets incorporated by

the recipient (e.g. Cassell et al., 1999; Beattie & Shovelton, 1999; Church, Garber, & Rogalski, 2007; Broaders & Goldin-Meadow, 2010).

A final indication for this idea is provided by a study applying a mismatch paradigm (Kelly, Özyürek, & Maris, 2010). Here, the gestures are either in accordance with the content of speech (match condition) or in violation with it (mismatch condition). In particular, the participants were first presented with a prime video of an action, e.g. chopping vegetables. Then they saw a target video of a speech-gesture combination. In one of the study's experiments, they were instructed to assess whether specifically the speech in the target video was related to the action in the prime video. For instance, when the person in the target video said, "chop", this was the case, but not when she said, "twist". Interestingly, the participants performance was worse, when the gesture, which was irrelevant for the task, mismatched speech. For instance, when the speech-gesture combination was chop-twist, the participants made more errors than when the speech-gesture combination was chop-chop. This suggests again, that recipients infer the meaning of iconics (see also Cassell et al., 1999). Please note, however, that iconics do not necessarily provide *precise* semantic information. Hadar and Pinchas-Zamir (2004), for instance, presented the participants with short videos, where a man uttered a phrase and conducted a gesture. The specific word, which was accompanied by the gesture, e.g. "snake", was never accessible. After each video, the participants had to pick the right word out of five candidates. Only in 40% of the cases the correct word was selected and in 7% of the cases (i.e. significantly above 0%) a completely unrelated candidate like "tree" was chosen. This suggests that recipients might have an idea about iconics' meaning, but at least sometimes a vague one that can lead to wrong interpretations (see also Krauss, Morrel-Samuels, & Colasante, 1991).

Despite this restriction, the bottom line is that behavioral research has provided a lot of evidence for the idea that recipients profit from iconics and can glean semantic information from them. The according studies outweigh the contradictory results, which is also backed up by a meta-analysis (Hostetter, 2011). With yet another technique – event-related potentials – even more evidence could be accumulated.

3.4.2 ERP research on iconics

Over the last couple of years the ERP method has also been used in order to explore the influence of co-speech gestures on communication. Again,

the lion's share of these studies has dealt with gestures, which have a very pictorial character like iconics. The core idea of the experiments is usually the same. On the one hand the participants are presented with gestures and on the other hand with speech (or another stream of information, which is known to have a communicative purpose, e.g. pictures). There is a congruent condition, where gesture and speech are in accordance with regard to the information they carry. In addition, there is a condition, where the gesture stream is in some way manipulated – either the gestures carry information, which does not fit to the information of the speech channel, or no gestures at all are presented to the participant. In case of an ERP difference between the congruent and the modified condition, this is interpreted as an indication that gestures are taken into account during communication.

An exemplar study is provided by Kelly, Kravitz, and Hopkins (2004). They used a priming paradigm, where videos of a man, who sat in front of a table, were presented. On the table, there was a tall, thin glass and a small, wide dish. In each trial, the target word consisted of the man uttering one of the just mentioned attributes; for instance, he said "tall". In the congruent condition, the preceding prime consisted of the man conducting a gesture, which referred to the same attribute (in this case depicting tallness). In an incongruent condition, the gesture prime referred to another attribute (e.g. depicting smallness). Please note, that the participants' task was related to speech only. They simply had to decide whether the target word referred to the glass or to the dish, so they could have neglected the gestures. Nevertheless the ERPs time locked to the word revealed a more negative N400 for the incongruent condition. These results suggest that gestures can influence how speech is processed. Based on the fact that an N400 effect was elicited, it seems that the lexical access for a word like "tall" is more difficult, when it is preceded by a gesture, which refers to another meaning like "small". Thus, iconics can apparently provide semantic information.

Since this experiment, further studies have found similar results (e.g. Wu & Coulson, 2007a; Lim et al., 2009; Kelly, Creigh, & Bartolotti, 2010). The basic logic of the experiments remains the same, but there is a rich variety regarding several experimental parameters and also regarding the specific ERP results. The following review is supposed to provide an overview of this research and also to give an impression of how an ERP experiment on abstract pointing could be designed.

Stimulus material

As often in psychological research there is a trade-off between controlled and natural stimulus material with each of the two sides having strengths and weaknesses. The already mentioned study by Kelly et al. (2004), for example, set the ground with rather controlled stimuli. They used only a small set of gestures and words, so that a participant was confronted various times with the same items over the course of the experiment. Furthermore, they applied a priming paradigm with just one gesture and one word per trial with both items being presented in sequence and not in parallel. An advantage of this approach is that there is only a small amount of noise in the EEG data, which makes it easy to decide whether there is a difference between the conditions. The external validity, however, is comparably low. A reason for this is that in everyday communication gestures usually do not precede single word statements, but occur simultaneously with the critical word as part of a sentence.

This was accounted for in other experiments like the one by Özyürek, Willems, Kita, and Hagoort (2007). In each trial, the participants listened to a full sentence like "He slips on the roof and rolls down." In the congruent condition, the utterance was combined with a video clip of a rolling down gesture – the index finger making circles and going downwards. In an incongruent condition, the sentence was combined with a semantically incompatible walking across gesture – wiggling fingers moving horizontally. The gesture was aligned with the critical words "rolls down". Again, the ERPs revealed a more negative N400 for the incongruent condition, which indicates that the impact of iconic gestures can still be found in a more natural setting. Yet, the stimulus material had still some artificial characteristics: The face of the gesturing person was not visible, the gestures were acted and not spontaneously produced, and the preparation and retraction phase of each gesture had been removed by means of video editing.

When aiming for a maximum of naturalness, an obvious idea is to use spontaneously produced gestures as stimulus material. This approach, however, has also a problematic aspect. For example, Wu and Coulson (2010) presented videos of a man uttering sentences like "Where there's a green parrot – fairly large." In the congruent condition, he conducted simultaneously an iconic gesture depicting a parrot sitting on a lower arm. Importantly, the video clips were taken from a continuous video recording, where the speaker explained various experiences to an inter-

locutor. The speaker was naive to the experiment's goal and he was not instructed to gesture on purpose. Therefore, it can be assumed that the gesture speech combinations appeared quite natural. The problem in such an experiment is that it is difficult to find equivalent stimulus material for the condition with the modified gesture channel. Ideally, the speaker should utter the same sentence again, but with no gesture or a wrong gesture. Of course, this is unlikely to occur in a natural conversation. Instead, Wu and Coulson (2010) created artificially a gesture absent condition. They combined the utterance with a still frame of the speaker, which was taken from a no-gesture phase. Similar to previous research, the ERPs at the gesture related word showed an N400 effect, i.e. the N400 in the congruent condition was smaller than in the gesture absent condition. It might seem plausible to attribute this ERP difference to the gesture manipulation. It should be considered, however, that the experimental conditions differed also in other regards. For instance, in contrast to the gesture absent condition, the congruent condition showed a dynamic stimulus of a moving person and this could have affected the ERP responses, too.

Experimental paradigm

Another difference between the ERP studies on iconics concerns the experimental paradigm. This aspect is closely connected to the conclusions that can be derived from an experiment. For instance, when finding an ERP difference between an incongruent and a congruent condition, this certainly suggests that gestures are processed and that they can have an impact on speech comprehension. The direction of this impact, however, remains unclear. Does an incongruent gesture interfere with speech processing and, thus, lead to an increased N400? Or does a congruent gesture facilitate speech processing and, thus, lead to an attenuated N400? One way to answer these questions is to include a baseline condition, where no gesture information is available. An example for this approach was already given through the experiment by Wu and Coulson (2010), where the gesture absent condition served as the baseline. The results showed a less negative N400 for the congruent condition compared to the baseline condition. Thus, it seems that speech is easier to comprehend when it is accompanied by a congruent gesture compared to a situation with no gesture information. This would suggest that congruent gestures have

a facilitatory effect. However, the caveat remains that the experimental conditions of this study differed considerably on a perceptual level.

This problem can be avoided by further modifying the experimental paradigm as shown by Holle and Gunter (2007). They took advantage of the fact that speech is sometimes ambiguous at a certain position, but that it gets disambiguated afterwards. In particular, they presented videos of a female uttering sentences like "She controlled the ball, which during the dance [...]". This statement is at first ambiguous, because "ball" is a homonym and can refer to a festive gathering or to a sports equipment. With the word "dance" in the following subclause, however, it becomes clear that the festivity is meant. In one of their experiments Holle and Gunter (2007) explored whether an iconic gesture can influence, which of the homonym's meanings gets activated. In order to test this, the homonym was accompanied by a gesture, which depicted someone dancing – this created the congruent condition (1) – or someone hitting a tennis ball – this created the incongruent condition (2).[1]

(1) She controlled the ball$_{[\text{dancing}]}$, which during the dance [...].

(2) She controlled the ball$_{[\text{tennis}]}$, which during the dance [...].

The ERPs were not taken at the homonym, where the gesture was conducted, but at the disambiguating word "dance". The results showed a more negative N400 for (2) compared to (1). This suggests that the ease of processing the verbal disambiguation "dance" was affected by the gesture at the preceding homonym. In turn, this indicates that the participants indeed used the gesture as a disambiguating cue. Based on these results, however, it is still difficult to decide whether the congruent gesture made the comprehension easier or whether the incongruent gesture made it more difficult.

Therefore, Holle and Gunter (2007) conducted a further experiment that included a baseline condition. In this condition, the gesture was substituted for a grooming movement like scratching one's neck, which was not suited to bias the participants to either of the homonym's meaning. In this context, the advantage of the disambiguation paradigm is that the position of the gesture manipulation (homonym) and the position of the ERP measurement (disambiguating word) are pulled apart. This way, all

[1] When presenting examples of stimulus phrases as in (1) and (2), the meaning of the gestures will be indicated via subscriptions throughout the whole thesis.

three conditions are identical in the relevant ERP epoch and a potential ERP difference cannot be attributed to a perceptual phenomenon. As in the preceding experiment, there was a less negative N400 for the congruent compared to the incongruent condition. Furthermore, the results showed no difference between the incongruent and the grooming condition, which suggests that an incongruent gesture makes speech comprehension not more difficult. In the congruent condition, however, a less negative N400 could be observed compared to the baseline condition. This means that a word like "dance" is easier to process when a preceding gesture biases towards this meaning compared to a meaningless gesture like grooming. Thus, iconic gestures have the potential to facilitate communication.

There is one restriction to this finding. Often, one of the homonym's meanings is dominant over the other one (Twilley, Dixon, Taylor, & Clark, 1994). For instance, when hearing "ball" most people would think of the sports equipment and not of the festive gathering. Holle and Gunter (2007) studied not only the effects of iconics, when the verbal disambiguation indicated the subordinate meaning as in (1) and (2), but also the other case, where the verbal disambiguation refers to the dominant meaning (e.g. "She controlled the ball, which during the game [...]"). In the experiment without the grooming condition, the N400 effect was found for the subordinate as well as for the dominant verbal disambiguation. In the experiment with the grooming condition, however, the impact of iconic gestures was not present for the dominant verbal disambiguation.

Task

Several of the above mentioned studies used tasks, where it was not necessary for the participants to analyze the meaning of the gestures (e.g. Kelly et al., 2004; Özyürek et al., 2007; Wu & Coulson, 2010). Therefore, gestures can influence speech processing even under shallow task conditions. In addition to this finding, some studies have explored in more detail how the task can affect the impact of gestures on speech comprehension. On the one hand there are studies, which show that the integration of iconic gestures is at least to some degree a task-independent process (Wu & Coulson, 2005; Holle & Gunter, 2007). In the relevant experiments, comparable results could be found for explicit task conditions and for shallow task conditions. A study by Obermeier, Holle, and Gunter (2011) demonstrates on the other hand that an explicit task can very well influ-

ence the ERP effects. The authors replicated basically one of the above mentioned experiments by Holle and Gunter (2007). In particular, they used the experiment without the grooming condition except for the fact that they did not present the full gestures, but only gesture fragments. These fragments stretched from the gesture onset up to the point, where the according information sufficed to disambiguate the homonym (this was explored in a pretest). This led to the fact that the fragments ended approximately 1 s before the homonym itself could be identified. In one of the conducted experiments the participants had a shallow task and were asked after some of the trials whether a word or a short video snippet had been part of the preceding stimulus. The according ERP results did not show any differences between the congruent and the incongruent condition. In another experiment, however, the participants had to assess after each trial whether the gesture fragment was compatible with the speech content. Now, the "typical" ERP result could be observed, i.e. a less negative N400 for the congruent condition. Thus, when setting up an ERP experiment on gestures it should be kept in mind that an explicit task can indeed boost the effects.

ERP effects

Another interesting aspect does not concern the experimental design, but the results of the ERP studies. As demonstrated by the above presented experiments the most common finding in the literature is a modulation of the N400 amplitude. In contrast to this, P600 effects are scarce. A notable exception are the studies by the group around Cornejo and Ibáñez. For example, Cornejo et al. (2009) showed short video clips, where a man uttered metaphoric expressions like "Those children are beasts." In one half of these expressions, the second noun was accompanied by a gesture, which matched the meaning of the word; for the remaining half of the expressions, the gesture was incongruent. In addition to a more negative N400, the violation condition elicited also a more positive P600. In line with the existing literature on the P600 component, the authors suggest that the conflict between speech and gesture triggered a reanalysis of the stimulus material. Further experiments have replicated the P600 effect (Ibáñez et al., 2010, 2011). Of course, this raises the question, why some studies report a P600 effect and some do not. So far, however, a convincing proposition to this issue has not been put forward.

Summary

Similar to the behavioral research, the strength of the existing ERP research on iconics lies in its diversity. The integration of semantic information coming from speech and gesture has been demonstrated with artificial and with natural stimuli. The integration process is carried out in different experimental paradigms and under various task conditions. In sum and despite the fact that there are some open questions, there is ample evidence that the recipient processes iconics and extracts meaning from them during communication. Moreover, the recipient can profit from iconics for speech comprehension. As of today, ERP research has therefore shifted from the question *whether* iconics have an impact on communication to the specifics of this usage, for instance the effect of iconics under different communicational conditions (Obermeier, Dolk, & Gunter, 2012).

3.5 The pilot study on abstract pointing

It appears tempting to generalize the effects found for iconics on other gesture types such as abstract pointing. Due to fundamental differences, however, this is not possible. It starts with the fact that abstract pointing's appearance can be regarded as rather subtle in comparison to iconics. With the latter gesture type whole pictures are drawn into the air, whereas abstract pointing is just an unspectacular pointing into the void. Due to this formal difference it could very well be that iconics draw more attention and are, subsequently, subject to more elaborated processing. Another aspect is that the valuable information of iconics is usually on the spot accessible. In contrast to that, abstract pointing requires two steps. First of all, the recipient needs to learn an association between a location in gesture space and a specific referent mentioned in speech. Only then, in a second step, a further pointing towards this location can potentially trigger or support the activation of the according referent. Thus, processing abstract pointing in its entirety requires presumably more investment from the recipient than processing iconics. In consequence, the question whether the recipient learns the gesture speech associations at all was addressed in a pilot study preceding this dissertation (Gunter, Weinbrenner, & Holle, 2015).

Regarding the nature of the stimulus material, we opted for a compromise that appeared comparatively natural, but allowed at the same

Figure 3.3: The three shots applied for the interview – a medium shot of the interviewer, a total shot and a medium shot of the interviewee (conducting an abstract pointing in this example).

time for a high degree of control. In particular, the participants were presented with a staged, ongoing conversation between an interviewer and an interviewee, whose face was blurred due to reasons of stimulus production[2] (see Figure 3.3). The interviewer asked questions about several dualistic topics like *cats and dogs*, *notebook and desktop* or *Donald and Mickey*. Each topic started with an establishing phase that provided the opportunity to learn gesture speech associations. For instance, when mentioning Donald, the interviewee conducted an abstract pointing to the left and when mentioning Mickey she pointed to the right. After two to four establishing gestures per side, it was tested whether the participants had indeed learned the gesturing order *Donald – left and Mickey – right*. Since we were specifically interested whether the gesture speech associations would be learned under conditions where this is not necessary, a mismatch paradigm was applied similar to the procedure used by Özyürek et al. (2007). The interviewer asked a last question like "These two characters – did Walt Disney create them at the same time or was there a gap?". In her response, the interviewee used once again a gesture speech combination. In the matching condition (1), the combination confirmed the gesturing order, whereas it was reversed in the mismatching condition (2).

(1) As far as I know, $Donald_{[\text{Donald}]}$ was created later.

(2) As far as I know, $Donald_{[\text{Mickey}]}$ was created later.

[2]Exactly the same issue is explained in detail for one of the thesis experiments in Section 4.2.5.

The task of the participants was not gesture related, instead the interview was stopped from time to time and they had to answer easy memory questions about the preceding content. Thus, task and experimental paradigm assured that there was never an obligation to process the gestures. Also, processing gestures was never beneficial for comprehension. Nevertheless, ERPs time-locked to the onset of the referent word showed a negative deviation for the mismatch condition from 200 ms to 450 ms and a positive deviation from 600 ms to 800 ms. In other words, the brain reacts differently to a violation of the gesturing order compared to a perpetuation. The negativity was interpreted as an N400 effect indicating that retrieving semantic information about the referent was affected by the manipulation. The positivity was interpreted as a P600 effect suggesting that the integration of the referent into the message is also impacted.

Comparable to the research on iconics, which was not finished after one experiment, several questions remain at this point. For example and as discussed for the ERP research on iconics (see Section 3.4.2), it is difficult to infer from mismatch experiments whether the matching gesture is helpful or whether the mismatching gesture interferes. In ERP terms, was the P600 amplitude, for instance, decreased by congruent abstract pointing or increased by incongruent abstract pointing?

Furthermore, the nature of the associations between abstract pointing and speech are not entirely clear. In particular, the mismatch condition did not only differ from the match condition, because it included a referent conflict. It also differed, because it presented a so far unfamiliar stimulation. Theoretically, it could be that the participants simply learned, that movement in a specific visual field comes with a specific acoustic pattern (e.g. movement in the left visual field comes with the sound of the word "Donald"). Under these circumstances, the brain response to a mismatch trial would deviate, because this complex, but rather perceptual rule is broken. To illustrate that this has nothing to do with processing referent information, consider that it is not necessary to understand the interview language German in order to distinguish a mismatch from a match situation. In this light, it would be reassuring to find additional evidence that abstract pointing is indeed connected to the processing of referent information.

Another open issue concerns the impact of abstract pointing gestures in a *Chicago situation* (McNeill, 2003, see Section 3.3.3). Particularly, it would be interesting to know whether abstract pointing is powerful enough to indicate the referent in a moment of ambiguous speech. The follow-up

experiment, which was the first one conducted as part of this dissertation, was carried out in order to start finding answers to these questions.

Part II

Experiments

4 Experiment 1 – a matter of reference

4.1 Introduction

In the preceding experiment it was explored how abstract pointing is processed, when speech itself provides sufficient referent information. In the present study the focus was shifted to what happens, when speech is underspecified as in the observation made by McNeill (2003), where two students talked about Chicago and it was not clear whether they refer to the city of Chicago or the University of Chicago (see 3.3.3). Is abstract pointing powerful enough in such situations to make the recipient select a referent?

To answer this, the idea of the disambiguation paradigm (Holle & Gunter, 2007) was picked up and a *Chicago situation* (McNeill, 2003) was experimentally generated. Particularly, the establishing phase remained untouched. For example, during the topic *Goethe – Shakespeare*, the interviewee conducted an abstract pointing gesture to the left when mentioning Goethe and an abstract pointing gesture to the right when mentioning Shakespeare. This way, the participants could again build up associations between gesture locations and specific words. The structure of the critical response, however, was modified. For instance, when being asked for her personal preference, the interviewee replied with (1).

(1) Then this classic would win, because I've rarely read something as beautiful as Goethe's Faust.

The first part of this statement is ambiguous, because it is not clear whether the interviewee refers to Goethe or to Shakespeare when she utters "this classic". Later on, the ambiguity is resolved with the explicit mention of the preferred author, in this case the word "Goethe's". Crucially, the interviewee accompanied the ambiguity with an abstract pointing gesture. This gesture was in theory suited to indicate the referent based on the associations from the establishing phase; by systematically

modifying the gesture channel three different versions of the interviewee's critical response were created.

(2) Then this classic$_{[\text{Goethe}]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.

(3) Then this classic$_{[\text{Shakespeare}]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.

(4) Then this classic$_{[_]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.

In the congruent version (2), she conducted a gesture to the side, where Goethe had been established. In the incongruent version (3), she pointed to the side where Shakespeare had been established. Finally, a version was included, where the gesture information was not available to the participants, as the camera was on the interviewer when the interviewee conducted the gesture; an example is given in (4).

If the participants utilize abstract pointing for reference tracking, their assumption about the referent should vary after the initial clause of the sentence. In turn, this assumption should modulate the ease, with which the explicit verbal mention of the referent is processed. For example, if the participants consider Goethe to be the discourse referent after having perceived "this classic" in combination with an abstract pointing to the left as in (2), this should render the later presentation of the word "Goethe's" rather easy. Thus, in order to measure abstract pointing's impact, the ERP at the position of the explicit verbal mention was taken. Please note, that all conditions were identical from a perceptual point of view during this epoch: The participants saw the interviewee uttering the word "Goethe's" while keeping her hands in a resting position. Therefore, in contrast to the preceding experiment potential ERP differences cannot be affected by perceptual processes.

Based on the experience from the first experiment the hypotheses concerned the N400 and P600 component. For the congruent condition (2), a small N400 amplitude for the word "Goethe's" seemed likely, because all relevant features regarding Goethe should already be activated due to the gesture. Likewise the P600 amplitude should be small, because the participants should already know that it is specifically the referent Goethe, whom the interviewee prefers, i.e. it should not be necessary to incorporate new data into the MRC.

In contrast to that, a more negative N400 and a more positive P600 for the incongruent condition (3) were expected. A more negative N400, because the semantic features of Goethe are not activated in the preceding part of the sentence and, thus, the lexical access for "Goethe's" should be more effortful. A more positive P600, because contradicting referent information is provided, which should render the MRC maintenance difficult.

The baseline condition (4) was included in order to make conclusions about the direction of abstract pointing's impact. For example, if congruent abstract pointing is beneficial for the understanding of a verbally ambiguous situation, the word "Goethe's" should be easier to process in the congruent condition compared with the baseline condition, where no gesture information is provided. In particular, an increased N400 and P600 for the baseline condition was to expect, because with the word "Goethe's" relevant features about this author have to be activated and because the mental representation of whom the interviewee refers to has to be updated from the rather vague *classic* to the more precise *Goethe*.

Intuitively, one might also expect that a bad gesture can interfere with speech processing. Hence, there should be an ERP difference between the incongruent and the baseline condition, too. A modulation of the N400 component did not seem likely in this comparison, because in both conditions the semantic features of Goethe have to be retrieved from memory at the position of "Goethe's". A difference in the P600 amplitude, however, was imaginable. Although in both conditions a maintenance of the MRC is necessary, one might speculate that this process is more demanding when two referent informations clash as in (3) than when the referent information has to be refined as in (4).

Of course, the question whether abstract pointing is used as referent indication and the question for the direction of its impact were closely intertwined. For example, it was hardly imaginable that abstract pointing signals the discourse referent to the recipient, but that this knowledge is not helpful.

4.2 Methods

4.2.1 Participants

36 volunteers (half of them females) entered the data analysis. They had a mean age of 23.9 years ($SD = 3.1$) and were all right handed with a mean

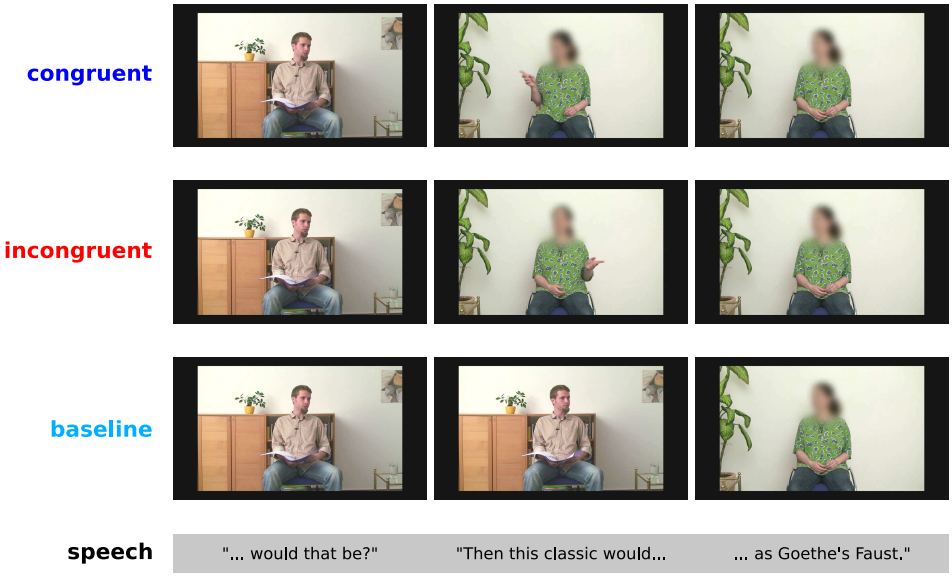

Figure 4.1: Sketch of how the three gesture conditions were created. In the baseline condition, the cut was postponed. In the congruent and incongruent condition, opposing gestures were presented.

laterality quotient (Oldfield, 1971) of 89.4 ($SD = 10.8$). All were German native speakers and had normal or corrected to normal vision. None of them had a hearing or neurological impairment. The participants were paid 7 EUR per hour for participation.

Notably, it was necessary to acquire data from 49 participants in order gain 36 participants for the final analyses. The reason for this high rejection rate was the a-priori decision to include only those participants, who had at least 20 valid trials in all three gesture conditions. Since they were presented with only 30 trials per condition and since it was not possible to tell them, when to avoid blinks and movements, this criterion was easily failed.

4.2.2 Experimental design

The participants watched a conversation between an interviewer (the experimenter) and an interviewee (an actress) about 90 topics. All topics were of dualistic nature, which means that they revolved around two objects like *Freud and Einstein*, *coffee and tea* or *swimming and jogging* (see

Appendix B for a complete list). Each topic started with the *establishing phase*, where the interviewee used abstract pointing consistently in order to establish a gesturing order. During the topic *Goethe and Shakespeare*, for instance, she conducted abstract pointing gestures to the left during her references to Goethe and abstract pointing gestures to the right during her references to Shakespeare. Potentially, this led to the simple order *Goethe – left and Shakespeare – right*. Then, a topic ended with the *critical phase*, where the experimental manipulation took place. The interviewee responded to a last question with a verbal utterance that was at first ambiguous, but then explicit. For example, asked for her preference, she said "Then this classic would win, because I've rarely read something as beautiful as Goethe's *Faust*." From the recipient's point of view it is at first not clear, who she is referring to due to the imprecise term "this classic", which could denote Goethe as well as Shakespeare. Shortly after this, however, the situation becomes clear, as she refers once again to the same object, but this time with the specific term "Goethe's". Please note, that the context always made it clear that the ambiguous term and the disambiguating term referred to the same object.

Aside from verbal aspects, the interviewee also conducted an abstract pointing gesture, which coincided with the verbal ambiguity. Based on the established gesturing order, this gesture indicated either the same referent as the later verbal disambiguation (congruent condition) or the opposite one (incongruent condition). Additionally, there was a third condition, where no gesture information was available to the participants. This baseline condition was achieved by means of video editing. During the critical question only the interviewer was visible on the screen and during the critical response only the interviewee. For the congruent and the incongruent condition, the cut between both shots was placed *before* the gesture was conducted. For the baseline condition, however, the cut was placed *after* the gesture, so that the participants could not observe it. A graphical illustration of all three gesture conditions is provided in Figure 4.1.

The variables gesture (congruent, incongruent or baseline) and side (abstract pointing to the left or to the right) were fully crossed. Therefore, there were six versions of each critical response, an example is given in Table 4.1 (original) and Table 4.2 (English translation). A single participant watched only one version of each topic. Between the participants the six versions of a topic were balanced. An additional variable concerns how many abstract pointing gestures were used for each object during

Table 4.1: Example for All Six Versions of a Critical Response in Experiment 1 – German Original

Gesture	Side	Response
congruent	left	Da würde der Klassiker$_{[\text{Goethe}]}$ gewinnen, denn so was schönes wie bei *Goethes* Faust habe ich selten gelesen.
congruent	right	Da würde der Klassiker$_{[\text{Shakespeare}]}$ gewinnen, denn so was schönes wie bei *Shakespeares* Romeo und Julia habe ich selten gelesen.
incongruent	left	Da würde der Klassiker$_{[\text{Goethe}]}$ gewinnen, denn so was schönes wie bei *Shakespeares* Romeo und Julia habe ich selten gelesen.
incongruent	right	Da würde der Klassiker$_{[\text{Shakespeare}]}$ gewinnen, denn so was schönes wie bei *Goethes* Faust habe ich selten gelesen.
baseline	left	Da würde der Klassiker$_{[_]}$ gewinnen, denn so was schönes wie bei *Goethes* Faust habe ich selten gelesen.
baseline	right	Da würde der Klassiker$_{[_]}$ gewinnen, denn so was schönes wie bei *Shakespeares* Romeo und Julia habe ich selten gelesen.

Note. Subscriptions contain the referent that was indicated by abstract pointing; ERPs were measured at the onset of the italicized word.

the establishing phase of a topic. In order to avoid that the participants could predict when the critical phase arrives, the interviewee conducted differing amounts of establishing gestures. In one third of the topics she conducted two per side, in another third three and in the last third four. Within a topic she always used the same amount of establishing gestures per side. Regarding this variable, there was only one version of each topic. Within a single participant the levels of all mentioned variables were balanced and pseudorandomized.

A last variable refers to the sequence in which the topics were presented to the participants. The experiment was carried out in two sessions. All odd-numbered participants watched the topics 1 to 45 in the first session and the topics 45 to 90 in the second session. For the even-numbered participants the opposite was the case.

Table 4.2: Example for All Six Versions of a Critical Response in Experiment 1 – English Translation

Gesture	Side	Response
congruent	left	Then this classic$_{[Goethe]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.
congruent	right	Then this classic$_{[Shakespeare]}$ would win, because I've rarely read something as beautiful as *Shakespeare's* Romeo and Juliet.
incongruent	left	Then this classic$_{[Goethe]}$ would win, because I've rarely read something as beautiful as *Shakespeare's* Romeo and Juliet.
incongruent	right	Then this classic$_{[Shakespeare]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.
baseline	left	Then this classic$_{[_]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.
baseline	right	Then this classic$_{[_]}$ would win, because I've rarely read something as beautiful as *Shakespeare's* Romeo and Juliet.

Note. Subscriptions contain the referent that was indicated by abstract pointing; ERPs were measured at the onset of the italicized word.

4.2.3 Interview preparation

102 topics of dualistic nature were chosen. They were sent to the actress, so that she could make herself familiar with them. Then, the critical question and the critical response were prepared for each topic. Two specifics about the critical response are noteworthy. First, there were always two versions of it – one referring to the left established object and one referring to the right established object. Second, in contrast to the rest of the material, the stock phrases for the critical response were scripted, so that both versions had the same wording until the verbal disambiguation, i.e. the position where the ERPs were measured. In a third step, the questions for the establishing phase were determined. In general, these were either fact questions ("In which religion is reincarnation an important aspect?" – taken from the topic *Hinduism and Islam*) or questions about personal opinions, preferences and experiences ("What kind of flooring do you

prefer?" – taken from the topic *carpet and parquet*). We also looked for arguments that the interviewee could potentially use during her responses.

Finally, a sequence was determined, in which the interview topics were supposed to be presented to the participants. This was carried out with two goals in mind. On the one hand, topics with a similar subject were grouped together into *meta topics* in order to create a coherent interview. For example, the topics *PCs and Macs* and *notebook and desktop* were put into the meta topic *Computers*. On the other hand, meta topics with related content were not presented in sequence in order to get a diversifying interview. For instance, the meta topic *computers* was not followed by the meta topic *cars*, as this could have been boring for a non-technophile participant.

4.2.4 Shooting

Three cameras were used for the shooting of the interview: one *Sony HDR-HC5E* and two *Sony HDR-HC7E*. All cameras used the recording format *HDV 1080i50*. The resulting videos had a resolution of 1440x1080, 25 frames per second, an aspect ratio of 16:9 and were interlaced. Audio was recorded by the internal microphones of the cameras. Due to the poor quality of these recordings, both interview partners were also equipped with lavalier microphones (a *Sennheiser MKE 2-4 GOLD-C* for the interviewee and a *Sennheiser MKE 2* for the interviewer), which were attached to their clothes. The according audio signal was recorded with a *Roland CD-2 CF/CD Recorder*. The output files had a PCM wav format with 44.1 kHz and 16 bit.

Each camera was installed at a fixed position, the resulting shots were a medium shot of the interviewer, a medium shot of the interviewee and a total shot; examples are provided in Figure 4.2. In general, the interviewer's questions were presented with the medium shot of him or with the total shot and the interviewee's responses with the medium shot of her.

The interview topics were recorded in the sequence, which was predetermined during the interview preparation. All material was recorded at least twice, so that the better variant could be chosen during post production. Altogether, the shooting spanned six days. Care was taken to achieve the same settings during all shooting days.

Before shooting a single interview topic, the interviewer and the interviewee talked informally about it and she was also informed about the

Figure 4.2: The three shots, which the interview was filmed with – a medium shot of the interviewer, a total shot and a medium shot of the interviewee (conducting an abstract pointing in this example).

upcoming questions. If she expressed lack of knowledge regarding a particular question, she was provided with according information. After the cameras and the audio recorder were started, a clapperboard was used. The sharp noise left a distinct marker on the audio recordings of all three cameras as well as of the external audio recorder. Then, the establishing phase and the critical question were recorded in a spontaneous manner. If this run was interrupted (e.g. due to a blooper), the present question answer chunk was repeated and the interview was continued. Subsequently, the establishing phase was again recorded, but this time the actress was told before each question, to which side she should point during the response. This was done in order to assure that it was possible to create a balanced establishing phase during post production, i.e. an establishing phase with the same amount of establishing gestures to each side. Finally, both versions of the critical response were filmed. In one version, she referred for example to Goethe and she pointed to the left side at the ambiguity. In the other version, she referred to Shakespeare and pointed right. The actress was free to modify the critical response after her liking as long as her personal wording until the verbal disambiguation was also identical in both versions. Please note, that the incongruent conditions were created by means of video editing during post production.

Altogether, approximately 55 h of raw footage were filmed.

4.2.5 Post production

The video editing software Final Cut Pro 5 (FCP) was used for the post production of the videos. A first task was to synchronize the video streams coming from the cameras and the good quality audio stream coming from the external audio recorder. In order to achieve this, the clap noise in the audio streams from the cameras and in the audio stream

from the recorder were visually aligned with a precision of half a frame, i.e. 20 ms. Since the video streams of the cameras were locked to their according audio streams, this led automatically to a synchronization of camera video and recorder audio. Finally, the audio streams from the cameras were removed.

After selecting the good material from the raw footage and combining it into coherent topics, another task concerned the generation of the incongruent versions of the critical response. To this end, the audio tracks of the congruent versions were simply switched; during this procedure the gesture stroke was used as reference point. For example, in order to create the version *incongruent – left*, the first step was to mark the frame of the gesture stroke in both congruent versions. Then the video track of *congruent – left* was combined with the audio track of *congruent – right*, so that the gesture stroke frames of both tracks were aligned. Due to this procedure the voice did not match the lips movement in the incongruent versions. In order to conceal this from the participants, a blur mask was put on the interviewee's face (see Figure 4.2, third image). During the experiment, the mask was accepted as a means to achieve anonymity for the interviewee.

90 of the 102 interview topics were selected to be presented during the experiment (11 were discarded and one was chosen as a demo trial). They were grouped into 18 meta topics, which were then exported into video files through a lossless procedure. Using the program FFmpeg these videos were converted so that they occupied less hard drive space. These final videos used the *Audio Video Interleave* container format (AVI) with *Xvid* as video codec and *PCM* (44.1 kHz and 16 bit) as audio codec. The whole interview lasted 120 minutes, the length of the meta topics/single video files ranged from four to eleven minutes.

4.2.6 Procedure

A memory task was integrated into the experiment with the purpose to keep the participants attentive over the course of the interview. After each meta topic, three questions were posed about the preceding interview. Each question was shown in the middle of the screen with one response alternative presented in the lower left corner and another response alternative in the lower right corner. Altogether 54 memory questions had to be answered via a corresponding button press. Immediately after the response, the participants were provided with feedback about whether it

was correct. During the instruction, the participants were informed that accuracy is essential and not speed. Please note, that the memory questions were never related to a critical phase or to visual aspects, including gestures, in general.

The instruction was given in written form with clarifying questions being possible. A cover story purported that the experiment aimed to study memory processes in light of pairwise comparisons. This way, a reason for the memory task and for the topics' dualistic nature was provided. After the preparation for the EEG, the participants were sat in a dimly lit, sound attenuating booth. Subsequently, they were informed about the EEG's susceptibility to body movement in general and eye movements and blinks in particular. Since a meta topic/video lasted several minutes, participants were not asked to avoid blinks completely. Instead they were supposed to keep the blink rate within reasonable limits. Before the experiment was eventually started, a demo trial was presented, which contained an interview topic that was not used for the actual experiment.

The meta topics were distributed over two experimental sessions. Between both sessions there was a gap of two to six days. In each session, the participants were presented with 45 interview topics in nine meta topics/videos. After a meta topic and the according memory task, there was a pause of self-determined length. At the end of the second session, a follow-up questionnaire had to be answered. Both sessions taken together and including the preparation as well as the debriefing procedure, the experiment lasted approximately five hours per participant.

4.2.7 EEG recording and data analysis

The EEG was recorded with 51 Ag/AgCl scalp electrodes, which were positioned according to sites defined in the 10-10-system of the American Clinical Neurophysiology Society (2006). A visualization is provided in Figure 4.3. In addition a ground electrode was attached to the sternum and two reference electrodes to the mastoids. During the data acquisition, the EEG was referenced against the left mastoid, which was recalculated to a linked mastoid reference during the data analysis. The horizontal as well as the vertical EOG were recorded with bipolar electrodes.

The EEG data were rejected off-line through an automatic artifact rejection, which used a 200 ms sliding time window on the EOG (40 µV) and EEG channels (30 µV), and through a manual artifact rejection. In addition, a blink correction was conducted. It was based on logistic regression

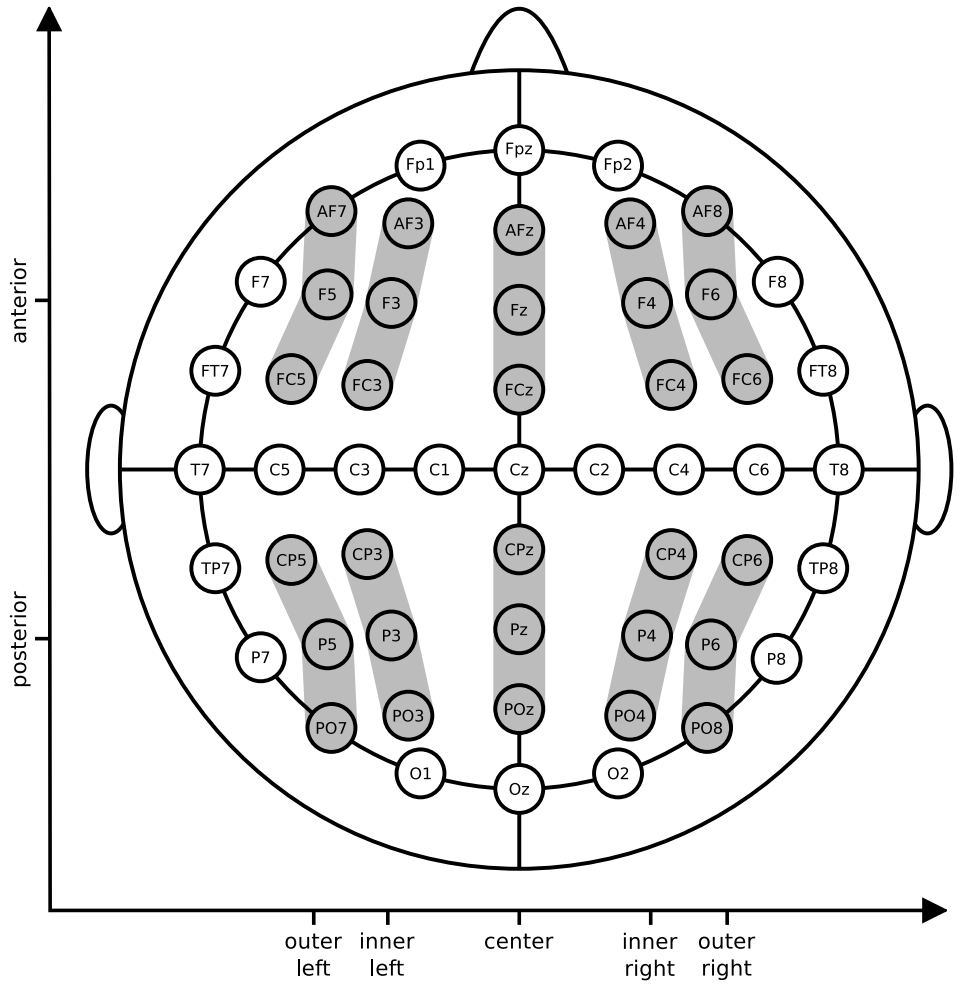

Figure 4.3: Electrodes and ROIs that were used for Experiment 1.

with 7.6% of all trials being corrected. In the end, 16.4% of all trials had to be rejected. On average, each participant provided 25.3 congruent trials ($SD = 2.6$), 25.2 incongruent trials ($SD = 2.7$) and 24.8 baseline trials ($SD = 3.1$).

The ERPs were time locked to the onset of the verbal disambiguation (e.g. "Goethe's"). The epochs lasted from -200 ms to 1000 ms. With the 200 ms prior to word onset serving as baseline, single subject averages were calculated for all three versions of the independent variable gesture (the other variables were not included in the data analysis, as this would have led to an insufficient signal to noise ratio). Ten regions of interest (ROIs) were defined: anterior outer left: AF7, F5, FC5; anterior inner left: AF3, F3, FC3; anterior central: AFz, Fz, FCz; anterior inner right: AF4, F4, FC4; anterior outer right: AF8, F6, FC6; posterior outer left: CP5, P5, PO7; posterior inner left: CP4, P3, PO3; posterior central: CPz, Pz, POz; posterior inner right: CP4, P4, PO4; posterior outer right: CP6, P6, PO8 (for a visualization, see Figure 4.3). After the calculation of the grand averages per ROI a repeated measures ANOVA was conducted. The independent variables were gesture (congruent, incongruent, baseline), coronal (anterior, posterior) and sagittal (outer left, inner left, center, inner right, outer right). All significant effects including the variable gesture will be reported; when appropriate, the p-values were corrected (Greenhouse & Geisser, 1959). After a significant omnibus test, post-hoc comparisons were carried out with adjusted alpha levels: When calculating only a subset of the potential comparisons (as in step down analyses, where one factor is held constant), adjustments were carried out based on Holm's sequentially rejective Bonferroni procedure (Holm, 1979, further on referred to as Bonferroni-Holm procedure). When calculating all potential comparisons, adjustments were instead conducted based on Shaffer's modified sequentially rejective Bonferroni procedure due to its greater power (Shaffer, 1986; Seaman, Levin, & Serlin, 1991; Keselman, 1998, further on referred to as Bonferroni-Shaffer-procedure).

In theory, it would also be interesting to compare the ERPs at the verbally ambiguous position. Here, a difference between the conditions, where a gesture can be seen (congruent/incongruent), and the baseline condition could be expected: In the first case a referent information should be processed by the participants whereas no referent is provided in the baseline condition. Due to methodological problems, however, these comparisons were not carried out. In the congruent and incongruent condition, the participants see the interviewee, who is talking and

in the middle of conducting a gesture. This is in harsh contrast to the baseline condition, where the participants still see the interviewee, who is listening to the interviewer and basically in a resting position. Thus, in addition to the aspect whether a referent is provided by gesture or not, the crucial conditions differ also on a mere perceptual level. Furthermore, one might expect that the participants evaluate the interviewer and the interviewee differently, which could lead to further different associations triggered by the according visual stimulus.

4.3 Results

4.3.1 Memory task

The participants selected in 94.3% of the cases the correct response ($SD = 4.8$; ranging from 75.9% to 100.0%).

4.3.2 ERPs

Figure 4.4 gives an overview of the ERP results. Looking at Figure 4.5, which provides a detail view and topographical maps, differences between the conditions can be observed at electrode CPz, for example. Starting around 550 ms the incongruent condition showed a positive deflection compared to the congruent and the baseline condition. The topographical maps indicate that both deviations were roughly focused between CPz and Pz. Furthermore, looking at electrodes like C3 (Figure 4.5) there seems to be a difference between the congruent and the baseline condition starting around 800 ms.

For the statistical analysis a time window from 550 ms to 750 ms and a time window from 800 ms to 1000 ms were chosen. In the former, the repeated measurement ANOVA Gesture x Coronal x Sagittal showed a significant main effect for gesture ($F(2,70) = 3.27, p = .04, \epsilon = .98$). Post-hoc tests were carried out with alpha adjustments based on the Bonferroni-Shaffer procedure. Since the omnibus test was significant and only three means were compared, the adjusted alpha levels had the same value as the original alpha, i.e. .05. The post-hoc tests revealed a significant difference between the incongruent and the congruent condition ($t(35) = -2.09, p = .04$) and between the incongruent and the baseline condition ($t(35) = -2.49, p = .02$).

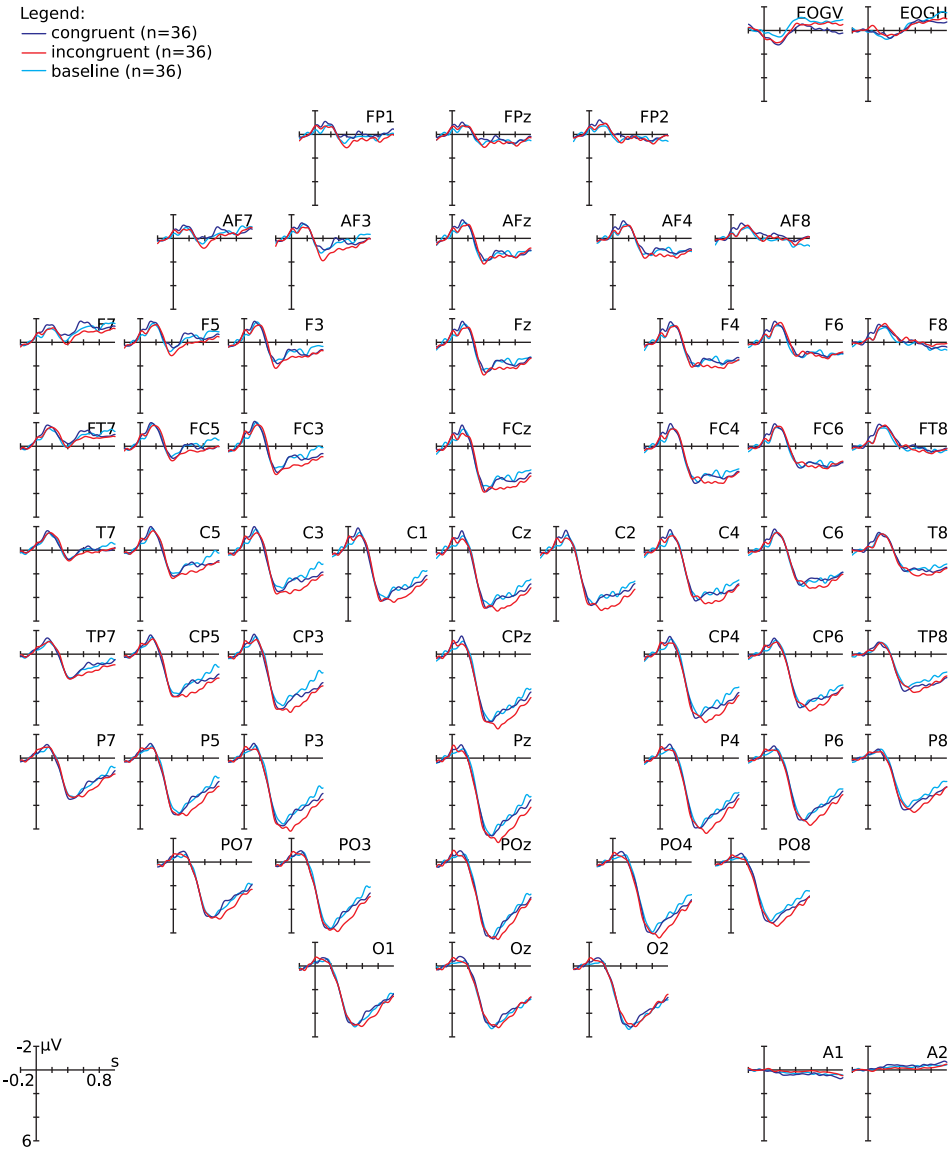

Figure 4.4: ERPs for Experiment 1.

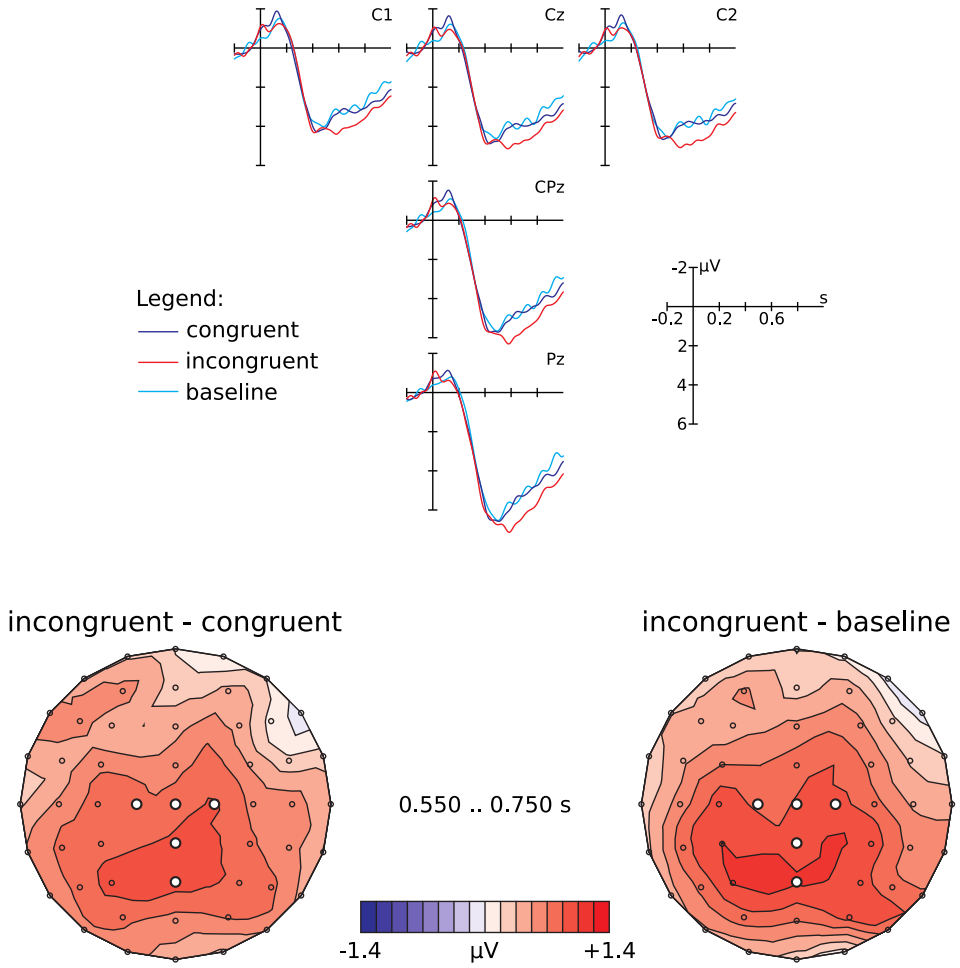

Figure 4.5: Detail view of relevant electrodes and topographical maps comparing the incongruent condition with the congruent and the baseline condition.

The statistical analysis for the time window from 800 ms to 1000 ms did not yield significant results.

To conclude, listening to the verbal disambiguation after an incongruent gesture accompanied the ambiguity, led to a positive deviation from 550 ms to 750 ms compared to the congruent and the baseline condition.

4.4 Discussion

First, it is worth to mention that the manipulation of abstract pointing affected again how an utterance is perceived. In particular, it was shown that an abstract pointing, which is conducted when speech is ambiguous about the referent, can influence how a later verbal mention of the referent is processed. After the prior experiment, this is a second demonstration that abstract pointing is taken into account by the recipient.

A notable difference to the mismatch experiment refers to the temporal relation between the critical gesture and the critical ERP epoch. In the prior study, it was shown that abstract pointing can modulate the brain response to a simultaneously uttered word, i.e. it was effective on a *local* level. In the present study, however, the gesture modulated how a word further downstream the sentence was processed. This demonstrates that the information coming from abstract pointing is kept in memory and, thus, that it is effective on a more *global* level.

Also in contrast to the mismatch experiment, the present result pattern cannot be influenced by any perceptual processes, as the stimulus material was perceptually identical during the relevant ERP epoch in all conditions. The only variable that was systematically modulated from condition to condition was the referent as indicated by abstract pointing. Thus, the fact that there were differences in the ERPs supports the idea that abstract pointing triggers information processing concerning the discourse referent.

When looking in detail at the data, there were both anticipated and surprising aspects. To start with the explicable results, between 550 ms and 750 ms and located around centro-parietal sites there was a positive deviation of the incongruent compared to the congruent condition. Based on latency as well as on topography I regard it as a P600 effect (cf. for example Neville, Nicol, Barss, Forster, & Garrett, 1991; Kuperberg et al., 2003; van Herten et al., 2005; Gouvea et al., 2010) and similar to the preceding mismatch experiment it can be explained with the MRC hypothesis

(Brouwer et al., 2012): In the congruent condition, the verbal mention of the referent, e.g. "Goethe's", states a mere repetition of the referent information that was provided by abstract pointing. Therefore, its integration into the MRC should be rather easy leading to a small amplitude of the P600 component. In contrast to that, the same process appears more demanding in the incongruent condition for two reasons. First, the verbal mention of the referent represents a novel piece of information, i.e. the MRC has to be updated with so far unknown information. Second, this bit of information is not only new, but it also contradicts the already existing MRC, as the preceding gesture had indicated that the interviewee fancies Shakespeare more. It seems suggestive that it is comparably difficult to keep up a coherent idea of what the speaker wants to get across when being confronted with such a dilemma situation. I assume that at least one of these reasons or both rendered the MRC maintenance more effortful resulting in an increased P600 amplitude.

The incongruent condition differed also significantly from the baseline condition. This effect showed the same polarity, temporal pattern and virtually the same topography as the just described incongruent vs. congruent deviation and, consequently, it is again assumed to reflect a P600 process. A notable difference is that the effect cannot be driven by novelty aspects, because in both conditions the critical word "Goethe's" added new information to the MRC. In contrast to the incongruent condition, however, there was no referent conflict in the baseline condition due to the absence of abstract pointing. This supports the idea that a referent conflict alone can significantly increase the effort for MRC maintenance leading to a more positive P600.

Researcher headache begins with the effects, which cannot be observed in the present result pattern. For example, a positive deviation for the baseline condition compared to the congruent condition was expected, because it has been shown that the mere introduction of a referent can result in an increased P600 (Burkhardt, 2006; Brouwer et al., 2012). Furthermore, based on the pilot study I also had expected N400 effects for the incongruent and the baseline condition compared to the congruent condition and they, too, did not show up. Just in general, there was no difference of the congruent condition compared to the baseline condition. This suggests that it does not make a difference whether a correct abstract pointing is provided or no gesture at all.

While this finding in itself would be fine, it is quite tantalizing in combination with the effects for the incongruent condition. In short, the data

suggest that abstract pointing is used in order to infer the referent and if it misleads the recipient about the referent, further comprehension becomes more difficult. However, if the gesture provides correct information about the referent, comprehension does not become easier. A bad gesture can interfere, but a good gesture cannot facilitate. A potential explanation for this counterintuitive finding lies in the fact that abstract pointing was not a reliable referent indicator. As a matter of fact tossing a coin would have had the same predictive power: Whenever an abstract pointing was visible during an ambiguous verbal situation, it was in 50% of the cases misleading about the referent. This unreliability could have affected how the participants processed the gestures and abstract pointing might be beneficial for speech comprehension, when the recipient can trust it. This hypothesis was tested in Experiment 2 of the dissertation.

5 Experiment 2 – reliability helps

5.1 Introduction

Experiment 2 was in almost all regards a replication of Experiment 1. The major modification was to omit the incongruent condition so that the trials were either presented in the congruent or in the baseline condition.[1] As a consequence, abstract pointing was a reliable cue and whenever it was observable in a critical response, it indicated the referent that was later stated by the explicit verbal mentioning. The hypotheses remained the same as in Experiment 1: In the baseline condition, where no gesture information is available, the verbal disambiguation represents the introduction of the referent within the utterance. Compared to the congruent condition, this should elicit a more negative N400, because the semantic features of the item Goethe are not preactivated, and a more positive P600, because the referent has to be incorporated into the MRC.

Grace to the omission of the incongruent condition it became possible to explore an additional issue, which concerned the gesturer's face. In experiments on gesturing one has to decide whether the gesturer's face is visible or not to the participants. The advantage of masking or hiding the face is that found effects cannot be attributed to facial cues that might cooccur with the gestures in a systematic fashion. A disadvantage could be that gesture effects might be artificially boosted. The face is a very informative stimulus. For example, we gain information about the speaker's emotional status just as we use facial signals as cues for auditory perception (Ekman, 1993; McGurk & MacDonald, 1976). Hence, it comes with no surprise that the face draws attention – an effect that can be shown from early childhood on (Ro, Russell, & Lavie, 2001; Valenza, Simion, Cassia, & Umiltà, 1996). It even remains the dominant fixation target when the speaker gestures (Gullberg & Holmqvist, 2006). Therefore, when elimi-

[1] Instead of terming the remaining conditions "congruent" and "baseline" it would be more fitting to speak of a "gesture" and a "no-gesture" condition. However, in order to make the comparison between Experiment 1 and 2 easy, the already established labels are kept.

nating this attentional focus by masking the face, participants could start directing their attention to "the next best thing", which might be gestures. Ultimately, this might produce evidence for gesture processing, which is only due to an unnatural experimental design.

While it has been possible to demonstrate gesture effects with both approaches, i.e. face visible or not (e.g. Kelly et al., 2004; Özyürek et al., 2007), it has not been systematically explored yet, whether the availability of facial information has an impact on the processing of gestures. At least partly this is due to specifics of the according stimuli. In the pilot study and Experiment 1, for instance, it was simply a must to mask the gesturer's face. Please remember, that the incongruent versions of a critical response were created by switching the audio tracks of the according congruent versions. The blur mask was necessary, as otherwise the participants would have noticed the resulting asynchrony between lip movement and speech. For the present experiment, however, with the incongruent condition being left out, there was the freedom to add a condition, where the face of the interviewee is visible. Please note, that this was accomplished with a between subject design. In other words, additional participants were invited, who were not only presented with reliable abstract pointing, but who could also see the interviewee's face. A very striking result would be, if a beneficial effect of abstract pointing would not be detectable for this group, but only for the blurred face group.

5.2 Methods

75 volunteers participated and were paid 7 EUR per hour. 11 participants had to be discarded due to excessive artifacts or problems during experimental presentation. The remaining 64 participants were all German native speakers, were right handed with a mean laterality quotient (Oldfield, 1971) of 92.0 ($SD = 9.7$) and had a mean age of 25.0 ($SD = 2.4$). None suffered from a known neurological or hearing impairment. Vision was normal or corrected to normal.

In general, the same methods as in Experiment 1 were applied. Notable changes are described below.

The stimulus material was modified in two major ways. First, no incongruent versions were presented, which was realized by creating according pseudorandomized trial lists. Second, half of the participants watched the

interview with the interviewee's face being visible. This was achieved by removing the blur mask with Final Cut Pro. A minor modification was that the amount of presented interview topics was reduced from 90 to 84 (cf. Appendix B). This was necessary, as there were only four condition combinations left after the elimination of the incongruent condition (congruent – left, congruent – right, baseline – left and baseline – right). Therefore, the amount of interview topics had to be a multiple of four (and not six as in Experiment 1) in order to achieve a balanced presentation within a participant.

During the EEG recording, additional scalp electrodes were measured (F1, F2, FC1, FC2, CP1, CP2, P1, P2; cf. Figure 2.1), but for the statistical analysis the same ROIs were applied as in Experiment 1 (Figure 4.3). The variable face (visible, blurred) was included as a between variable in the repeated measurement ANOVAs. Artifact rejection was conducted as in Experiment 1 except that it was not necessary to apply a regression based blink correction in order to achieve an acceptable amount of trials per condition and participant. Eventually, 30.6% of the trials had to be discarded and participants contributed on average 29.0 trials to the congruent condition ($SD = 4.9$) and 29.3 trials to the baseline condition ($SD = 5.1$).

5.3 Results

5.3.1 Memory task

The participants responded in 93.7% correct to the questions from the memory task ($SD = 4.5$). The individual performances ranged from 70.4% to 100.0%.

5.3.2 ERPs

Figure 5.1 gives an overview of the found ERPs in Experiment 2. Similar to Experiment 1, a sharp positive deflection occurred in all conditions and at most electrodes roughly between 200 ms and 400 ms. Aside from this common ERP pattern for both conditions, there seem to be two potential candidates for a significant difference between the conditions. First, there was an early negative deflection for the baseline condition, which was most prominent at centro-parietal electrodes. Second, there was a late

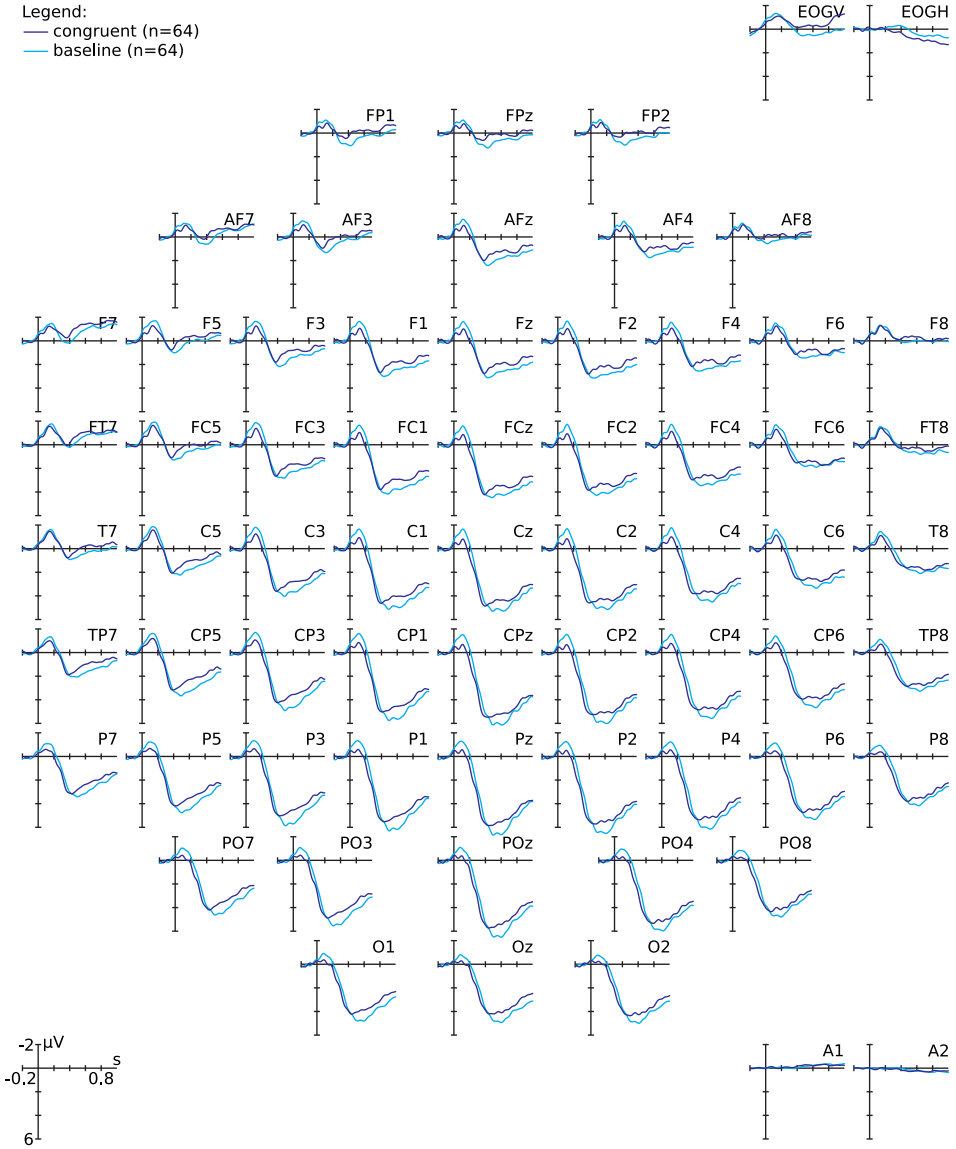

Figure 5.1: An overview of the ERPs that were found in Experiment 2. All measured scalp electrodes are displayed.

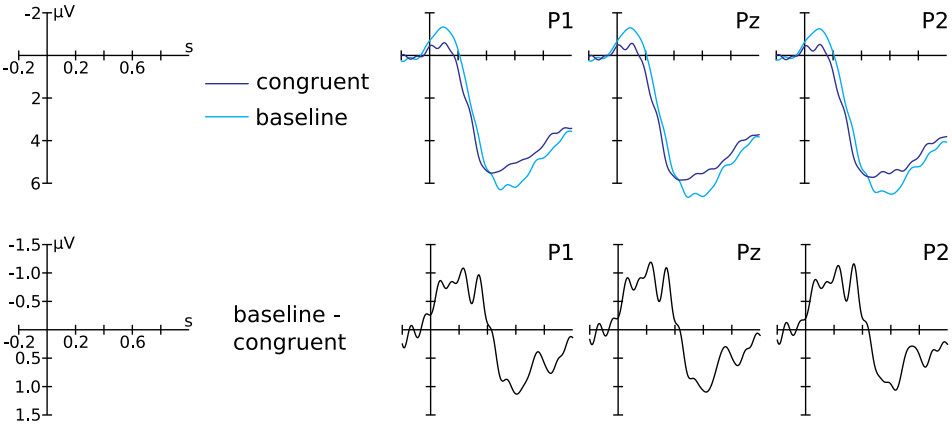

Figure 5.2: The first row provides a detail view of the ERP waves at selected electrodes. The second row shows the difference waves at the same electrodes. Note the different scaling of the ordinates.

positive deflection for the baseline condition, which was best observable at centro-parietal sites over the left hemisphere.

Figure 5.2 provides a closer look at selected electrodes and an additional look at the according difference waves. Solely based on the ERP waves one might have judged that the early negativity was apparent from 50 ms to 150 ms only. The difference waves revealed, however, that the negativity persisted roughly until 400 ms. In the ERP waves this was poorly visible due to the fact that both conditions showed the strong positive deflection in the relevant time span.

When looking at the difference waves it is difficult to tell whether the early negativity represents a single deviation or two negativities smearing into each other. In order to further pursue this issue, the development of the voltage distribution over the scalp was checked, which can be seen in Figure 5.3. When comparing the early part of the negativity (first map) and the late part (third map), it appears that the topographies can be seperated from each other: While the early part is most prominent over central sites with no clear bias to either side, the late part is focused around centro-parietal sites with a bias to the right. This could be used as an argument for two separate effects. However, when taking the whole time series into account it appears that the late positivity is the cause of this putative distinction. The positive deviation started at frontal sites on

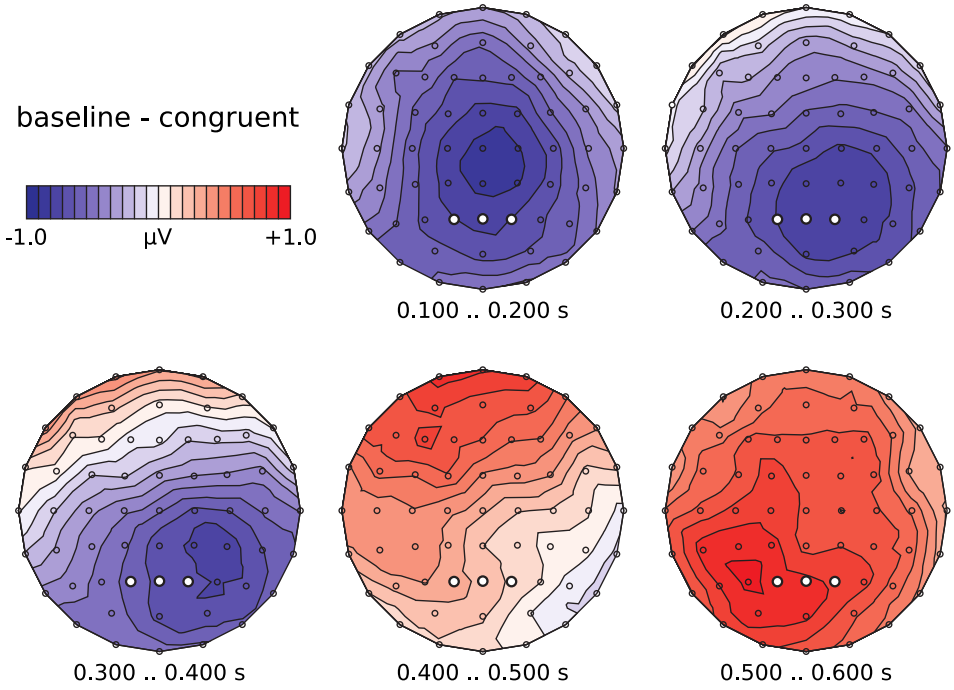

Figure 5.3: Development of the voltage distribution on the scalp in steps of 100 ms. The highlighted electrodes are displayed in Figure 5.2.

the left hemisphere (second to fourth map). I assume that it superseded from this point on the negativity and changed this way its focus.

In the end, I decided to treat the negativity as a single deviation. For the statistical analyses a time window from 50 ms to 400 ms was used. For the late positivity, a time window from 450 ms to 700 ms was chosen. The topographical maps of both time windows can be seen in 5.4. The negative deflection was focused at centro-parietal sites and had a slight bias to the right hemisphere; CP2 and P2 are electrodes, where the effect was most prominent. The positive deflection was focused at posterior sites and had a bias to the left hemisphere being centered around electrodes like CP3 and P3.

50 ms – 400 ms

The repeated measurement ANOVA with the between variable face and the within variables gesture, coronal and sagittal revealed a significant

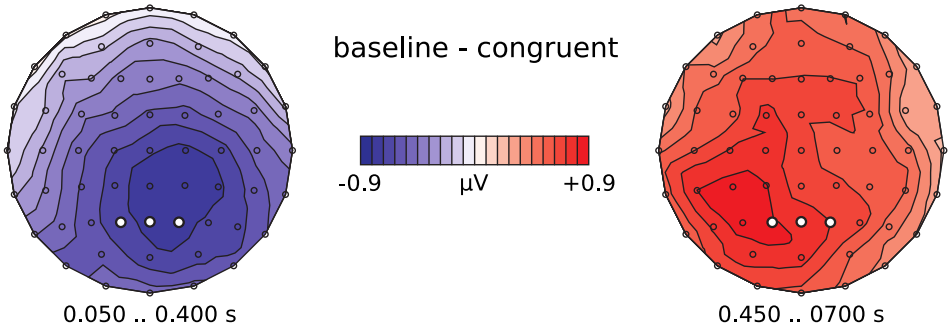

Figure 5.4: Topographical maps depicting the difference of the baseline and the congruent condition for both time windows. The highlighted electrodes are displayed in Figure 5.2.

main effect for gesture ($F(1,62) = 6.53, p = .01$), a significant interaction of gesture with coronal ($F(1,62) = 6.07, p = .02$) and an interaction of gesture with sagittal ($F(4,248) = 3.81, p = .03, \epsilon = .47$). A step-down analysis of the first interaction holding the variable coronal constant was carried out with adjusted alpha levels according to the Bonferroni-Holm procedure. It showed that the difference between the congruent and the baseline condition was present over posterior sites ($t(63) = -3.47, p = .001, \alpha_{adj} = .025$), but not over anterior sites ($t(63) = -1.44, p = .156, \alpha_{adj} = .050$). An according step-down analysis of the second interaction holding the variable sagittal constant showed that the difference between the congruent and the baseline condition was significant over central, inner right and outer right sites (central: $t(63) = -2.77, p = .007, \alpha_{adj} = .010$; inner right: $t(63) = -2.72, p = .008, \alpha_{adj} = .013$; outer right: $t(63) = -2.59, p = .012, \alpha_{adj} = .017$), whereas it slightly failed to reach significance over inner left and outer left sites (inner left: $t(63) = -2.29, p = .026, \alpha_{adj} = .025$; outer left: $t(63) = -1.99, p = .051, \alpha_{adj} = .050$).

450 ms – 700 ms

There was a main effect for the variable gesture ($F(1,62) = 5.11, p = .03$).

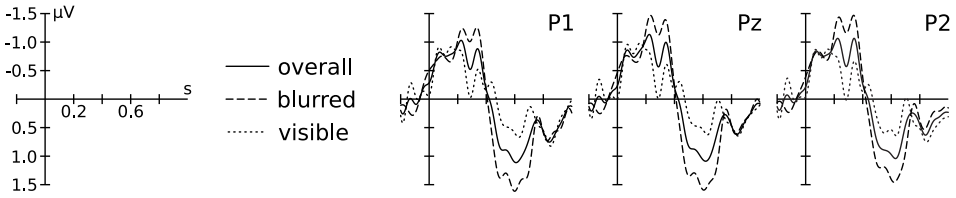

Figure 5.5: The difference waves (baseline - congruent) for both conditions of the between variable face.

Impact of face variable

Figure 5.5 depicts the difference waves for the participants, who did not see the interviewee's face due to the blur mask, and for the participants, who could see the face. Based on this illustration it is very tempting to assume differences between both conditions. In particular, the negative as well as the positive deviation seem to have a smaller amplitude in the face visible condition. In contrast to this, the statistical analyses showed no significant interaction including the between variable face. An explanation for this discrepancy between the visual impression and the statistics lies in the fact that the data showed a high amount of variance. For example, when looking at the late positivity, the difference between the baseline and the congruent condition had a mean of 1.1 and a standard deviation of 2.4 in the face blurred condition; in the face visible condition the mean was 0.3 and the standard deviation 2.3 (calculated over all ROI electrodes).

Summary

When participants were not presented with incongruent abstract pointing, there were clear effects for the baseline as compared to the congruent condition at the position of the verbal disambiguation. In particular, there was an early negativity from 50 ms to 400 ms over posterior sites and a broadly distributed positivity from 450 ms to 700 ms. These results did not interact with whether the interviewee's face was visible or not.

5.4 Discussion

The main conclusion from Experiment 2 is that a congruent abstract pointing can be beneficial for the recipient. This interpretation is based on the

fact that the participants processed the same sentence differently depending on whether they observed an abstract pointing or not. Strictly speaking, finding a difference in the ERPs demonstrates only that congruent abstract pointing has an impact, but it is not clear whether the gestures supported communication or interfered with it. The view, that abstract pointing is facilitatory is driven by two aspects. First of all, given the experimental paradigm I think it is the more intuitive assumption. The ERPs were measured at the word disambiguating the sentence on the verbal level. In the baseline condition this word provided brand new information. For example, until the verbal mentioning of the referent it had not been clear in this condition that the interviewee preferred "Goethe". In contrast to that, the same word provided well-known information in the congruent condition, because the preceding gesture had already revealed that Goethe is the referent. In my opinion, it seems reasonable to assume that already known information is easier to process than new information. In addition to this intuitive argument, I think that congruent abstract pointing is beneficial due to the specific ERP deviations that were elicited in the current experiment.

5.4.1 Early negative deviation

The negative deviation of the baseline condition reflects in my opinion an N400-like effect similar to the one, which was demonstrated in the pilot study on abstract pointing (Gunter et al., 2015). In the congruent condition, an abstract pointing accompanied the verbal ambiguity, which indicated one of the potential referents, for example Goethe. Presumably, such a gesture triggered the retrieval of information about Goethe from semantic memory. Shortly after the ambiguity, the interviewee explicitly mentioned Goethe at the verbal disambiguation. I propose that this second retrieval was comparatively easy, because the relevant information had already been activated. In contrast to this, the information about Goethe had not been preacitivated in the baseline condition. As a result lexical access was more effortful leading to a more pronounced N400 component.

The interpretation of the early negativity as N400 effect is challenged by two aspects. First, the scalp distribution of the negative deviation is more posterior than the prototypical N400 effect. I suggest, however, that this difference can be attributed to to the specifics of the stimuli, as it has been repeatedly shown that the N400 topography varies with stimulus

type (e.g. visually presented words: Kutas & Hillyard, 1984, auditorily presented words: McCallum et al., 1984, concrete vs. abstract words: Holcomb, Kounios, Anderson, & West, 1999, pictures: Ganis et al., 1996, emblems: Gunter & Bach, 2004).

That the negative deviation started already around 50 ms is the second challenging finding, as the N400's onset latency is usually around 200 ms (Lau et al., 2008; Kutas & Federmeier, 2011). One potential explanation is that the applied experimental design allowed for unusually early word identification, which in turn resulted in early lexical access processes. In particular, the following scenario is proposed: During the establishing phase of each topic the participants noticed, which two items the conversation revolved around. Then, in the critical response, the verbal ambiguity signalled that the interviewee will shortly thereafter mention one of these referents explicitly, i.e. she will either utter "Goethe" or "Shakespeare". As a consequence, word identification at the verbal disambiguation was basically reduced to a dual-choice task, which could already have accelerated this process. Furthermore, within this dual-choice task it usually sufficed to hear a fraction of the word in order to decide between both alternatives. Most of the times, even the first phoneme was sufficient as in the example of "Goethe" and "Shakespeare", where both words have quite distinct phonological onsets. Thus, word identification could have been further sped up, as it was possibly based on the initial phoneme only. On top of that, one might even speculate that co-articulation effects allowed the participants to infer the referent based on the transition from the word before the verbal disambiguation to the disambiguation itself. Admittedly, such changes in articulation are subtle, but it has been shown that recipients are sensitive to them (Dahan, Magnuson, Tanenhaus, & Hogan, 2001). Altogether, early word identification seems reasonable. In combination with the finding by van Petten, Coulson, Rubin, Plante, and Parks (1999), who could show that the N400 onset varies depending on how early a word can be identified, this could explain the early N400 effect in the present experiment.

Another potential explanation for the N400's early onset is that the negative deviation comprises not only an N400 effect, but in addition another negativity, which precedes the N400 (for reviews on such early components, see for example Näätänen, Kujala, & Winkler, 2011; Bendixen, SanMiguel, & Schröger, 2012). As noted in the results section it was difficult to decide between these alternatives. In my opinion, a potential candidate is the N100, which peaks in general between 80 ms and 120 ms and

whose amplitude can be affected by selective attention – the more processing resources a participant allocates to a certain stimulus, the more negative the N100 (e.g. Hillyard, Hink, Schwent, & Picton, 1973; Hansen, Dickstein, Berka, & Hillyard, 1983; Lange, Rösler, & Röder, 2003). Specifically in speech perception it has been found that the N100 is increased for the initial syllable of a word compared to subsequent syllables, which suggests that recipients focus their attention on the first phonemes of a word (e.g. Sanders, Newport, & Neville, 2002; Sanders & Neville, 2003; Astheimer & Sanders, 2009). This behavior is explained by the fact that the initial syllable of a word represents its least predictable and, thus, its most informative part (Aslin, Saffran, & Newport, 1999). Importantly, the group around Sanders could also show that the selective attention to word-initial syllables is not an unalterable effect (Astheimer & Sanders, 2011). In particular, they presented their participants with an artificial language, where it was initially neither possible to segment the speech stream into single words nor to predict upcoming information. After a learning process, however, the participants could differentiate the words of the language and they had realized that some words are predictable, because they were always preceded by the same word. The results showed that the N100 was increased for the unpredictable words when comparing the post-learning measurement with the pre-learning measurement. Replicating earlier results this means that as soon as participants can separate words they dedicate more attention to their onsets. The predictable words, however, did not show this increase in N100 amplitude. This suggests that recipients pay special attention to the onset of an upcoming word unless they can predict this word. The described characteristic of the N100 is also applicable to the current experiment. In the baseline condition, the verbal disambiguation was not predictable and, hence, the participants allocated selective attention to its first syllable resulting in a strong N100 component. In contrast to that, the participants could predict the verbal disambiguation in the congruent condition, as they already knew the referent from abstract pointing leading to an attenuated N100.

Mirroring the results section it is difficult to decide for the moment whether the sole N400 account or the N100-N400 account explains the data better.

5.4.2 Late positive deviation

The positive deflection of the baseline condition resembles in terms of topography and latency a P600 effect and is similar to the positivities that were found in the mismatch experiment (see Section 3.5) as well as in Experiment 1, where the incongruent condition differed from the congruent and from the baseline condition (see Section 5). In all three of these instances, the effect was related to a referent conflict between the gesture channel and the speech channel. Obviously, the present deviation cannot be explained with such a gesture-speech conflict, simply because there was no gesture information in the baseline condition. In this light, please remember a core idea of the MRC hypothesis, i.e. that you do not need an exceptionally problematic parsing situation in order to elicit a P600 component (Brouwer et al., 2012). It is rather assumed that the P600 is a response to all words that provide information about what the speaker wants to get across. The more difficult the integration of the respective information into the MRC, the more positive the P600. Speaking in these terms, the positive deviation for the word "Goethe's" in (2) compared to (1) is not assumed to reflect a sincere problem in speech processing, but only an everyday difference in the ease of MRC maintenance.

(1) Then this classic$_{[\text{Goethe}]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.

(2) Then this classic$_{[_]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.

In particular, it is proposed that the participants had already the information whom the interviewee prefers in (1) thanks to abstract pointing. The repetition of this referent information via the word "Goethe's" did not require any modulations of the MRC. In contrast to that, the participants were clueless about the preferred author in (2) up to the explicit verbal mentioning, which made it necessary to add this important information to the MRC. This resulted in the increased P600 for the baseline condition or – depending on from what angle you look at it – in the decreased P600 for the congruent condition, respectively.

5.4.3 Summary of processing events

Taken together, the results of Experiment 2 strongly suggest that abstract pointing influenced how the referent information at the verbal disam-

biguation is processed. Specifically, the following sequence of processes is proposed for the baseline condition: Possibly, the participants paid increased attention to the verbal disambiguation, because they awaited the verbal disambiguation with the referent information. It may be that this led to a more pronounced N100. Then, lexical access was carried out and this retrieval of semantic information was more difficult compared to the congruent condition, as relevant features regarding the referent had not been preactivated. This was reflected by a more negative N400. Lastly, the participants had to incorporate the information about the identity of the referent into their MRC, which elicited a more positive P600.

5.4.4 Face visibility

The current data showed a null effect for the face variable indicating that it made no difference for the processing of abstract pointing whether the speaker's face was visible or not. On first sight, this finding could be regarded as counterintuitive. As noted, the face is an important social cue that draws attention (e.g. Ekman, 1993; Valenza et al., 1996; Gullberg & Holmqvist, 2006) and this could be taken as indication that it should have an influence on the processing of abstract pointing whether the face is visible or not. Especially in case of abstract pointing, however, it appears plausible that the gesture information can even be extracted when the recipient fixates the face and not the hand, as this gesture type presumably does not require a detailed visual analysis. In the current experiment, for instance, all the participants had to do was to differentiate between a gesture to the left and a gesture to the right to gain the relevant information. Of course, it would be interesting to find out whether there would be an interaction for other gesture types such as iconics, as they might require a more fine grained visual analysis, at least in order to perceive detail aspects. For the moment and in case of abstract pointing, however, the hypothesis can be maintained that experimental effects are not boosted by artificial face invisibility or, to look at it from the opposite angle, that they are not endangered by natural face visibility.

6 Interim discussion of the ERP experiments

When learning about the existence of abstract pointing, the initial goal was to find out whether it can really be used for spatial reference tracking. Experiment 1 and 2 together with the pilot study (Gunter et al., 2015) strongly suggest that this is the case. All experiments shared the same basic logic: The participants were given the opportunity to learn associations between abstract pointing gestures and referents. Then a critical phase followed suit, where these associations were either confirmed or not. Importantly, the participants were in all experiments sensitive to this manipulation and while the basic logic was the same, the experiments differed profoundly on other matters such as the experimental paradigm.

In the pilot study (Gunter et al., 2015), the participants were either presented with a referent-gesture combination they already knew from the establishing phase (e.g. *Donald – left*) or with a mixed-up combination (e.g. *Donald – right*). The ERPs deviated in the latter condition, which suggests that the associations were built up and that referential meaning was inferred from the gestures. A counterargument could be that the participants did not learn associations between a gesture location and a referent, but rather and on a more perceptual level the combination of a visual event (movement in left visual field) with an auditory event (something like the spectrogram of the word "Donald"). The ERP deviations could then be explained by the fact that the participants were presented with novel audio-visual stimuli (*Donald – right*) versus already known ones.

Experiment 1 and 2 of this dissertation disprove this perceptual hypothesis. Here, the experimental paradigm was changed to a disambiguation paradigm and accordingly the stimulus material was always the same during the critical epochs. Thus, no perceptual processes can be held responsible for the found ERP effects. In particular, the participants were confronted with speech, which was referentially ambiguous. This situation was either accompanied by abstract pointing or not. Later on, the

referential ambiguity was resolved by speech. When measuring the ERPs at this verbal resolution, different brain processes depending on whether an abstract pointing had been presented or not could be found. Three conclusions can be drawn from this result: First, recipients receive some information from abstract pointing. Second, they keep this information in memory. Third, this information has an impact, when the discourse referent is verbally mentioned. This strongly suggests that abstract pointing can indeed be used to infer the discourse referent. Alternative explanations are hard to imagine. For example, if abstract pointing had no function at all for the recipient, there should be no ERP differences in either of the experiments. The same should be true, if abstract pointing had only the function to highlight a contrast to the recipient.

After the pilot study (Gunter et al., 2015), another question emerged: In what way does abstract pointing influence language comprehension? More specifically, can it make language comprehension easier? To answer this question it is necessary to take the characteristics of the affected ERP components in account, the N400 and the P600. The N400 is related to the retrieval of information from semantic memory – the easier this process, the smaller the N400 amplitude. The P600 is related to the integration of a word into the MRC – again, the easier, the smaller the component. Let's focus on Experiment 2 first: Here, abstract pointing resulted in a smaller N400 and a smaller P600 at the moment of the verbal referent resolution. This suggests, that abstract pointing was not only used to infer the referent at the verbal ambiguity, but that the later verbal resolution was in consequence easier to comprehend, in particular the retrieval from semantic memory and the MRC maintenance were apparently facilitated. In other words, abstract pointing has definitely the potential to be beneficial for speech comprehension.

The picture gets more complicated with the results of Experiment 1, where incongruent gestures were added, which misled about the referent. Here, a more positive P600 compared to the baseline and the congruent condition was found. I.e. the recipient gets into trouble integrating the verbal referent into his idea of what is being communicated, when the referent is finally uttered in the speech stream. This result seems straightforward. However, there was no difference between the baseline condition and the congruent condition. Apparently, the beneficial potential of abstract pointing gets lost, when abstract pointing becomes unreliable.

On a very basic level, this just means that abstract pointing was differently processed in Experiment 1.

6.1 Different processing of abstract pointing

The fact that abstract pointing can be differently processed is insofar not surprising, as this has been repeatedly shown for other gesture types (e.g. Holle & Gunter, 2007; Kelly, Ward, Creigh, & Bartolotti, 2007; Kelly, Creigh, & Bartolotti, 2010). Of course, it would be interesting to know what *different processing* means exactly for Experiment 1. Obviously, it cannot mean that abstract pointing was completely ignored, as gesturing led to a deviation in the brain response. Also, it cannot mean that abstract pointing was treated alike regardless of the experimental condition, because the cognitive system differentiated between the congruent and the incongruent condition. Finally, it cannot mean that selectively the incongruent gestures were processed and their information kept in memory. When the gesture was executed, i.e. at the verbal ambiguity, there was no way for the participants to know whether this particular gesture would turn out to be of the congruent or incongruent type. Hence, even the information from the congruent gestures must have been kept in memory. The only difference then is – to state the obvious – that the information coming from incongruent abstract pointing had a measurable impact at the verbal disambiguation while the information from congruent abstract pointing had not.

In sum, every approach as to what *different processing* means will have to deal with two questions. First, why was the beneficial potential of abstract pointing lost in Experiment 1? More specifically, why was there no difference between the congruent and the baseline condition? Second, why did the incongruent gestures in contrast to the congruent gestures affect speech comprehension, i.e. why was there still an effect for the incongruent gestures? At least three scenarios are thinkable.

6.1.1 Rare processing

The basic assumption of the first scenario is that maybe only a fraction of the participants engaged in spatial reference tracking and/or that the participants utilized abstract pointing in only a fraction of the trials. Considering that abstract pointing was not a reliable referent indicator both variants seem reasonable from the perspective of the recipient. Please remember in this regard, that the ERPs in our experiments are based on an averaging process, i.e. in this scenario they would include trials where abstract pointing was utilized and trials where they were not. In

consequence, differences between conditions such as between the congruent and the baseline condition could become smaller and potentially too small to be detectable.

This would explain the lack of an effect for the beneficial gesture, but so far the question for the origin of the effect for the misleading gesture remains unanswered. This could be due to the incongruent condition prompting the most intensive processing of all conditions. As explained for the hypotheses of Experiment 1 (see Section 4), the MRC is presumably more challenged when a referent conflict arises (incongruent condition) than when a referent is finally revealed (baseline condition) and this should be reflected in a stronger deviation of the P600. Hence, despite of not being processed over the course of all trials/by all participants, abstract pointing was maybe often enough processed to observe the relatively strong P600 deviation for the incongruent condition.

6.1.2 Shallow processing

The assumption of the second scenario is that abstract pointing could have only been processed on a shallow level. This admittedly vague term tries to capture two possible characteristics. In particular it could be that, on the one hand, we can neglect abstract pointing information and, on the other hand, some of this information still seeps in. Support for this idea comes primarily from the field of selective attention in speech processing. In a seminal study on what he termed the *cocktail party problem*, Cherry (1953) could demonstrate both aspects: First, when participants are presented with two different speech streams simultaneously, they can successfully reject one of the speech streams in order to focus on the other one. Second, the information from the rejected stream is not to a 100% blocked and even semantic processing seems to be carried out (for a review, see Bronkhorst, 2015). According evidence has mostly been accumulated using emotional salient stimuli; for instance, a considerable amount of participants notices the appearance of the own name (Cherry, 1953; Moray, 1959; Wood & Cowan, 1995; Straube & Germer, 1979; Nielsen & Sarason, 1981). However, even otherwise neutral stimuli can gain enough subjective importance to show such an effect. For example, city names that had been associated with an electric shock still trigger a galvanic skin response when presented in the rejected stream (Corteen & Wood, 1972; Corteen & Dunn, 1974; von Wright, Anderson, & Stenman, 1975).

Different models have been put forward to explain these phenomenons (e.g. Broadbent, 1966; Treisman, 1969; Deutsch & Deutsch, 1963). What most of them include is at least one filtering stage. This means, at first all incoming information is supposed to receive preliminary processing. Then, one or more filters block a part of the information from entering higher processing while this is granted to the other part of information (for reviews see Pashler, 1998; Driver, 2001).

Applied to Experiment 1, I consider it possible that participants started to reject the incoming information stream from abstract pointing just as a needless speech stream can be rejected. This could account for the missing beneficial effect of congruent abstract pointing, as it has repeatedly been shown that ERP effects such as the N400 or P600 can diminish or disappear due to a lack of attention (Carey, Mercure, Pizzioli, & Aydelott, 2014; Bentin, Kutas, & Hillyard, 1995; Hahne & Friederici, 2002; Kolk et al., 2003). Nevertheless, abstract pointing was maybe still processed to some degree just as rejected speech can be thus far semantically processed as to notice a relevant word. Part of this processing must have been that the fit between gesture and verbal referent was checked. This does not necessarily mean that true referent information was compared. Instead something like the perceptual gist would have sufficed – does movement in the left visual field fit to the spectrogram of the word "Shakespeare's" based on the experience from the establishing phase? When there were signs of an incongruency, the filter let the according gesture information pass to higher processing. Subsequently, the referent conflict was detected leading to problematic MRC maintenance and an increased P600.

This scenario provokes the question why the information of a fit between gesture and verbal referent was not allowed to pass the filter. Potentially, because it is not relevant enough, as it solely confirms the speech referent. In contrast to that, incongruent information carries the threat of a misunderstanding in communication. It is easy to imagine that the cognitive system encounters such a threat with all due respect, i.e. on a higher processing level (cf. van de Meerendonk et al., 2009). With regard to the studies on words associated with electric shock (Corteen & Wood, 1972; Corteen & Dunn, 1974; von Wright et al., 1975), the speculative analogy is that congruent abstract pointing resembles rather a neutral city name whereas incongruent abstract pointing resembles rather the shock associated city name.

A critical difference to classic cocktail party phenomenons shall not be concealed. When noticing that one's own name is uttered, you simply

have to process the rejected speech stream on a word-by-word basis. In the present scenario, the rejected gesture information has to be kept in memory until the verbal disambiguation is uttered. The daring assumption is in other words that the cocktail party phenomenon can also work on a pragmatic level, which would have to be proven in the future. In the end and despite open questions, however, the basic idea still seems plausible to me: Abstract pointing might be rejected on a general level when unreliable, but still to such a degree processed that information of crucial communicational value can be detected.

6.1.3 Reliability updating

Monitoring the reliability of abstract pointing is a mandatory procedure in every attempt to explain the data of Experiment 1. Without monitoring, there would be no notice of unreliability. Without notice of unreliability, there would be no reason to switch the way how abstract pointing is processed. While the preceding scenarios solely implied the idea of monitoring as a prerequisite, the third scenario focuses on it.

In particular, the monitoring itself could be realized via a decay mechanism that takes only incongruent abstract pointing into account. Every time an incongruency is produced by the speaker, this event is put in relation to time leading to an increased P600. If the according frequency of incongruent abstract pointing rises above a certain threshold, the cognitive system completely abandons utilizing abstract pointing for the purpose of reference tracking. This would explain the lack of beneficial effects for the congruent condition.

Despite not being used for reference tracking, the cognitive system might still keep on monitoring its reliability. The reason for this could be that the cognitive system acts on the assumption of appropriate communicative behavior (cf. Grice, 1989). In other words, it assumes that abstract pointing is produced in order to enhance and not to hinder communication. If no instances of incongruency follow suit, the evidence to disregard abstract pointing decays. When the frequency falls below a critical threshold, abstract pointing is again taken into account as referent indicator.

How much delay is needed for such a return to usual processing is difficult to tell. It would certainly depend on the time unit against it is related. Furthermore, other factors could play a role such as the individual variation in the gesturing of a particular speaker, the importance

of the gesture cue for the communicative situation, etc. Looking at Experiment 1, however, a reasonable hypothesis would be that for the first couple of trials of each session, the participants utilized abstract pointing triggering according beneficial effects. Unfortunately, this cannot be explored due to an insufficient signal-to-noise ratio.

6.2 Ways of facilitation

The above presented proposals differ especially regarding the kind of abstract pointing information that is still taken into account and to what end this is done. In *rare processing*, abstract pointing is processed as under reliable conditions, i.e. for the purpose of reference tracking, it simply happens more seldom. In *shallow processing*, only incongruent information passes the assumed filter in order to detect a potential misunderstanding. In *reliability updating*, it is again only incongruent gesture information that is taken into account, but this time solely for the goal of updating abstract pointing's reliability status. Importantly, all scenarios do also share a key element: Abstract pointing was continuously processed.[1] This is insofar remarkable, as it means that the cognitive system invested processing resources, although the information coming from abstract pointing was at best redundant (establishing phase) or useless (critical phase).

An obvious explanation for this continuation would be that the cognitive system expects to profit from abstract pointing in the future. For example and as demonstrated, abstract pointing could facilitate speech comprehension in times of ambiguous speech after becoming reliable again. While this is possible, it is arguable whether such situations happen often enough to make continued processing efficient (cf. So et al., 2009). An alternative explanation is that, perhaps, abstract pointing is also beneficial when it conveys the same information as speech. This would considerably increase its facilitatory capacity.

Looking at all three ERP experiments, 82% of the gestures were conducted when they indicated simultaneously the same referent as speech (all establishing gestures and the congruent critical gestures of the pilot study). Thus, they were redundant and this could be thought of having no beneficial effect. Recent research on the phenomenon of *code-*

[1] To clarify, continued processing means not necessarily that abstract pointing was processed in all trials. Instead, this term denotes the fact that at least a considerable amount of participants kept on processing abstract pointing in a considerable amount of trials (cf. *rare processing*).

blending, however, suggests otherwise. Speakers, who are proficient in a spoken and in a sign language, can simultaneously produce items of identical meaning. For instance, they can say and sign "bacon" at the same time (Emmorey, Borinstein, Thompson, & Gollan, 2008; Bishop, 2010; Emmorey, Petrich, & Gollan, 2013). This code-blending is in clear contrast to code-switching, where a speaker alternates between languages of the same modality, e.g. from spoken English to spoken German. While code-switching is known to generate additional comprehension costs (e.g. Grainger & Beauvillain, 1987; Thomas & Allport, 2000; Alvarez, Holcomb, & Grainger, 2003), there are first results suggesting that code-blending is beneficial: Emmorey, Petrich, and Gollan (2012) conducted a study, where bimodal bilinguals had to fulfill a semantic categorization task, in particular they had to decide whether an item is edible or not. The items were presented as words, signs or code-blends. Results revealed that code-blends were significantly faster processed than word-only or sign-only items. Furthermore, the participants were more accurate in the code-blend compared to the sign-only condition. In short, language comprehension was easier in the code-blend situation.

I propose that something very similar happens in the case of redundant abstract pointing, as it is also conducted simultaneously with the according speech item and refers to the same semantic content. The only difference is, that sign-vocabulary is already known whereas gesture-speech associations have to be learned. Of course, this learning process requires effort, but it might be outweighed by the possible profit. Presumably, redundant abstract pointing represents in the end a code-blend and can similarly facilitate language comprehension, which would add to the reasons why abstract pointing is processed. In Experiment 2, for instance, it could have created additional value by also being helpful in situations, where speech was perfectly clear about the referent. For Experiment 1 one might speculate that the prospect of reliable abstract pointing in ambiguous as well as in redundant code-blending situations represented a significant motivation for continued processing. Finally, in case of the pilot study (Gunter et al., 2015) the idea of beneficial redundant abstract pointing appears to be the only reason for processing the gestures, as abstract pointing was never disambiguating, but only redundant or even confusing in this experiment.

A final implication of the code-blend hypothesis is that abstract pointing could already be beneficial in situations, where the referent is absolutely clear due to specific speech and due to the fact that no other

referent candidate is available, as the discourse revolves around just one referent (for an example of abstract pointing with just one referent candidate see Stec & Huiskes, 2014).

To summarize, the fact that code-blending between sign and spoken language is beneficial in language comprehension supports the idea that redundant abstract pointing is likewise not useless, but facilitatory. In addition to abstract pointing's disambiguating capacity, this could explain the recipient's motivation to utilize abstract pointing.

6.3 Considerations about the P600

6.3.1 The functional role behind the P600

In the introduction on the P600 it was mentioned that there are at least three theories about what kind of processing is reflected by the P600 effect (see Section 2.3). The original view links it to processing faced with syntactic problems (Osterhout & Holcomb, 1992; Hagoort et al., 1993; Kaan et al., 2000). Then there is a family of various multi-stream models, which all assume that the speech stream is analyzed by more than one process and that a mismatch between these processes results in an increased P600 (e.g. Kuperberg, 2007; Bornkessel-Schlesewsky & Schlesewsky, 2008; van de Meerendonk et al., 2009). Finally, there is the MRC approach, which relates the P600 to the maintenance of the mental representation of what is being communicated (Brouwer et al., 2012; Brouwer & Hoeks, 2013). Although the present ERP experiments and the pilot study on abstract pointing (Gunter et al., 2015) were not intended to test these theories, together they appear to shed some light on this issue. For a better overview, the critical situations leading to an increased P600 amplitude are summarized. As usual, italics indicate the words where the ERPs were taken (taken from Gunter et al., 2015, Experiment 1, Experiment 2).

(1) As far as I know, *Donald*$_{[\text{Mickey}]}$ was created later.

(2) Then this classic$_{[\text{Shakespeare}]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.

(3) Then this classic$_{[_]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.

First of all, it seems difficult to bring these examples in line with a syntactic view of the P600. In all experiments the different conditions

featured exactly the same syntactic structure until and including the critical word that elicited the P600 effect. In the pilot study (1), the only difference to the congruent condition is that speech and gesture indicate different referents. While this might be remarkable for the cognitive system, it does not appear to be a true syntactic violation as in, "... the spoilt child *throw* the toy" (Hagoort et al., 1993). In (2) and (3), a grammatically correct word has to be integrated into a sentence, so again there is no reason to assume syntactic difficulties. Overall, the experiments on abstract pointing add to the existing literature, in which it is hard to attribute the P600 effect to a syntactic problem (e.g. van Herten et al., 2005; Regel et al., 2011; Sanford et al., 2011).

Regarding the multi-stream models, it appears that a slightly modified version of the *monitoring hypothesis* (van de Meerendonk et al., 2009) is better suited to explain at least some of the data. According to this model, a sentence like "The fox that hunted the *poacher* ..." elicits an increased P600, as the parser detects a conflict between a purely linguistic analysis and a heuristic analysis, which comes to the conclusion that it should be rather the poacher that hunts the fox. On a more general level, this means that two separate processes come to a different conclusion about the meaning of the phrase and something very similar is imaginable in (1) and (2). Under the assumption that there is also a process, which analyses the gesture stream, a conflict with the speech analysis would result due to a disagreement about the referent. In (3), however, there is no gesture information available that could contradict speech. Therefore, the *monitoring hypothesis* is not applicable to this case and seems in the end not to reflect the processes responsible for the P600 effects at hand.

Altogether, the MRC approach is in my opinion the only one that can explain all of the P600 results. In (1), the MRC maintenance is a more effortful process, because there are two referent candidates. In (3), it is more demanding, because a referent has to be integrated in the MRC as opposed to the other condition of Experiment 2, where the referent is already integrated. For (2), three ideas were sketched trying to explain the P600 effect. *Rare processing* and *shallow processing* ultimately assume problematic MRC maintenance due to a referent conflict similar to (1). *Reliability updating* relates the increased P600 to the recalculation of abstract pointing's reliability. Certainly, it is arguable whether this kind of processing would belong to the maintenance of the MRC, it might be that it reflects a process of its own. However, as the according P600 effect of (2) is similar to the effects of (1) and (3) regarding topography as well

as latency and as *reliability updating* shares with MRC maintenance the characteristic that message relevant information has to be integrated on a pragmatic level, it is suggested that *reliability updating* represents a part of MRC maintenance.

In sum, I am under the impression that the MRC approach explains the P600 effect of the pilot study Gunter et al. (2015), Experiment 1 and Experiment 2 best.

6.3.2 The robustness of the P600

Considering all of the ERP studies on abstract pointing the most reliable effect was a modulation of the P600 amplitude (Gunter et al., 2015, Experiment 1, Experiment 2). This appears insofar surprising, as this effect was not found in the vast majority of the existing gesture ERP studies. In my opinion, this contrast is largely due to the fact that most gesture experiments were simply not designed to elicit a P600 effect.

One reason is that a large body of the experiments utilized a semantic priming paradigm. Most of the times, a single gesture and a single word are presented in these experiments with either matching or mismatching semantic content (Bernardis, Salillas, & Caramelli, 2008; Gunter & Bach, 2004; Habets, Kita, Shao, Özyürek, & Hagoort, 2011; Kelly et al., 2004, 2007; Kelly, Creigh, & Bartolotti, 2010; Wu & Coulson, 2011). Sometimes a more complex procedure is applied, where the prime consists for example of a full phrase (Wu & Coulson, 2007a, 2007b, 2010). In general, such experiments are not known to elicit a P600 effect in the domain of pure speech comprehension, so without further assumptions it cannot be expected in gesture experiments. A potential explanation for the lack of P600 effects is that at the time of the ERP measurement, i.e. at the target, the parser is not put in a situation, where it tries to maintain an MRC. Consider, for instance, an example of complex priming (Wu & Coulson, 2007a): The prime consists of a video showing a man who says, “It’s actually a double door.”, while conducting a fitting gesture. Subsequently, the target is presented, in particular a picture of a single door. While the prime certainly triggers the construction and maintenance of an MRC, the parser does presumably not try to integrate the rather detached presentation of a picture into this MRC. The proposition in short is that the target is not considered to be part of the message and, thus, no MRC/P600 effect can be found in these experiments.

Furthermore, the rarity of P600 effects is probably in parts due to task specifics. As outlined in the discussion of *shallow processing* (see Section 6.1.2), the P600 effect elicited by speech is sensitive to how hard the participants try to understand what is communicated; the harder they try, the more likely the P600 effect (Brouwer et al., 2012; Hahne & Friederici, 2002; Kolk et al., 2003). Applied to gesture studies, some of the P600 absences could be explained by an according lack of effort. For instance, in the second experiment of Holle and Gunter (2007), participants were solely asked to judge whether they had heard a particular word or seen a particular arm movement in the preceding trial. To be successful in such a task a rather superficial processing suffices (see also Obermeier et al., 2011, 2012). Likewise, in experiments where participants are asked to "listen and watch carefully" with otherwise no specific task and no measurement of how well they truly paid attention, the participants' effort might not be strong enough in order to elicit a P600 effect (Özyürek et al., 2007).

As opposed to the above stated, in all abstract pointing experiments the ERPs were taken during the presentation of an ongoing conversation between two speakers, so at a position where the current item belonged obviously to the preceding discourse/message. Moreover, due to the memory task the participants were forced to follow the content of the conversation and the according accuracy rate was always above 90%. Hence, it is safe to say that the participants were engaged in updating a continued MRC during the critical ERP epochs. In other words, the prerequisite to find a P600 effect was given.

Finally, a third explanation for the lack of P600 effects might lie in the nature of the stimulus material itself. In most experiments so far, iconics were studied, which are insofar known to have a vague meaning, as they can often be interpreted in more than one way (Hadar & Pinchas-Zamir, 2004). For example, when the hands of a speaker make a roundish shaped gesture, such an iconic gesture could refer to an apple, a ball, the globe, etc. Additionally, iconic gestures can depict semantic aspects that are not covered by speech. For instance, when making a type writing movement while saying, "... and then he wrote a letter", the gesture channel reveals that a keyboard was used (Cassell et al., 1999). Since iconics are so variable in their relation with speech, it is possible that they readily trigger lexical access, but are not used by the parser to restrain the MRC substantially. This could explain some of the remaining sole N400 results (Holle & Gunter, 2007; Obermeier et al., 2011, see Experiment 1 in both

cases). Abstract pointing in contrast is quite precise about its meaning. After the establishing phase it is clear, for example, that a pointing to the left refers to Goethe. When such a gesture is paired with speech referring to Shakespeare, there is no room for an integration free of conflict.

In the end, I do not believe the reliable emergence of P600 effects with abstract pointing as surprising. Instead it is suggested that their robust appearance is due to the experimental design and the nature of the stimulus material.

6.4 Conclusion

The most important finding of the current ERP experiments is that abstract pointing indeed conveys referent information. Under reliable gesture conditions, the recipient takes advantage of this capacity and utilizes this gesture type to infer the referent. Under unreliable conditions, referential information is still taken into account, but in a different manner. While the exact purpose of this continued processing is not clear, it can be assumed that it is at least done in order to monitor abstract pointing's reliability.

7 Experiment 3 – flexibility and interindividuality

7.1 Introduction

The idea for Experiment 3 originated from conversations that I had with different researchers on the preceding ERP data. A recurring question was whether the recipient requires specifically abstract pointing in order to engage in spatial reference tracking or whether other stimulus types, for example non-natural stimuli, would also trigger this behavior. Such dissociations exist in other areas. For example, while the underlying neural networks seem to be the same, it appears easier to process the action of grasping, when the stimulus is a human hand and not a mechanical claw as shown by studies on infants as well as on adults (Gazzola, Rizzolatti, Wicker, & Keysers, 2007; Daum & Gredebäck, 2011a, 2011b). Most important for the present purpose is a study on beat gestures by Holle et al. (2012). The main finding was that this gesture type can influence the syntactic processing of a sentence. Particularly, it can facilitate the comprehension of a non-preferred word order by indicating the noun phrase. Interestingly, however, this beneficial effect disappeared, when the video stream of the clips, which served as stimulus material, was exchanged for a black screen that depicted solely a moving dot, which followed the up and down trajectory of the original beat gesture.

Just in general it can be asked for the specific characteristics of a stimulus in order to make its comprehension easier. In this line it has been shown for *concrete pointing*, that spatial attention is modulated to a greater extent when the gesture is carried out with an extended index finger and not, for example, with an extended pinky (Ariga & Watanabe, 2009). Furthermore, at least for infants there is also an interaction of gesture with speech. At the age of twelve months, they can usually shift their attention towards the direction of a pointing gesture. Interestingly, this effect is largest when supported by an appropriate utterance (“Look, there!”), but it is absent without the speech (Daum, Ulber, & Gredebäck, 2013). This

finding might be related to the fact that in most cases concrete pointing is accompanied by speech, often in order to clarify the intention of the gesture (cf. Tomasello et al., 2007, see also Section 3.2.1). Such results cannot be attributed to differences in *naturalness*. After all, it can happen in a natural environment that pointing is executed with the pinky or without speech. On a more general level, however, the experiments suggest that there is something like a prototypical concrete pointing gesture that can be processed best.

Specifically in case of abstract pointing the question about the effectiveness of non-protoypical stimuli can be translated into terms of learning psychology. Procedurally, the association of words and locations in gesture space is equal to *classical conditioning* (Pavlov, 1927/1940; Mitchell, de Houwer, & Lovibond, 2009). The fundament of this learning type is an unconditioned stimulus, which has already some kind of relevance for the cognitive system and, therefore, triggers the unconditioned response. Learning is achieved by pairing the unconditioned stimulus with a to-be conditioned stimulus. After repeated pairing, the conditioned stimulus alone evokes a conditioned response, which is identical or similar to the unconditioned response. In the most famous example, the unconditioned stimulus food was jointly presented with the conditioned stimulus of ringing a bell. Initially, only the food led to the unconditioned response of increased salivation in dogs, but after the pairing procedure, the ringing bell sufficed to elicit the conditioned response salivation (Pavlov, 1927/1940). In spatial reference tracking, the unconditioned stimulus is represented by a referent word like "Schwarzenegger". The utterance of this word will usually lead to the unconditioned response, which consists of retrieving semantic information about Schwarzenegger and the integration of the referent into the sentence. Abstract pointing plays the role of the conditioned stimulus and after pairing it will lead on its own to the conditioned response of lexical access and referent integration as suggested by the ERP experiments.

One of the early ideas about classical conditioning was the equipotentiality premise. In a nutshell it suggested that all combinations of conditioned stimulus and unconditioned stimulus are approximately learned with the same ease. Accordingly, it should not matter whether you pair a specific audio-visual stimulation or a specific taste with the unconditioned stimulus of feeling sick (e.g. elicited via the injection of a toxin). Experimental evidence, however, suggests otherwise, as rats show the conditioned response of avoidance behavior mainly to the taste (Garcia

& Koelling, 1966). Findings such as these led to the notion of *selective associations*, which means that certain stimulus pairings are more easily learned than others (Seligman, 1970). Especially in clinical psychology this idea had considerable impact, for example when trying to explain why phobias are rather related to spiders than to cars (Seligman, 1971; Öhman & Mineka, 2001).

Hence, speaking in terms of classical conditioning, the research question of the present experiment was whether the pairing of a referent with an abstract pointing gesture is a selective association and, thus, easier to learn for the recipient than the pairing with another stimulus type.

This issue was explored with a behavioral experiment, where abstract pointing was put into contrast with an artificial stimulus similar to the one used by Holle et al. (2012). In a natural gesture condition, the participants watched interview videos as in the preceding experiments. In a non-natural dot condition, the visual input was drastically reduced and the gestures were replaced by dots flashing up on the left or right side. In order to find out whether the participants engaged in spatial reference tracking, they were asked for the present discourse referent, when the interviewee was ambiguous about it. An example is provided in (1), which is taken from the topic *Schwarzenegger and Stallone*, where the interviewee had to state, whom she prefers as an actor.

(1) Well, then this muscleman$_{[\text{Schwarzenegger}]}$, because at least [cut]

The video was stopped at the position of the cut. Based on the ERP results, it was expected that a recipient of the gesture condition would respond in line with the location of abstract pointing and pick Schwarzenegger as the presumed referent. If the association of referents and abstract pointing is a selective one, participants of the dot condition should show a significantly weaker or no response bias at all.

While the idea of this experiment seems attractively easy, the difficulty of the task could pose a problem. In fact, it might be that – regardless of the experimental condition – all participants always pick the correct referent, which would make it impossible to find differences between the conditions. Please put yourself in the position of the participant. In the dot condition, you are staring at an almost empty screen while listening to a conversation about Arnold Schwarzenegger and Sylvester Stallone. Whenever the interviewee utters "Schwarzenegger", a clearly visible dot flashes up on the left side of the screen, whenever she mentions Stallone, the dot appears on the right side. Then the interviewee says that she

prefers *this guy* as an actor with a dot popping up on the left side... At least to me as the experimenter, the referent seemed very easy to determine. In the end, however, I was taught otherwise and an experiment on the selectivity of gesture speech associations was enriched by findings on interindividual differences.

7.2 Methods

7.2.1 Participants

96 participants entered the data analysis, half of them were female. Their mean age was 23.9 years ($SD = 2.8$) and 14 were left handed. Their vision was normal or corrected to normal and they did not suffer from known neurological or hearing impairments. All participants volunteered and were paid 7 EUR per hour for participation.

7.2.2 Experimental design and stimulus material

Twelve interview topics in the congruent version were selected from Experiment 1 (see Appendix B) and presented in a pseudorandomized order to each participant. The amount of establishing gestures (two, three or four per side) was balanced between these topics. Each topic was modified so that the interviewee's critical response was not observable to its full extent. In particular, the ambiguity including the abstract pointing was still shown, but then the video stopped so that the verbal disambiguation was not presented. This way, the critical response remained ambiguous from a verbal point of view. The only means to infer the referent was provided by abstract pointing during the ambiguity. Regarding this critical gesture, half of the participants watched a specific interview topic with an abstract pointing to the left and the remaining half with an abstract pointing to the right. Within a participant the amount of topics with the critical response to the left and to the right was balanced.

The experimental variable of interest was that half of the participants watched regular videos and the remaining half videos, where the video track was modified.[1] In particular, only a grey frame was shown and whenever a gesture had occurred in the original video track, a dot was shown. Due to their size and color the dots were well noticeable against

[1]The interviewee's face was blurred in the gesture condition, as the effect of the face variable in Experiment 2 had not been analysed when Experiment 3 was launched.

Figure 7.1: Screenshots of the gesture and the dot version.

the background. They were presented either on the left or on the right side of the screen depending on the side of the according gesture in the original video track. A comparison of the gesture condition and the dot condition is provided in Figure 7.1. The dots were presented as long as the according gestures from the original videos, i.e. from the first frame that depicted a noticeable arm or hand movement to the frame, where the hand went into resting position again. The audio track of the dot videos was not altered.

The grey frame in the dot condition was included in order to signal to the participants that the screen was actually used throughout the experiment. Otherwise the screen would have been completely empty until the first appearance of a dot, which could have suggested to the participants that it is not necessary to observe it. The frame was present from the beginning of the video presentation and had exactly the same size as the videos with the gestures.

7.2.3 Procedure

Each participant was sat in a dimly lit room with a distance of approximately 120 cm to the computer screen. A button box, where the buttons were not placed in a left to right manner, but in a top to bottom manner, was placed before them on the table with the screen. The participants were orally instructed that they would be presented with a video of an interview, that they should follow the interview attentively, that they would be asked questions about the interview from time to time and that reaction time was not important, but accuracy. Then, the experimenter left the room and the participants were free to start the interview.

A trial began with a progress bar, which lasted 1000 ms and indicated that a video was loaded into memory. Then a video with a single inter-

view topic was presented. During the critical response the video stopped between the ambiguity and the verbal disambiguation. This was followed by a screen with a question, which referred to the ambiguous situation; it was placed in the upper half of the screen. Simultaneously, both response alternatives were presented vertically aligned in the lower half of the screen. Coming back to the topic *Schwarzenegger – Stallone*, for instance, the question to the participants was "Who does she think is the better actor?" and the response alternatives were "Schwarzenegger" and "Stallone". Whether a specific response alternative was presented at the upper or lower position was balanced between the participants. The question and the response alternatives were presented until the participant indicated a response via a button press. Then, the next trial was started.

At the end of the experiment, I asked three questions to the participants. First, I wanted to know whether they had started at some point throughout the experiment to use a systematic strategy when picking the responses. If they confirmed this, they were asked to describe this strategy. Finally, the experimenter explained to the participants how they could have answered correctly, i.e. through spatial reference tracking. When it was clear that the participants had understood this potential strategy, they were explicitly asked whether they had applied it or not. Participants, who did so, were termed *detectors* (irrespective of when they had started to use the gestures/dots as a cue) and participants, who did not use the gesture/dot strategy, were termed *non-detectors*.

7.2.4 Data analysis

As alluded to in the introduction, most of the participants were thought to be detectors, i.e. sooner or later they were expected to engage in spatial reference tracking in order to answer the questions. I saw a chance for differences, however, during the first trials of the experiment. In consequence, analysing overall values (e.g. all correct responses of the dot group vs. all correct responses of the gesture group) did not seem appropriate and, hence, all trials were separately analysed with two types of statistical tests. First, chi-square tests were conducted for each trial in order to infer whether the amount of correct responses and the cue type (gesture/dot) were dependent of each other. Correction for multiple testing over twelve trials was accomplished via the Bonferroni-Holm procedure (Holm, 1979). Second, binomial tests were conducted per trial

for the gesture and the dot condition separately. This was meant to infer whether the performance of a group was significantly different from chance level. In this case, the Benjamini-Hochberg procedure with q being set at .05 was applied in order to correct for multiple comparisons (Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001). The change in the correction procedure was motivated by two aspects. First, the Benjamini-Hochberg procedure has more power. Second, considering the expected ceiling effect after a couple of trials, a lot of significant results were anticipated. In this light a potentially increased rate of Type I errors seemed bearable.

7.3 Results

In Figure 7.2 the learning curves for the gesture condition and the dot condition are presented. Both groups show an increase in correct responses over the trials. While only 37.5% of the gesture group and 60.4% of the dot group gave the correct response in the first trial, 77.1% and 70.8% answered correctly in the last trial. The gesture group showed a significant above chance performance for the first time in trial 5; the dot group did so in trial 2. Subsequently, both groups returned to chance level. A stable performance with significantly more correct than incorrect responses was achieved in trial 10 for the gesture group and trial 4 for the dot group. Please refer to Table 7.1 for an overview of the amount of correct responses and according p-values per group and trial. From a descriptive point, the greatest difference between both groups appeared in trial 1, where the dot participants showed a higher accuracy. Though the according p-value was below the conventional alpha-level, it failed to reach significance after Bonferroni-Holm correction, which resulted in an adjusted alpha level of .004 ($X^2(1, N = 96) = 5.04, p = .025$). Therefore, I suggest to regard the difference only as a tendency. No other trial showed similar results. In sum, the performances of both groups were always comparable.

Table 7.2 reveals that roughly half of the participants in the gesture group and exactly half of the participants in the dot group did not detect the strategy how to answer the questions correctly (a chi-square analysis of this distribution indicated no significant difference: $X^2(1, N = 96) = 0.04, p = .84$). Since there were enough cases in each cell of the table, the detector and the non-detector group were analyzed independently.

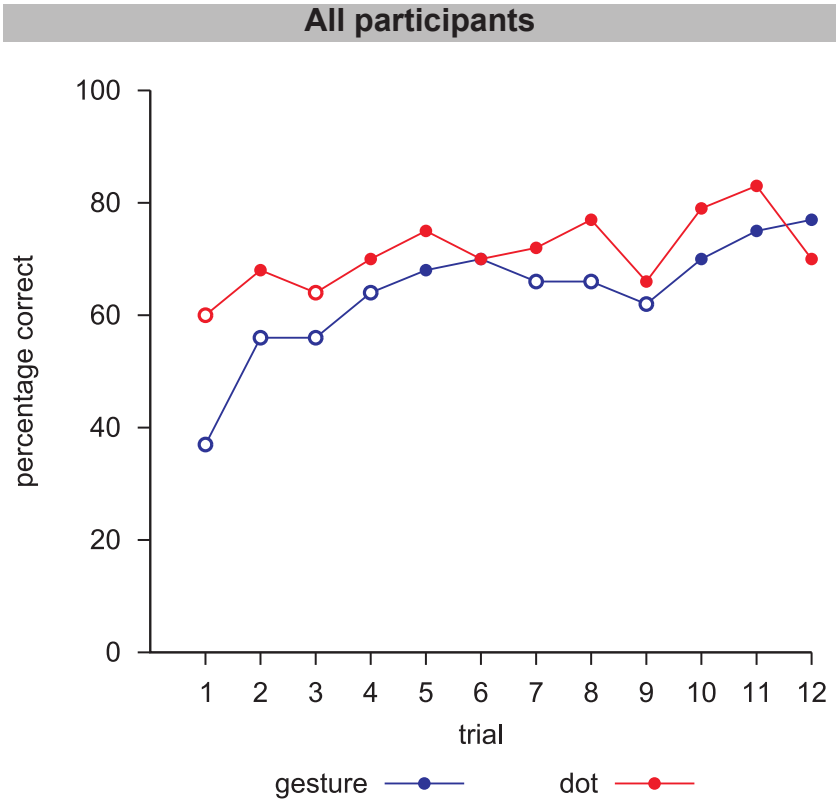

Figure 7.2: Percentage of participants, who gave the correct response. Filled data points indicate that the performance was significantly different from chance.

Table 7.1: Correct Responses and p-Values for the Binomial Tests

Trial	Gesture			Dot		
	N_{cor}	p-value	α_{adj}	N_{cor}	p-value	α_{adj}
1	18	.111	–	29	.193	–
2	27	.471	–	33	.013*	.038
3	27	.471	–	31	.059[a]	.046
4	31	.059	–	34	.006*	.033
5	33	.013*	.021	36	.001*	.017
6	34	.006*	.017	34	.006*	.029
7	32	.029	–	35	.002*	.021
8	32	.029[a]	.025	37	<.001*	.013
9	30	.111	–	32	.029*	.042
10	34	.006*	.013	38	<.001*	.008
11	36	.001*	.008	40	<.001*	.004
12	37	<.001*	.004	34	.006*	.025

Note. Maximum of correct responses was 48 per group and trial
[a] $p > \alpha_{adj}$ leading to abortion of Benjamini-Hochberg procedure
*significant at false discovery rate of .05

Table 7.2: Contingency Table depicting Detectors and Non-Detectors per Condition

	Detector	Non-detector	Sum
Gesture	25	23	48
Dot	24	24	48
Sum	49	47	96

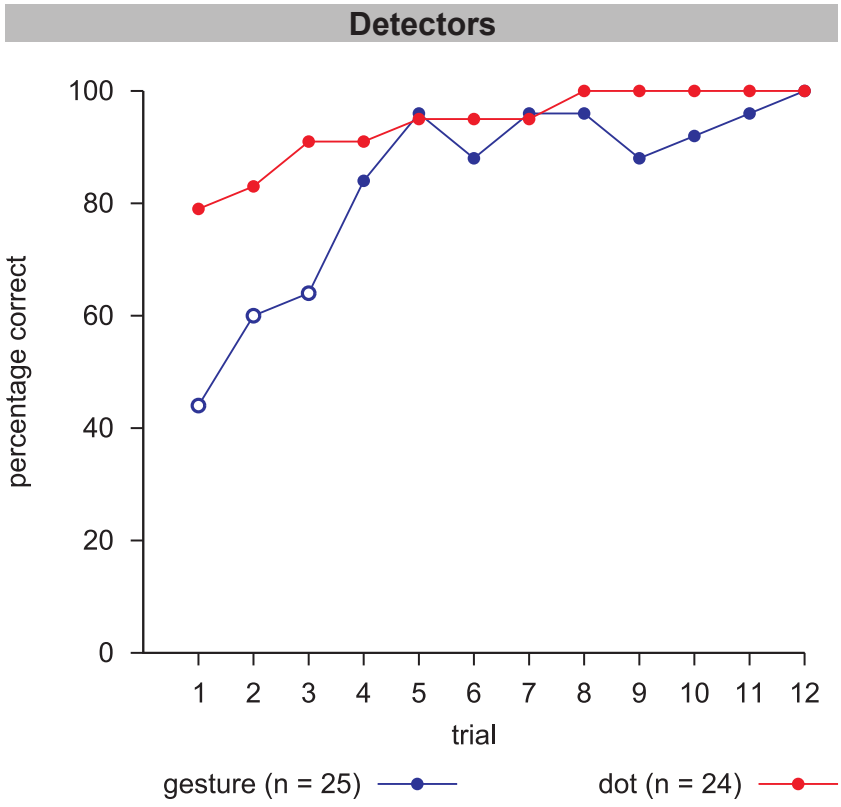

Figure 7.3: Learning curves for the subgroup of detectors. Filled data points indicate that the performance was significantly different from chance.

Table 7.3: Correct Responses and p-Values for the Binomial Tests within the Detector Subgroup

Trial	Gesture			Dot		
	N_{cor}	p-value	α_{adj}	N_{cor}	p-value	α_{adj}
1	11	.690	–	19	.007*	.050
2	15	.424	–	20	.002*	.046
3	16	.230[a]	.042	22	<.001*	.042
4	21	.001*	.038	22	<.001*	.038
5	24	<.001*	.021	23	<.001*	.033
6	22	<.001*	.033	23	<.001*	.029
7	24	<.001*	.017	23	<.001*	.025
8	24	<.001*	.013	24	<.001*	.021
9	22	<.001*	.029	24	<.001*	.017
10	23	<.001*	.025	24	<.001*	.012
11	24	<.001*	.008	24	<.001*	.008
12	25	<.001*	.004	24	<.001*	.004

Note. The gesture group contained 25 participants and the dot group 24 participants.
[a] $p > \alpha_{adj}$ leading to abortion of Benjamini-Hochberg procedure
*significant at false discovery rate of .05

Within the subgroup of the detectors, the pattern was approximately the same as for all participants, but with increased overall performance (see Figure 7.3). Again, the participants managed to improve with time: The gesture and dot group started with 44.0% and 79.2% in the first trial and ended both with 100.0% in the last trial. The gesture group started to be significantly above chance level with trial four and never fell back to chance. Notably, the participants of the dot condition were better than chance from the beginning on. The amount of correct responses and according p-values are presented in Table 7.3. Looking specifically at trial 1 and trial 3, one might suspect a difference between the conditions. However, while the according p-values were below the conventional alpha level, they failed to be significant after Bonferroni-Holm adjustments (trial 1: $X^2(1, N = 49) = 6.38, p = 0.012, \alpha_{adj} = .004$; trial 3: $X^2(1, N = 49) = 5.38, p = .020, \alpha_{adj} = .005$). Altogether there was never a

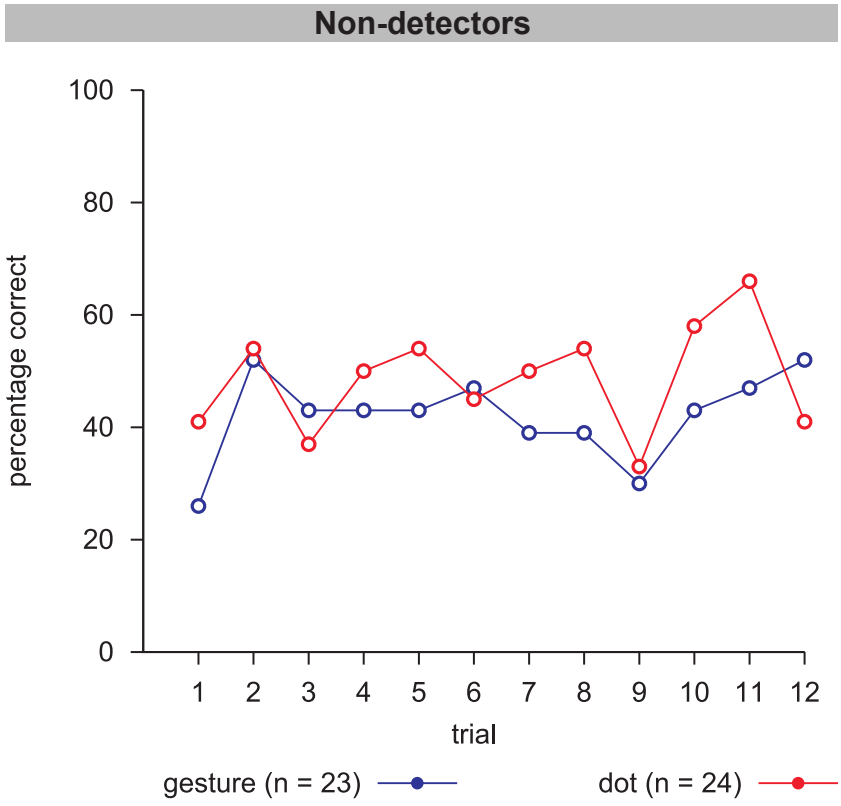

Figure 7.4: Learning curves for the subgroup of non-detectors. All data points unfilled, indicating a performance not significantly different from chance.

significant difference between the groups; the better performance of the dot group in trial 1 and 3 could be regarded as a tendency.

Within the subgroup of non-detectors the participants showed apparently a random performance (see Figure 7.4). None of the groups showed an accuracy rate significantly different from chance in any of the trials. Again, the groups did not differ significantly in any of the trials in their performance.

In the post-hoc questionnaire 30% of the non-detectors reported that they had applied a strategy in order to answer the questions. These strategies were written down in a non-standardized form by the recipients and can be roughly classified in the following way: Some participants chose a linguistic based strategy in order to identify the referent. For example,

Table 7.4: Strategies that were Applied by the Non-Detectors

	None	Linguistic	Content	Alternating	Sum
Frequency	33	6	6	2	47
Percentage	70%	13%	13%	4%	100%

they assumed that the referent of the critical response would be the one the interviewee talked about before. Other participants tried to interpret the content of the interviewee's statements regarding a specific interview topic very thoroughly. This way they built up hypotheses about the interviewee's preferences and picked the according response alternative. Two participants were in this regard not classifiable as they switched between different strategies until the end of the experiment (see Table 7.4).

To summarize the main results: First of all, it is important to note that roughly half of the participants did not detect the system how to answer the questions correctly. The accuracy of the non-detectors was generally at chance level. In contrast to this, the accuracy of the detectors was most of the times significantly above chance – for example, all participants of the gesture subgroup as well as all participants of the dot subgroup indicated the correct response in the last trial. A clear performance difference between the gesture group and the dot group was never found.

7.4 Discussion

Experiment 3 showed two major findings. First, spatial reference tracking cannot only be accomplished with abstract pointing. Second, engaging in spatial reference tracking might not be a general phenomenon.

7.4.1 Flexibility and high associability

The original research question of the present experiment was whether the pairing of a referent with an abstract pointing gesture is a selective association. This idea would have been supported, if participants had shown a preference to learn such associations as compared to the pairing of referents with artificial dots. This, however, was not the case. For example, the rate of detectors per group did not differ, i.e. regardless of stimulus type roughly half of the participants started to infer the referent based on spatial reference tracking. Also, the results within the group of detectors

do not support the idea of abstract pointing as a preferred stimulus. Both detector groups depicted a similar learning curve and acquired the associations to such a degree that they showed an impeccable accuracy rate at the end of the experiment.

On the contrary, if anything the data slightly favor the notion of an advantage for the artificial stimulus, because at least in the beginning there was a tendency for the dot participants to show a higher accuracy rate. In my opinion, this could be due to the fact that the dots constituted aside from the grey frame the only visual stimulation. Thus, their appearance was probably always striking and attention drawing, which made it possible to quickly establish a link between the dots and speech. In contrast, the gesture participants had to process a far more complex visual stimulation including other movements such as head tilting or changes in posture. Certainly, the gesturing itself was also easy to detect, but I assume that it was more demanding to isolate the gestures as a key element for the task compared to the dots. Altogether, the tendency for an earlier usage of the artificial dots is attributed to a higher salience of this stimulus type.

Despite the just mentioned trend for a difference, the results show in essence a null effect. While this indicates that abstract pointing is not preferred over artificial dots in spatial reference tracking, it does not necessarily contradict the idea that the association between abstract pointing and discourse referent is a selective one. Instead and based on the quick learning rate of the detectors in the gesture and in the dot group I assume that the discourse referents were selectively associated with both types of conditioned stimulus. In consequence, the data suggest that the cognitive system is flexible and accepts different kinds of spatial stimuli in order to engage in reference tracking. For experimental matters, this finding could be leveraged, as some future studies on spatial reference tracking could be carried out with artificial stimulus material that requires a drastically reduced amount of resources in preparation. Thinking about everyday conversations, it supports the idea that the recipient infers the referent not only from spatially consistent abstract pointing, but also from other gesture types such as iconics.

The current results stand in contrast to those from studies on other gesture types such as concrete pointing or beats, where dissociations between a prototypical and an atypical stimulus were found (Ariga & Watanabe, 2009; Holle et al., 2012). Since the present experiment does not only differ regarding the gesture type that was studied, but also other critical as-

pects like the dependent variable, there is, of course, the possibility that the mentioned contrast is caused by differences in experimental design. I propose, however, that the cognitive system's flexibility regarding spatial reference tracking can be explained when looking from a functional perspective on classical conditioning.

It has been repeatedly stated that classical conditioning usually works best, when there is an ecological relation between the conditioned and the unconditioned stimulus (Hollis, 1997; Domjan, 2005). Let's reconsider, for example, the study of Garcia and Koelling (1966). It appears safe to assume that there is a high ecological relation between the conditioned stimulus of a specific taste and the unconditioned stimulus of feeling sick. After all, in everyday life people get sick after eating or drinking something of a specific taste. Sickness is not, however, preceded by a specific audiovisual stimulation. As a result, rats learn the association between taste and sickness much better. In case of the present experiment, I speculate that the ecological relation is established through the dimension of *space*. When encountering a referent in real life, one of its most relevant attributes is its location. For example, if an infant wants to make a request at his mother, he will usually turn towards her location. If he wants to give something to her, the location is even essential. Obviously, spatial location is also an important attribute of abstract pointing and based on this ecological relation it is easily associated with discourse referents. However, spatial location is also the most distinctive attribute of the artifical dots. In other words, without intending it a kind of non-natural stimulus was created that perfectly captured the ecological relation between abstract pointing and discourse referents. Proposedly, this led in sum to the high learning rate within both detector groups indicating a high associability for each stimulus type.

7.4.2 Non-detectors

Unexpectedly, a major finding of the present experiment is that half of the participants did not engage in spatial reference tracking and showed a random performance. This is in contrast to the a-priori assumption that every participant would sooner or later utilize the gestures or dots to infer the referent. It is also surprising in light of the quick learning of the detector groups. Various reasons of different origin might account for this result.

Lack of grounding

The high rate of non-detectors might be driven by the experimental design. Naturally occurring communication is a joint activity, which means that there is a high degree of interaction between the interlocutors. A major purpose of the interaction is to *ground* the communicative acts; in other words, the discourse partners try to reach a state, where every partner has a sufficient understanding of what is being communicated (Clark & Wilkes-Gibbs, 1986; Clark, 1997). Consider, for example, the critical situation of the present experiment, where the interviewee is asked whom she believes to be the better actor and she replies with (1).

(1) Well, then this muscleman$_{[\text{Schwarzenegger}]}$, because at least [cut]

Let's assume for a moment that this is a natural conversation between the interviewee and one of the participants. At this moment in time, there would be plenty of means for both parties to ground the communicative act. If secure about the referent, the participant could produce a head nod. If insecure, he could ask, "Who do you mean?", he could make an irritated expression with his face, or he could stutter "Eh?". All this would signal to the interviewee, that there is a need for clarification. The interviewee herself could monitor the participant and if he signals problems or does not signal understanding, she could repair her statement by repeating it in a more explicit manner. In case the participant simply continues the conversation, the interviewee would also have further means to ground. If he continues, "Agreed, Stallone is way better", she could interrupt and put the matter straight. If he continues, "Agreed, Schwarzenegger is way better", she could utter "Yeah", which would work as an affirmation for the participant (for exemplary means of grounding see Sacks, Schegloff, & Jefferson, 1974; Goodwin, 1986; Clark & Krych, 2004). Using different terminology, one could say the participant would receive *feedback* about whether he is right or wrong in his referent assumption.

Coming back to the high rate of non-detectors in the present experiment, I speculate that it might be due to a lack of grounding. There was no communicative interaction between interviewee and participant and just in general there was no feedback that informed the participants about the correctness of their responses. In turn, it was not possible for the participants to evaluate whether their task strategy was successful or not. This could have led them to stay with whatever strategy they had applied in the first place and if this was not spatial reference tracking,

random performance would be the result. Alternatively, the participants could have started to switch constantly between various strategies or to refrain from using any strategy. Either way, the result would once again be a performance at chance level.

One might object that the establishing phase could have served as a "sandbox", where the participants had the opportunity to test their strategies and receive feedback about them. This is certainly true for the strategy of spatial reference tracking. If the participants suspected the location of gestures/dots to play a key role in inferring the referent, they could have monitored during the establishing phase whether their hypothesis of fixed associations between speech and spatial stimulus is correct. Every utterance of a referent accompanied by the spatial stimulus would have served as positive feedback and on average this would have happened six times in each trial. For other strategies, however, it is difficult to maintain the idea of the establishing phase as a sandbox. According to the follow-up interviews, a part of the participants tried to assess the interviewee's character during the establishing phase and picked the referent of the critical phase according to their subjective assumption of how the interviewee would respond. This means probably, that these participants did not regard the establishing phase as a testing arena, but as a time where they had to pay utmost attention to the interview's content in order to form the necessary beliefs about the interviewee's character. Another part of the participants pursued the strategy to pick the critical referent based on who the interviewee was referring to in the preceding utterance. It seems safe to assume that the participants did not think that this strategy would work throughout the whole conversation, but only in case of ambiguous speech. Since the establishing phase did not include verbal ambiguities, it did again not qualify as a sandbox.

In sum, I hypothesize that the rate of detectors should be considerably higher in a natural conversation that includes the element of grounding.

Lack of amplitude

At least in case of the gesture condition, another reason to miss the opportunity of spatial reference tracking could be rooted in formal aspects. As mentioned in the Introduction, pointing can vary in amplitude (see Section 3.2). You can make a flick with the index finger or you can lift the whole arm, extend the elbow and all fingers. All degrees in between are imaginable. As suggested by Enfield et al. (2007), amplitude could be

correlated to the communicational importance of the gesture: the greater the amplitude, the more important for the recipient to analyze the gesture. It is as in verbal speech, where crucial words can be stressed by the speaker.

While the according observations regarding gesture amplitude were made for concrete pointing, it is easily imaginable that the same goes for abstract pointing, i.e. that exceptionally relevant abstract pointing gestures are carried out with a greater amplitude. In contrast to that, the abstract pointing gestures of the current experiment were all conducted in roughly the same manner with a rather low amplitude, at least when comparing the gestures with sample images from other publications (Enfield et al., 2007; Stec & Huiskes, 2014). In my opinion, this might have lured some of the participants into not considering the gestures for the purpose of reference tracking.

In case of the dot condition, this line of reasoning is hardly applicable, as the dots were quite salient. However, a potential problem could have been that the dots were *always* salient. In verbal communication we do not stress every occurrence of "Schwarzenegger". If we did, it could be that the stressing loses its potential to trigger a plus of attention. In consequence, one might speculate that there might have been more detectors in the dot group, if the dot on the ambiguous term "this muscleman" had been more salient than the establishing dots and, thus, signaled a greater importance.

Interindividual differences

Despite the potential impact coming from lack of grounding or lack of amplitude half of the participants did nevertheless become detectors. This suggests that differences between participants are alternatively or in addition a contributor to the found split in task performance. Such differences could be located in cognitive skills or in personality traits.

Three cognitive capabilities appear comparatively important in comprehending spatial reference tracking. First, the capability to divide attention between two modalities in order to gather visual and auditory information simultaneously seems mandatory and its extent could have a direct influence on the comprehension of spatial reference tracking (Duncan, 1980; Driver & Spence, 1998). Second, recipients need to process the spatial relations adequately. In case of two distinct referents this task appears not too difficult, but when imagining a greater number of referents it be-

comes obvious that visuospatial cognition could be a key capability and at least for gesture production, it has been shown that greater spatial skills are correlated with a higher gesturing rate (Hostetter & Alibali, 2007). Third, working memory could be important. This hypothesis is mainly based on the fact that the comprehension of speech benefits significantly from good working memory (Daneman & Merikle, 1996; Caplan & Waters, 1999; Baddeley, 2012). In consequence, it could very well be that the comprehension of spatial reference tracking benefits as well; the only difference might be that it is not only verbal working memory, which is of importance, but also visuospatial working memory (cf. Wu & Coulson, 2014).

Regarding personality traits, specifically extraversion (McCrae & Costa, 1997) could foster the ability to detect spatial reference tracking. This assumption is based on two findings. First, extraverted people seem to gesture more than introverted people (Gifford, 1994; Hostetter & Potthoff, 2012). In other words, these people have a high proficiency in the production of co-speech gestures. Second, at least in case of sign language, it has been shown that proficiency can play an important role for the ability to understand signs (e.g. Morford & Carlson, 2011; Emmorey et al., 2012). Considering the relatedness of sign language and co-speech gestures, it seems a valid working hypothesis that the same goes for co-speech gestures, i.e. proficiency in producing gestures leads to better understanding of gestures. In consequence this would mean, that specifically extraverted people are readily able to detect spatial reference tracking.

Taken together, the proposition is that one or more of the skills divided attention, visuospatial cognition, working memory or the personality trait extraversion are crucial for the comprehension of spatial reference tracking. Accordingly, differences in these measures might have contributed to whether a participant became a detector or not.

The relevance of non-detectors

A posteriori it is not possible to decide, which one of the speculations about the existence of non-detectors might be correct. It is worthwile, however, to think about their potential implications. A first group of explanations attributes the existence of non-detectors to some external cause. This might be the lack of amplitude or the lack of grounding, others are probably imaginable. In extreme form, external explanations suggest that in case of naturally occurring communication there should

be only detectors. Spatial reference tracking should be carried out by everybody, if the circumstances are right.

A second group of explanations states that the non-detection of spatial reference tracking is caused by an aspect within the participant itself. For example, non-detectors might be introverted or might have trouble to divide their attention between two modalities. As noted, it is reasonable to assume that such interindividual differences did indeed play a crucial role in the present experiment, otherwise there should be no split into detectors and non-detectors at all. The intriguing question that arises from this reasoning is the following: Are there non-detectors in natural communication, i.e. people who do not engage in spatial reference tracking even when the circumstances are right?

This would be important to know for further research. If a researcher intended, for instance, to find out where abstract pointing is processed in the brain, it would be wise to include only such participants in an MRI experiment who actually process abstract pointing. Non-detectors would add to the noise and make it harder to detect potential effects. As a matter of fact, it cannot be ruled out that such an effect attenuation happened in the preceding ERP experiments. It would also explain some of the statistical variance in these experiments (described in an exemplar fashion when presenting the results on the face variable in Experiment 2, see Section 5.3.2). Thus, if the present finding of non-detectors is not a singular phenomenon, it would be desirable to have a means in order to filter them out before conducting an according experiment.

Aside from such methodological aspects, the reliable existence of non-detectors would pose further questions. The results so far on the comprehension of abstract pointing are quite diverse. We know that a considerable amount of recipients engages in spatial reference tracking, even when speech is always explicit (Gunter et al., 2015). We also know that a considerable amount of recipients does not engage in spatial reference tracking, even when speech does not suffice and they are asked for the referent (the present experiment). Studies on gesture production are also diverse in their findings as has been discussed in the Introduction (see Section 3.3.2). Depending on the specific conditions, the results reach from that the consistent usage of gesture space only sometimes takes place to the result that it almost always takes place (e.g. Gullberg, 2006; So et al., 2005). For the moment this suggests that there is not only flexibility regarding what kind of stimulus is used for spatial reference tracking, but also flexibility for whether spatial reference tracking is used at all. An

interesting question is then how the speaker and the recipient come to terms. How does a recipient notice that the location of gestures deserves attention? When can a speaker be sure that the recipient makes use of his consistent usage of gesture space?

These questions tap into interactional processes between speaker and recipient. Experiment 4 was intended at making a first step towards this research field.

8 Experiment 4 – a first step towards interaction

8.1 Introduction

In natural communication, the roles of recipient and producer are not stable. Instead the interlocutors interact and constantly switch their roles (cf. Clark, 1997). In the experiments so far this aspect was not explored, instead the focus was on the pure comprehension of abstract pointing, the participants were basically regarded as input systems that receive and process information. With the current experiment, however, the transition when recipient turns producer was put into the spotlight. A guiding question was whether such a newborn producer would also engage in spatial reference tracking after having observed abstract pointing by the discourse partner.

8.1.1 Interpersonal coordination

There is reason to assume that this could happen. In social interaction, people have a tendency to align with all kinds of behavior they observe in others. This includes the phenomenon of *behavioral mimicry*, a term that is most often used when people unwittingly imitate or repeat a particular motor behavior such as foot shaking or body posture (for a review, see Chartrand & Lakin, 2013). Mimicry has also been demonstrated for co-speech gestures (Tabensky, 2001; Kimbara, 2008). In an experiment by Holler and Wilkin (2011), for example, participants had to work jointly on a task. Crucially, they could either see each other or they could only hear each other. The dependent variable was the proportion of gestures that appeared similar to gestures, which the interlocutor had conducted before. The major result was that the proportion of such mimicked gestures was significantly higher in the face-to-face condition. This means, that the imitation of motor movements takes also place when semantic meaning is conveyed. Please note that the authors of this study focussed

on iconics and metaphorics. Pointing gestures in general were excluded from the analysis, as they look all very similar and, thus, it would have been difficult to tell mimicked pointings from non-mimicked pointings.

Importantly, interpersonal coordination has also been shown on a more complex level, i.e. when it is not only about observing motor behavior and repeating (almost) the same behavior within a short amount of time. Specifically in case of speech, for instance, it has been shown that people are prone to align with vocabulary or with linguistic structures (Garrod & Anderson, 1987; Pickering & Ferreira, 2008; Fusaroli et al., 2012). For example, when listening to a sentence like (1), this has an impact on how you describe a picture of a generous mother later on. In particular, the likelihood rises that you adopt the prepositional construction and say something like (2) instead of (3) (Bock, Dell, Chang, & Onishi, 2007).

(1) The team owner told an offensive joke to the columnist.

(2) A mother gives some ice cream to her boy.

(3) A mother gives her boy some ice cream.

The facts that people mimic co-speech gestures and that they align with their communicational behavior even on complex levels make it thinkable that a recipient might engage in spatial reference tracking via abstract pointing after observing it in the interactant. To illustrate the rationale of the current experiment, imagine for a moment that this takes really place. Let's say that somebody is in a conversation about dogs and cats. In the beginning he is in the role of the recipient and perceives, how the speaker is setting up a gesturing order via abstract pointing. This could be dogs on the left and cats on the right. Then, the speaker directs a question to the recipient, so it is suddenly his turn to be the producer and to state a response. If he is about to apply spatial reference tracking in this response, it is mandatory for him to access or reactivate the established gesturing order – otherwise he could end up conducting an abstract pointing towards the wrong location. Based on this reasoning, the present goal was to find out whether there is a traceable influence of the gesturing order, when the former recipient is in the process of producing a response. To this end, applying an *interference task* seemed most appropriate.

8.1.2 The Simon task

Interference tasks such as the Stroop or the Simon task have been widely and successfully used tools when it comes to studying the interaction of perception and action (Stroop, 1935; MacLeod, 1991; Simon & Small, 1969; Hommel, 2011). Despite crucial differences between all variants of interference tasks (Kornblum, Hasbroucq, & Osman, 1990; Hommel, 2011) most of them share a basic idea: The participants are presented with a stimulus set, which can either consist of a couple of items or of a single item that depicts various attributes. The stimulus set contains relevant and irrelevant information for the task at hand. The question of interest is whether the irrelevant information affects the task performance.

Most important for the present purpose is the Simon task. In a prototypical variant of this paradigm, participants are asked to press a left button when perceiving a high tone and a right button when perceiving a low tone. Put differently, the response dimension is of spatial nature (left/right) and the relevant stimulus dimension is pitch (high/low). Additionally, the stimulus set varies just like the responses on a spatial dimension, as the tones can be played via a headphone's left or right speaker. Although this attribute of stimulus location is irrelevant for the task, participants' performance is worse when it does not correspond with the response location. They need more time and make more errors, when a high tone demanding a left response is presented via the right speaker (Simon & Small, 1969). This interference effect is thought to originate from the overlap between the response dimension (left/right button to press) and the irrelevant stimulus dimension (left/right location of tone) (Kornblum et al., 1990).

An expansion of the pure Simon task, which is important for the current experiment, refers to the matter of timing. Usually and as in the study by Simon and Small (1969), irrelevant stimulus dimension, relevant stimulus dimension and response dimension are virtually simultaneously activated. The source of the tone is revealed together with its pitch and the participant is asked at this very moment to respond as fast as possible. In rare instances, however, Simon like effects have also been shown, when the irrelevant stimulus dimension is only available in memory (Tlauka & McKenna, 1998; Hommel, 2002; J. X. Zhang & Johnson, 2004). In an experiment by Pellicano, Vu, Proctor, Nicoletti, and Umiltà (2008), for instance, participants had to fulfill two tasks. First, they were asked to count the occurrences of two color stimuli with each of them only being presented

on one side of the screen, e.g. blue was always left and green always right. In the second task, the same stimuli were presented in the middle of the screen, i.e. they did not vary on a spatial dimension anymore, and participants had to indicate the color by pressing a left or right button. For one half of the participants, the mapping between the response buttons and the color stimuli matched the association of the stimulus locations and the color stimuli from the first task (e.g. left button for blue color). For the other half the mapping was flipped (e.g. left button for green color). The results showed at least for the first trials of the second task an interference effect. Participants of the flipped group depicted an increased reaction time and decreased accuracy. This means, although the irrelevant spatial information about the color stimuli (left/right) existed only in memory, it still influenced the response performance for the second task.

For the current experiment on abstract pointing, the Simon task was adapted in a similar way: Just like the above imagined recipient, who is in the middle of a conversation about pets, participants were presented with a dualistic interview topic and its according gesturing order (*dogs – left and cats – right*). Then, they were asked a dual-choice question. The response alternatives were the main referents of the topic, dogs and cats, and they were aligned in a left-right fashion on the screen. Responses were accordingly indicated with the left or right hand. The critical manipulation was to pit the irrelevant gesturing order of the bygone establishing phase against the presently relevant alignment of the response alternatives. For example, when being asked, "What animal does bark?", the correct response alternative *dogs* could either show up on the left side of the screen, which would be in accordance with the irrelevant gesturing order or on the right side, which would be in violation with the gesturing order. If the participants access the gesturing order when being in the process of preparing the response, they should be affected by the manipulation. In particular, an increased reaction time and/or decreased accuracy rate for the violation condition was expected.

The core idea of this experimental design is based on a pure Simon task such as the one by Simon and Small (1969). On side of the stimulus there is a relevant dimension, which is in the present case of referential nature (dogs/cats), and an irrelevant spatial dimension, which comes from the gesturing order (left/right). The latter overlaps with the response dimension (left/right). Similar to the study by Pellicano et al. (2008) the irrele-

vant stimulus dimension is only available in memory when the response is demanded.

Obviously, however, there are further modifications in the present design that are necessary to note. For example, the critical stimulus itself is not presented. What is presented instead is a question and participants have to infer the correct response alternative, i.e. the critical stimulus that varies on a relevant and an irrelevant dimension. Only then, a potential interference could be generated. This is opposed to most Simon task experiments, where tones or color stimuli are overtly presented, when participants have to give their response. Another modification refers to the fact that the stimulus-response mapping is not known beforehand in the current design, but revealed when the response is due. In contrast to this, a classic Simon task depicts a fixed mapping throughout the experiment (e.g. react with left response to high tone and vice versa). This modification is mainly caused by the fact that the interview topics and, hence, the referents change constantly in the present experiment. Accordingly, a fixed mapping of referents to response sides is not possible, but has to be updated with every topic. In a way, the current experimental design does not represent a single Simon experiment with many trials, but many Simon experiments, where each contains only one trial.

I assumed these modifications to be a source of additional variance in the data potentially making the detection of an interference effect more difficult. Other modifications, however, were meant to make interference detection easier. Based on the fact that a lack of grounding might have been partly responsible for the high amount of non-detectors in the preceding experiment, the general goal was to increase the feeling of interaction and naturalness for the participants. For example, feedback was included in the present experiment. Furthermore, the questions to the participants did not appear in written form on the screen, but they were asked by the interviewer. Finally, participants did not have to press rather artificial buttons, but they could respond with simple flicks of their hands, which resembled very much pointing gestures. More details about these means are presented in the following sections.

8.2 Methods

8.2.1 Participants

34 volunteers participated in Experiment 4. Two of them were discarded from further analyses due to problems during experimental presentation. The remaining 32 participants (half of them females) had a mean age of 25.1 years ($SD = 2.4$) and were all right handed with a mean laterality quotient (Oldfield, 1971) of 91.7 ($SD = 13.1$). None suffered from a known neurological or hearing impairment. Vision was normal or corrected to normal. The participants were reimbursed with 7 EUR per hour.

8.2.2 Experimental design

The participants watched 72 interview topics taken from the recordings for the first experiment. Crucially, the last question answer chunk – the one with the verbal amibiguity – was removed and instead the interviewer directed a dual choice question to the participants. The response alternatives were always the objects that the interviewee had established in gesture space. For example, in the topic *vodka – tequila* the response alternatives were also "vodka" and "tequila". One response alternative was presented on the left side of the screen and the other one on the right side. The dependent variable of interest concerned the order of the response alternatives: Either it was in accordance with the established gesturing order or in violation with it. For instance, from the participant's point of view the interviewee had established the referent *vodka* on the left side of gesture space and the referent *tequila* on the right side. In the accordance condition, the word "vodka" was also presented on the left side and "tequila" on the right side. In the violation condition, the order of the response alternatives was reversed.

In order to maximize an interference effect between the (potential) activation of the gesturing order and the selection of the correct response side, both processes were aligned: For the activation of the gesturing order, I hypothesized that it would coincide with the moment, where the participants inferred the correct response object. By approximation, this moment was close to the onset of that word of the question that provided enough information to infer the answer. For example, when asking "What do you rather drink in Moscow?", the crucial bit of information was provided by the word "Moscow", because with this word the participants were enabled to decide that "vodka" is the correct response object.

For the selection of the correct response side, I assumed that this moment would be close to the onset of the presentation of the response alternatives. Consequently, in order to align the activation of the gesturing order and the selection of the response side, two means were applied: The questions were in such a way constructed that the response object could only be inferred with the last word of the question and the response alternatives were always presented with the onset of this last word.

Two more variables were used for randomizing purposes. First, each interview topic existed in two versions – either the correct response to the participants was the left established or the right established object. Each participant watched only one version of a topic, between participants the versions of a particular topic were balanced and within a participant "left questions" and "right questions" were also balanced. Second, half of the participants watched the topics in the sequence 1 to 72, while the other half watched them in the sequence 37 to 72 and 1 to 36.

8.2.3 Stimulus material

In a first step, the questions directed to the participants were developed. All 102 topics originally recorded for the first experiment were considered during this process. For some of the topics the question to the participants referred back to the content of the particular video (e.g. "What is in Sabine's opinion a bit more 'disreputable'?" taken from the topic *tattoo – piercing*). For other topics, it was possible to answer the question based on world knowledge (e.g. "What animal does bark?" taken from the topic *dogs – cats*). As mentioned, all questions had in common that the correct response could only be inferred with the last word of the question.

In theory the questions could have been presented in text form on the screen after the establishing phase. As it was the goal, however, to give the participants the feeling to be part of the conversation, the questions were presented in video form, where the interviewer addresses specifically the participant. In order to combine these new sequences with the existing videos in a harmonic manner, great care was taken to create the same "look" as before. To this end, the shooting situation was reestablished in a detailed manner (location, light, recording devices, interviewer's clothes). Each question was recorded with a close-up shot of the interviewer in the same way: At the beginning, he looked at the position, where the interviewee had been seated in the original recordings. Then he turned his head towards the camera, looked directly into

it and uttered the question. Thereafter, he kept looking into the camera for several seconds. This was necessary, as the response alternatives were supposed to be presented beyond the offset of the question's last word. The intention of using a close-up – a distinct shot that was not used during the original recordings – and of the interviewer turning towards the camera was to make it clear to the participants that the question was addressed to them and not to the interviewee. An example can be seen in Figure 8.1.

The post production was once more carried out with Final Cut Pro: The critical phase of the original videos was removed and replaced by the question directed at the participants. Also, the blur mask over the interviewee's face was removed. Again, this was intended to increase the feeling of the participants to be part of a natural conversation.

In a last step, 72 topics were selected for experimental presentation (see Appendix C for a complete list) and six for the demo version.

The response alternatives were not part of the videos, instead they were displayed on the videos with the experimental software *Presentation*.

8.2.4 Response measurement

In contrast to most reaction time studies, where the participants have to press a button in order to give a response, used two 3-axis accelerometers (*STMicroelectronics LIS3L02AQ5*), which were attached to the back of the most distal phalanx of both index fingers (see Figure 8.2). This way, the necessary response movement looked very much like a pointing gesture: The participants sat with the hands resting on their lap and when they wanted to give a response, they simply had to lift the according index finger. When they thought that the response alternative on the left side of the screen was correct, they had to lift the left index and vice versa. The intention behind this procedure was again that the setting should resemble more a conversation (making a gesture) than an experiment (pressing a button).

The sensitivity was for both sensors and between participants identical. They were manually calibrated so that a usual finger lift reliably elicited a response in all participants. A potential downside of using acceleration sensors instead of button presses is that accidental and slight movements elicit a response, too, which can make it difficult to detect the true response of a participant. One approach to avoid this problem in the current experiment was that the sensors showed a refractory phase,

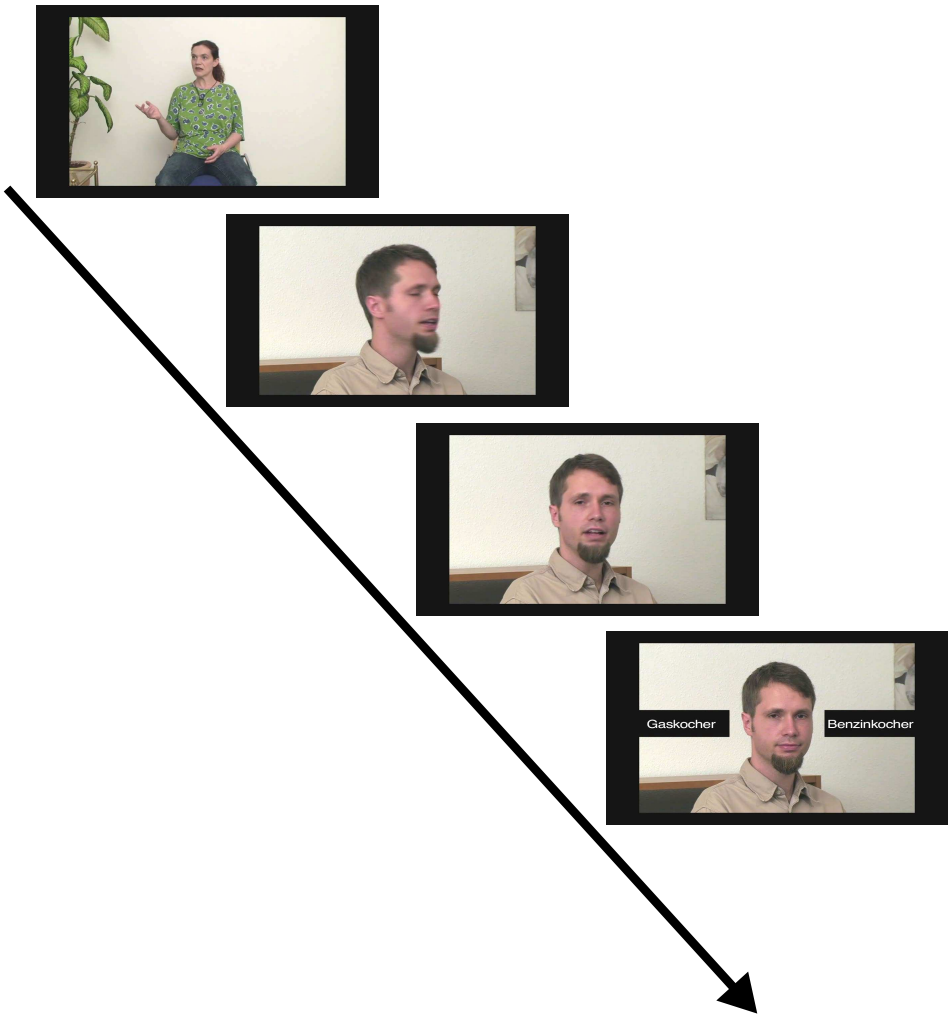

Figure 8.1: Going from upper left to lower right: After the last response of the interviewee (still frame 1), the interviewer turned to the participant (still frame 2) and asked a question (still frame 3). The response alternatives were visually presented on the screen at the onset of the question's last word, where they remained for 2000 ms (still frame 4).

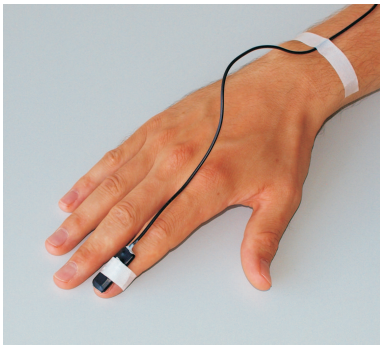

Figure 8.2: Accelerometer attached on the backside of an index finger.

i.e. they could not be activated again for a few seconds after a response had been elicited. The participants were informed about this behavior and were, therefore, motivated to keep still prior to giving the response. In addition, the participants were informed about the sensors' sensitivity, which was meant as a motivation to give a response with a single clear movement of the intended finger.

8.2.5 Procedure

The participants sat in a dimly lit room with a distance of approximately 120 cm to the computer screen. The accelerometers were attached to the tip of the index fingers with medical tape. In addition, the sensor's cable running to the measuring device was attached with clearance to the back of the hand so that the weight of the cable was not pulling at the sensor (Figure 8.2). The participants were orally instructed about how they should sit (hands on the lap), about how they were supposed to indicate a response and about the characteristics of the sensors (refractory phase and sensitivity). They were asked to give the responses as quickly and as accurate as possible. Subsequently, six demo trials were presented. If a participant had no further questions, the experimenter left the room and the participant was free to begin with the experiment.

A trial started with a progress bar indicating that the next video was prepared for presentation; it lasted for 1000 ms. Then the video with a particular topic started. At the end of the video the crucial question was posed to the participants and the response alternatives were presented with the onset of the last word of the question. The participants had

2000 ms to indicate a response, otherwise it was labelled as a miss. After the response, the participants were immediately provided with a feedback, which lasted for 2000 ms. In case of a correct response, the reaction time was presented in green font on a black background. In case of a miss or an incorrect response the according label was presented in red font.

8.2.6 Data analysis

Only correct responses entered the data analysis. Subsequently, data points with extreme reaction times were removed based on a box plot analysis (Tukey, 1977), i.e. all data points which were outside the data range $Q1 - 1.5 * IQR \leq x \leq Q3 + 1.5 * IQR$ were discarded. In the end, each participant provided on average 64.3 trials ($SD = 3.2$), which could be used for further analyses. The mean reaction times and the percentage of correct answers were calculated per subject and per condition. The achieved data were analyzed with separate two sided dependent t-tests.

8.3 Results

The participants indicated the correct response more quickly in the accordance condition than in the violation condition, which can be seen in Figure 8.3. The according mean reaction times were 906.0 ms ($SD = 76.0$) and 924.0 ms ($SD = 78.2$). The decrease in reaction time in the accordance condition was 18.1 ms ($SD = 46.7$). This difference was significant with $t(1, 31) = -2.2$ and $p = .04$.

As depicted in Figure 8.4 the participants gave also more correct responses in the accordance condition than in the violation condition. The respective mean percentages are 91.1% ($SD = 4.4$) compared to 87.4% ($SD = 7.1$). Though the difference of 3.6% ($SD = 7.5$) represents in absolute values only 1.3 trials, the difference was nevertheless significant ($t(1, 31) = 2.7$, $p = .01$).

8.4 Discussion

By applying a Simon-like task it was demonstrated that people reactivate the gesturing order they observed in the interlocutor when they become the producer during a conversation. Although a previously observed gesturing order was irrelevant for the participants' task, it had nevertheless

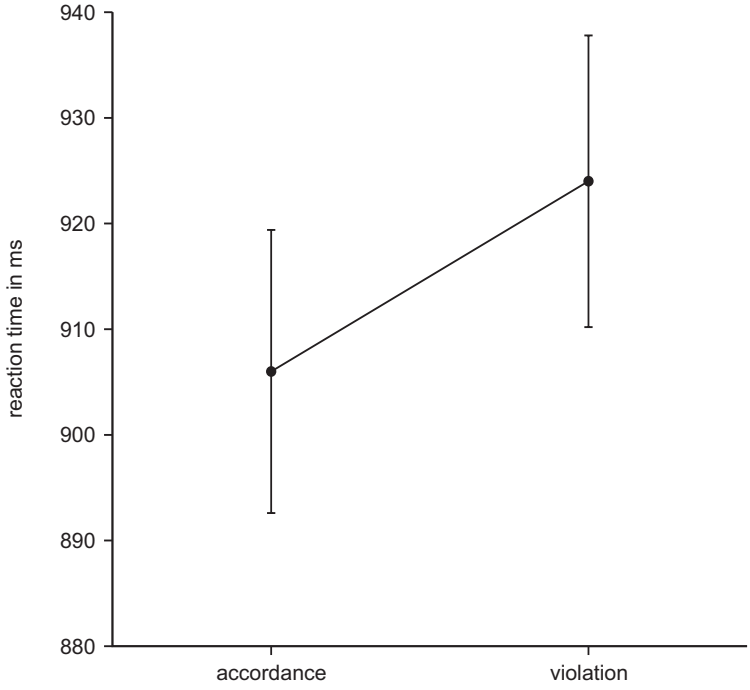

Figure 8.3: Mean reaction times depending on whether the response alternatives were presented in accordance or in violation with the established gesturing order. Error bars indicate the standard error of mean.

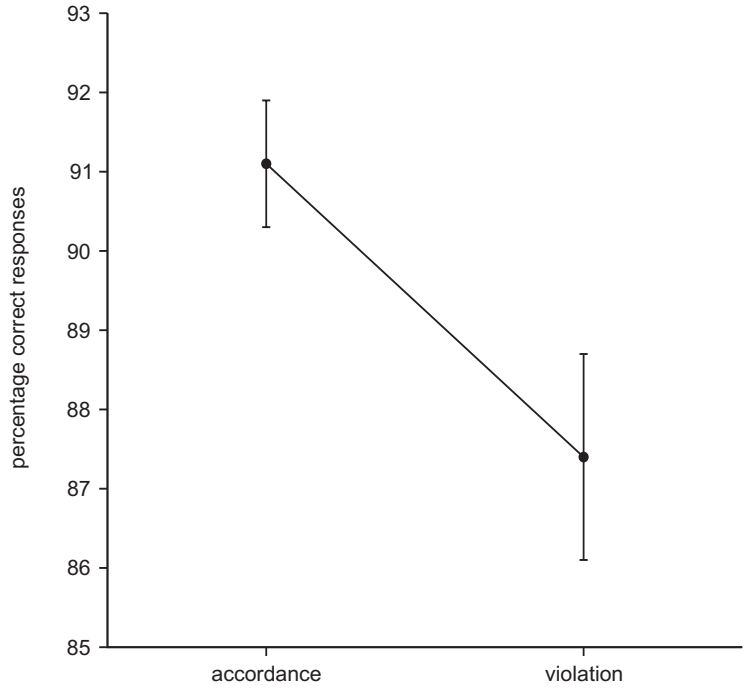

Figure 8.4: Mean percentage of correct responses depending on whether the response alternatives were presented in accordance or in violation with the established gesturing order. Error bars indicate the standard error of mean.

a significant influence on their response performance. This is a first indication that abstract pointing is also subject to interpersonal coordination.

Before turning into details about this result, it is worthwhile to consider that there was an interference effect at all in the present experiment, although the well established Simon task was modified in three regards. For one, participants had to infer the critical stimulus. For two, the stimulus-response mapping changed necessarily with every topic and was presented together with the stimulus set. For three, the irrelevant stimulus dimension existed only in memory.

That aside, two other characteristics of the present experiment render the demonstrated interference effect even more remarkable. The first one refers to the association strength between relevant and irrelevant stimulus dimension. In the present case, the association between referent and location was based on average on three events (two to four establishing gestures per referent). This is in harsh contrast to other studies showing an influence of memory based associations for Simon-like tasks. In the already presented study by Pellicano et al. (2008), for instance, there were 25 associative events per color stimulus (see also Tagliabue, Zorzi, Umiltà, & Bassignani, 2000; Vu, 2007). In addition to the results regarding the detector group in Experiment 3, this suggests that the associations between discourse referents and gesture locations are easily acquired.

The second characteristic is that there was never a reason for the participants of the current experiment to consider the irrelevant stimulus dimension, i.e. the spatial order of the referents. To elaborate this aspect, a comparison with the study by Pellicano et al. (2008) is once again helpful: Here, the associations between the color stimuli and the according locations were achieved by letting the participants count the appearance of the stimuli. As nicely observed by the authors, participants often applied spatial strategies for this task. For example, they visualized numbers at the locations of the color stimuli and increased this number with every appearance. The bottom line is that the irrelevant spatial dimension of the stimuli gained this way a task induced relevance. In contrast, there was no such relevance in the present experiment. Participants did not have to pay attention to gestures in general during the establishing phase. Instead they probably tried to memorize the content of the video in order to be prepared for the subsequent task and it is hard to imagine a consciously chosen, spatial strategy that could have helped them in this endeavor. Furthermore, in contrast to Experiment 1 till 3 the gestures never had the potential to disambiguate speech and there was never a po-

tentially attention raising conflict between speech and gesture. Of course, it is still imaginable that the unusual high amount of abstract pointing by the interviewee gave the gestures and their locations relevance. It appears likely, however, that this effect should have worn off after a couple of trials, where gesturing did not play a helpful role. In sum, I propose that the associations between referents and locations were learned in a spontaneous and incidental manner. If so, this would be further evidence that the recipient shows an affinity to take abstract pointing into account – even when there is no incentive to do so.

8.4.1 Origin of interference

A classic question when using an interference task is the one for the origin of the interference. For a pure Simon task such as the one by Simon and Small (1969), the answer is easy, as it is clear that the interference is caused by the overlap between the irrelevant stimulus dimension (tone location) and the response dimension (button location). For instance, the right location of a tone might prime a right response even though its pitch is high and, thus, requiring a left response (Kornblum et al., 1990; Hommel, 2011). This mechanism could also explain the present interference effect. For example, when participants wanted to indicate *dogs* as the response, the irrelevant gesturing order *dogs – left and cats – right* might have primed a left hand action, which competed with the right hand action that was primed by the location of the word "dogs" on the screen.

In the present Simon-like task, however, at least two other origins are thinkable. First, the interference could be due to an overlap between the irrelevant and the relevant stimulus dimension. In particular, it is imaginable that the gesturing order virtually placed *cats* on the right side of the screen, which could have impeded the encoding of the word "dogs" at this location. In this regard, the present experiment resembles another well known interference paradigm, i.e. the Stroop task, where further research was necessary to pinpoint the interference source, which is in this case a mixture of stimulus-response overlap and stimulus-stimulus overlap (H. Zhang & Kornblum, 1998; de Houwer, 2003).

Second, the current result pattern could be rooted in the allocation of attention. Then, the experimental design could be regarded as a kind of Posner cueing task. In the name giving study, participants had to react to peripherally presented targets (Posner, 1980). Prior to this, a cue was presented indicating the location of the target. Invalid cues, i.e.

cues indicating the opposite location, led to slower reactions to the target. Similarly, when wanting to indicate the response *dogs*, the participants of the current experiment could have used the information of the gesturing order as a cue in order to shift their attention to the left side of the screen. Just as observable in the current data, this should have led to a better performance in the accordance and a worse performance in the violation condition.

In sum, future studies will be necessary to tell *what* exactly caused the current interference effect, stimulus-response overlap, stimulus-stimulus overlap or a systematically varying allocation of attention. Regardless of the precise mechanism, however, I suggest that the effect is caused by the irrelevant gesturing order. Moreover, I am quite certain about *when* the crucial access to the gesturing order took place.

8.4.2 Time of reactivation

On first sight, one might suspect that the participants activated the gesturing order for the last time, when they observed the final abstract pointing gesture of the establishing phase. One might further speculate that this activation lasted long enough to influence the response performance at the end of the trial. While theoretically possible, this scenario seems unlikely due to the large time gap between the stroke of the last observable gesture and the presentation of the response alternatives. Specifically, there were on average 9.4 s ($SD = 5.9$ s) between both points in time. Though I am not aware of existing literature about temporal asynchronies of this magnitude for interference tasks such as the Simon or the Stroop task, it has repeatedly been shown that much smaller offsets in the range of 500 ms decrease an otherwise detectable interference effect in a considerable manner (Dyer, 1971; Glaser & Glaser, 1982; Hommel, 1993) and offsets in the range of 3 s can even eliminate it (Lu & Proctor, 2001). A similar reasoning holds for the case that attention allocation is the crucial mechanism. Since the validity of the cues was at chance level in the present experiment, the proposed attention shifts are presumably of exogenous nature, i.e. driven by an involuntary system (Giordano, McElree, & Carrasco, 2009; Carrasco, 2011). A crucial aspect about the exogenous deployment of attention is that it is not known to last longer than a couple of hundred milliseconds (Müller & Rabbitt, 1989; Carrasco, 2011). Thus, there is strong support for the idea that participants reactivated the irrel-

evant gesturing order when they inferred the correct response alternative and prepared their response.

This temporal aspect is the core finding of the present experiment. A recipient is in a conversation and amongst all the other information, which is mostly of verbal nature, he also perceives abstract pointing gestures. The previous experiments had already shown that he can build up associations between speech referents and gesture locations and that he can use these associations to infer the referent. The present experiment shows that the associations are kept in memory not only for as long as the interactant speaks, but long enough so that they can get reactivated when the recipient becomes a producer, i.e. when being in the process of preparing a response regarding one of the established referents. This finding is no evidence that the former recipient will show a tendency to actually engage in abstract pointing, but it is evidence that the according foundation is laid. The newborn producer is theoretically capable of utilizing the gesturing order for his own needs and goals.

Part III

General Discussion

9 General Discussion

9.1 Summary

When the work on this dissertation started, few things about abstract pointing in particular and spatial reference tracking in general were clear. It was known that the recipient notices, when abstract pointing is not used in a spatially conistent manner (Gunter et al., 2015). It was also known that the producer can engage in the consistent usage of gesture space when forbidden to talk (So et al., 2005). Other questions were still open. For instance, it was not clear whether the recipient can take advantage of abstract pointing by inferring the discourse referent. The aim of the present thesis was to answer this and other questions about abstract pointing in order to create a basis for further research. The following paragraphs are intended to recollect the major results.

All experiments were based on the same stimulus material, i.e. an interview in video format, which was supposed to look as natural as possible while allowing for as much control as necessary in order to run experiments with systematic and focussed manipulations of the independent variable. While the establishing phases remained the same for all experiments, they differed regarding the subsequent critical phase, where the experimental manipulation took place. In Experiment 1, the critical phase presented the participants with a verbal statement that was initially ambiguous and only later on explicit. An example is given in (1), which represents the interviewee's response to the question for her preferred author, Goethe or Shakespeare.

(1) Then this classic$_{[\text{Goethe/Shakespeare/—}]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.

Subscript indicates the gesture information of the congruent, incongruent and baseline condition. ERPs taken at the verbal disambiguation "Goethe's" showed a more positive P600 for the incongruent condition suggesting that the cognitive system required more processing resources in order to deal with referent conflict. Hence, the data seemed to support

the idea that the referent can be inferred from abstract pointing. Interestingly, there was no difference between the congruent and the baseline condition indicating that abstract pointing had no beneficial potential. This seemed paradox, as it would mean that the recipient takes abstract pointing into account, although it can only be of disadvantage.

With Experiment 2 this issue was further explored. Here, abstract pointing was a reliable cue, as the incongruent condition was removed from the experimental design. Now, the ERPs taken at "Goethe's" showed a more negative N400 and a more positive P600 for the baseline condition as compared to the congruent condition. This suggests that participants had to initiate a lexical access regarding the referent and that they had to update the MRC with the referent information – two processes that apparently had already been carried out in the congruent condition with the help of the gesture information. Thus, when abstract pointing is reliable, it can make language comprehension easier.

The comparison of both experiments suggests that the participants of Experiment 1 did at least execute a monitoring process regarding abstract pointing in order to update its reliability status. Without such a monitoring no change in how the gestures were processed should have happened. It was suggested that the monitoring itself could be based on the occurrences of incongruent gesturing.

Regardless of how exactly abstract pointing was processed in Experiment 1, it remains that it was continuously processed, although it repeatedly failed to be informative about the referent. The reason for this affinity might be that the cognitive system expects to profit from abstract pointing. Aside from disambiguation when speech lacks clarity, I speculated that abstract pointing might facilitate language comprehension even when its information is completely redundant with speech. This would put it close to the phenomenon of code-blending, where people knowing a sign and a spoken language simultaneously sign and utter equivalent semantic content. Importantly, there is first evidence indicating that code-blends are beneficial for the comprehension process (Emmorey et al., 2012) and the same might be true for spatial reference tracking via abstract pointing.

By conducting Experiment 3, the spotlight was moved from abstract pointing's effect on the discourse to its appearance. In particular, the research question was whether spatial reference tracking is also possible with artificial stimuli that are not encountered in natural communication. The short answer is yes. It did not matter whether participants were

presented with abstract pointing as part of a natural appearing video or with colored dots on a black screen. Their accuracy in identifying the correct discourse referent in verbally ambiguous situations was in essence the same. Hence, it appears that the recipient is quite flexible regarding what kind of stimuli are accepted for spatial reference tracking.

The unexpected finding of Experiment 3 was that regardless of stimulus type half of the participants did not engage in spatial reference tracking, i.e. they were not able to tell the momentary discourse referent based on the spatial cue. Potential explanations for this result pattern were discussed. For example, it might be due to interindividual differences, e.g. the ability to divide attention between the auditory (speech) and the visual (gesture) modality. It might also be due to specifics of the experimental design such as the fact that the participants were not grounded in form of feedback about their referent hypotheses.

Interactional aspects regarding abstract pointing became a major interest for Experiment 4, which was set up to see whether the perceived gesturing is to such a degree internalized that a recipient is still influenced by it when he becomes producer during a conversational process. This is indeed the case. Although the gesturing order was never of relevance for the participants, they reactivated the spatial information about the referents, when they had to indicate a response including one of the referents. This was demonstrated with a Simon-like task where the spatial gesturing order was pitted against the spatial alignment of response alternatives on the screen. When both spatial orders did not correspond, the responses of the participants were slower and more error-prone compared to the control condition.

9.2 Outlook

Since the present studies represent the first systematic approach to explore abstract pointing, the just summarized results represent but a mere fraction of what could be known about this gesture type. Before turning to potential future research projects, please allow for two rather general remarks. First, it is essential to remember that all findings of this dissertation are based on the study of *abstract pointing*. Necessarily this puts *spatial reference tracking* into the focus and while I switch often between those terms, it is important to realize that they are not interchangeable. Due to its form abstract pointing seems to be *the* gesture type being suited

for spatial reference tracking. It bears virtually no other feature than its location. This is different, however, for all other co-speech gesture types such as iconics. They depict more semantic features and, thus, more information that has to be processed. In consequence, the present findings on abstract pointing cannot be generalized and it would be a worthwhile endeavor to find out whether they also hold for spatial reference tracking that is conducted with iconics, for example. The results of Experiment 3, which stress the flexibility of the cognitive system, hint at least to the fact that this might be the case.

Second, it is not possible to just generalize onto abstract pointing what is known about iconics. This is mainly due to the fact that abstract pointing works in a quite different manner by requiring an establishing phase. Hence, theoretically there is a massive amount of topics that could be approached regarding abstract pointing. The following sections provide a subjective selection of some of these topics, but also topics that are uniquely interesting for the case of abstract pointing.

9.2.1 The bare necessities

It would certainly be interesting to further study the technical aspects of abstract pointing. For instance, while the present experiments demonstrate a high associability between abstract pointing and referent information (cf. Rescorla, 1988), the necessary amount of establishing gestures is not clear. With slightly modified designs it would be possible to further narrow down this number. For example, one could carry out an experiment like the pilot study with the only difference that there are just enough trials for data analysis, which depict solely one establishing gesture. If the results showed nevertheless the N400-P600 pattern for the mismatch condition of these trials, this would turn abstract pointing into an example for one-trial learning, which has mostly been shown for cases of high evolutionary relevance such as rapid aversion of toxins (Garcia, Kimeldorf, & Koelling, 1955; Ackroff, Dym, Yiin, & Sclafani, 2009; Dunlap & Stephens, 2014).

Other technical aspects that appear worthwhile to study refer to abstract pointing's amplitude, the body part it is accomplished with or the mechanism of *reliability updating*. Is there a necessary minimum amplitude? Can abstract pointing be conducted with head nods as it is the case for concrete pointing? Can a producer show an inconsistent usage of gesture space for a long time, but when he starts using it consistently,

the recipient would nevertheless quickly pick up on it? The basic interest beneath these questions is always to get an impression of the criteria that abstract pointing has to meet in order to be utilized for reference tracking. Obviously, interactions between different parameters seem likely. For instance, based on the knowledge about classic conditioning, one might speculate that an abstract pointing of big amplitude can trigger spatial reference tracking immediately, while small-amplitude gesturing requires maybe a couple of establishing gestures (cf. Rescorla & Wagner, 1972).

Finally, it could be that not only the gesture but also the recipient has to fulfill some requirements so that abstract pointing is successfully applied in spatial reference tracking. As discussed for Experiment 3 this could concern cognitive abilities or personality traits. In order to explore this issue, quasi-experiments could be carried out, where participants are measured on a selected scale such as *divided attention* and assigned to different groups according to whether they score high or low. When overtly or discretely confronted with the task to make use of abstract pointing, significant performance differences between the groups would suggest that the selected skill or trait is of relevance for the ability to process abstract pointing. For instance, when rerunning Experiment 3 one might imagine that there is a higher percentage of detectors within the group, who shows a good ability to divide attention. Research along this line would be insofar invaluable, as it seems absolutely necessary to be able to identify detectors and non-detectors for future experiments. An analogy with research on speech comprehension might illustrate this further: Who would run an experiment on the processing of the German language, when it is not sure how many of the participants actually understand German? Furthermore, the issue of non-detectors could be relevant for other gesture types, too. If the ability to divide attention between the visual gesture and the auditory speech stream proves to be crucial for abstract pointing processing, the natural follow-up question would be whether this is also true for other gesture types such as iconics. After all, the latter gesture type requires as well to divide attention between modalities.

9.2.2 Interaction revisited

Regarding interactional aspects this dissertation leads to two follow-up questions. The first one is based on the results from Experiment 4, where it was demonstrated that recipients are still affected by the gesturing order, when they become producer. As noted, however, it is not clear

whether they also have a tendency to actually engage in abstract pointing. Due to the fact that abstract pointing is conducted far less frequently than other gesture types like iconics (McNeill, 1992), it appears difficult to further explore this issue by analysing overt behavior of participants, simply because it would be necessary to gather an uneconomically huge amount of data. In contrast the lateralized readiness potential (LRP) might prove to be a suitable method. It is thought to reflect the preferential preparation of motor activity regarding a specific side of the body such as s moving just the left or just the right hand. Importantly, the LRP is sensitive enough to pick up preparational processes that never lead to overt behavior (Eimer & Coles, 2003). Thus, it might serve as an indicator for abstract pointing preparation without actual gesture execution. A potential experiment could be very similar to Experiment 4. Particularly, participants would again become producers at the end of a trial by asking them a question, e.g. "What animals do bark?" for the topic *dogs and cats*. This time, however, no response alternatives would be presented, instead participants would have to utter the response. Based on the established gesturing order *dogs – left and cats – right*, the hypothesis is that the LRP should indicate a preparation of left hand movement. If so, this would suggest that former recipients of abstract pointing indeed have a tendency to continue on abstract pointing.

The second interactional question resulting from this dissertation is the one for how people coordinate. Sometimes the producer conducts gestures in a spatially consistent manner, sometimes not. Sometimes the recipient processes abstract pointing in such a way that it can be beneficial, sometimes not. It would be quite inefficient, if both parties would not somehow come to terms on whether they presently engage in spatial reference tracking. A parsimonious postulation is based on the results of Experiment 1 and 2. Perhaps the producer gives the initial impulse by starting the consistent usage at some point in time. As noted, the recipient seems to monitor at least the reliability of abstract pointing. Hence, when the producer points consistently, the reliability index should rise and when reaching a certain threshold the recipient would start to fully process abstract pointing. In theory, nothing more would be necessary for successful alignment.

In light of the results from Experiment 3, however, this theory would suggest that half of human mankind is not able to make use of abstract pointing, because in this study there were roughly 50% non-detectors despite the fact that all gestures were conducted in a spatially consistent

manner. Certainly, this might be the case. Alternatively, it could be that the above postulation is too parsimonious and that there are other mechanisms between producer and recipient that assure alignment. I discussed the possibility that one of them is grounding and, indeed, it is easily imaginable that this basic principle of verbal communication applies for gestures as well. A first way to test this would be to simply rerun Experiment 3 with one slight alteration: giving feedback at the end of each trial. If grounding is essential for coordination, the ratio of non-detectors should decrease significantly due to this manipulation.

Even after such an experiment, the exploration of coordination in abstract pointing would be far from over. For instance, it has been shown that eye gaze is an important cue in the usage of iconics (Senju & Johnson, 2009; Holler et al., 2015) and this might go for abstract pointing as well.

9.2.3 Abstract pointing's value revisited

Especially the differences between Experiment 1 and 2 nurtured considerations about the communicative value of abstract pointing. An intriguing theory is that abstract pointing might be quite similar to the phenomenon of code-blending that is found in bilinguals who know a spoken and a sign language (Bishop, 2010). Interestingly, it has been suggested that bimodal code-blends facilitate comprehension when compared to unimodal presentations (Emmorey et al., 2012). On a broader level, code-blends could therefore be regarded as a naturally occurring variant of the *redundant signals effect*, in which it has been shown that participants show a better performance when two stimuli, which demand the same response, are presented at the same time via different modalities as compared to a situation, where only one stimulus is presented (Raab, 1962; Miller, 1986; Schröger & Widmann, 1998). A code-blend works presumably in the same way, as word and sign require the same response, i.e. to comprehend the (redundant) semantic content. The same might be true for abstract pointing, where pointing gesture and speech require to comprehend the semantic content of an item like *Goethe*. Just as in code-blends or the redundant signals effect, this could lead to an accelerated response, specifically to a faster comprehension of the word.

Remarkably, there are further similarities between abstract pointing and sign language. As noted earlier, co-speech gestures differ usually profoundly from sign language (see Section 3.1.2). For example, sign languages have a lexicon and they have a standard of well-formedness, while

co-speech gestures do not. Abstract pointing disarranges this convenient distinction at least partly. As soon as speech gesture associations are built, a pointing to the left can indeed be looked up in a virtual lexicon and the recipient would find the entry *Goethe*. Furthermore, after the establishing phase there is insofar a standard of well-formedness, as the producer has to point to the left (and not to the right) when marking Goethe as the referent. While this represents no evidence that abstract pointing facilitates through code-blending, it shows that abstract pointing and signs are not as different as they might appear on first sight.

If the code-blending hypothesis is true, this could shed light on a couple of issues. For instance, it could explain the early onset of the N400 modulation in the initial ERP study on abstract pointing (Gunter et al., 2015), as the code-blend potentially allowed for an earlier retrieval of semantic information. It could also explain, why the recipient shows a stubborn adherence to abstract pointing, even when it is notoriously unreliable as in Experiment 1. Maybe this happens, because abstract pointing is even beneficial in situations of speech gesture redundancy. In a similar manner, the code-blend hypothesis could explain, why recipients do not show a preference for consistent usage of gesture space in times of verbal underspecification (So et al., 2009). While such a situation seems on first sight to be the prime example for facilitation by spatial reference tracking, it might be just one of many beneficial situations.

For the moment, it remains to be seen whether abstract pointing can be compared with signs in code-blends. This hypothesis would be supported, if typical redundant signal effects could be demonstrated with this gesture type. Ideally, an according experiment should comprise two groups of participants: a bimodal group, where participants know a spoken and a sign language, and a unimodal group, where participants have no knowledge of a sign language. When bimodals are presented with code-blends and with speech alone in a semantic decision task, this should lead to a similar facilitating effect of code-blends as the one demonstrated by Emmorey et al. (2012). Crucially, if unimodals show a comparable effect of facilitation when speech accompanied by abstract pointing is put in contrast with speech only, this would support the code-blend hypothesis for abstract pointing. Additionally, the unimodal group should also be presented with pairings of abstract pointing and speech, where the preceding establishing gestures are not presented to the participants. This way, it could be assured that a potential beneficial effect of abstract pointing is not caused by random visual stimulation.

If such a behavioral experiment shows significant effects, the code-blend idea could be further explored using ERPs. Basically, the same design as above could be used and one of the core questions would be the one for the affected ERP components. For instance, when participants are presented with a conversation including spatial reference tracking, how does the processing of a discourse referent differ depending on whether it is accompanied by (established) abstract pointing? I would not expect that the P600 component shows any deviations, because the according referent has to be integrated in the MRC either way. The N400 component, however, might be less pronounced in the condition with abstract pointing, as lexical access could be eased when being carried out via two modalities. Again, including a condition with abstract pointing, where the establishing gesture speech pairings are not visible to the participants, seems mandatory in order to rule out that potential effects are solely due to visual stimulation. Finally, it would be once more interesting to compare the result with the difference between a code-blend and a speech-only condition in bimodals to gain further insight in the relatedness of signs in code-blends and abstract pointing in spatial reference tracking.

9.2.4 The neural underpinning

Since a couple of years the functional neuroanatomy behind gesture comprehension has started to draw more and more scientific attention (e.g. Skipper, Goldin-Meadow, Nusbaum, & Small, 2007; Willems, Ozyürek, & Hagoort, 2007; Holle, Gunter, Rüschemeyer, Hennenlotter, & Iacoboni, 2008). The findings are diverse and differ, for example, depending on the studied gesture type or the context. In recent meta-analyses, however, a consensus emerges on brain areas that usually show an increased activation for gesture comprehension (Andric & Small, 2012; Yang, Andric, & Mathew, 2015). Accordingly, areas like the premotor cortices, the inferior and superior parietal lobules, and the posterior part of the left superior temporal sulcus are primarily involved in recognizing gestures as gestures. This includes to identify gestures as biological motions that are to be separated from otherwise movement and, more specifically, to identify them as movements originating from the hands. Then there are areas like the posterior part of the medial temporal gyrus and left inferior frontal gyrus that are supposed to be involved in the semantic processing of the gestures.

Of course, it would be interesting to know how the comprehension of abstract pointing fits into this pattern. Perhaps this would help to further narrow down the brain regions that are essential for gesture processing in general. That aside, especially comparisons with other gesture types might promote the understanding about the neural underpinning of gesture comprehension. For instance, it has been suggested that gestures like emblems, i.e. gestures with a predefined meaning, show a stronger activation specifically in the posterior part of left superior temporal sulcus, which is involved in the comprehension of written and spoken language (Yang et al., 2015). Since abstract pointing presumably has a predefined meaning after the establishing phase, one might speculate that its comprehension should also show a stronger activation in this area when being compared with the comprehension of a gesture type like iconics.

Coming back to the code-blend idea, functional imaging may also be helpful to further explore this particular issue. Weisberg, McCullough, and Emmorey (2015) found in an fMRI study that code-blends lead to a decreased activation in auditory association areas of the cortex as compared to a speech-only condition. This suggests that the facilitatory effect of code-blends may at least partly be due to a more efficient processing on a sensory level. If abstract pointing is indeed processed as part of a code-blend, a similar reduction in activation should be observable.

9.3 Conclusion

In this dissertation, a rather neglected type of co-speech gesture and its potential mechanism to impact communication were studied. My curiosity about abstract pointing arose with the discourse that was referred to as the *Chicago Conversation* (McNeill, 2003). Could it be that discourse partners set up referents in empty space with simple pointing gestures and that the referent information coming from these locations has a considerable influence on the course of the conversation? In retrospective it is obviously difficult to tell what happened in this particular conversation. However, after this thesis project I am positive that it could have happened this way. Recipients are capable of inferring the referent from the location of abstract pointing and the communicational process can benefit from this inference. While it is not clear under what exact circumstances people engage in spatial reference tracking, they show flexibility regarding the type of spatial stimulus they accept and they can quickly

learn the associations between spatial stimulus and speech. Furthermore, people observing a gesturing order show a tendency to align with this order beyond the utterance of the interactant, i.e. when it is their turn to respond.

Part IV

Bibliography and appendix

Bibliography

Ackroff, K., Dym, C., Yiin, Y.-M., & Sclafani, A. (2009). Rapid acquisition of conditioned flavor preferences in rats. *Physiology and Behavior*, *97*(3-4), 406–413. doi: 10.1016/j.physbeh.2009.03.014

Alamillo, A. R., Colletta, J.-M., & Kunene, R. N. (2010). Reference-tracking in gesture and speech: a developmental study on french narratives. *Rivista di Psicolinguistica Applicata*, *X*(3), 75–95.

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*(2), 169–188. doi: 10.1006/jmla.2000.2752

Alvarez, R. P., Holcomb, P. J., & Grainger, J. (2003). Accessing word meaning in two languages: an event-related brain potential study of beginning bilinguals. *Brain and Language*, *87*(2), 290–304. doi: 10.1016/S0093-934X(03)00108-1

American Clinical Neurophysiology Society. (2006). Guideline 5: guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *23*(2), 107–110.

Andric, M., & Small, S. L. (2012). Gesture's neural language. *Frontiers in Psychology*, *3*, 1–12. doi: 10.3389/fpsyg.2012.00099

Ariga, A., & Watanabe, K. (2009). What is special about the index finger? the index finger advantage in manipulating reflexive attentional shift. *Japanese Psychological Research*, *51*(4), 258–265. doi: 10.1111/j.1468-5884.2009.00408.x

Aslin, R., Saffran, J., & Newport, E. (1999). Statistical learning in linguistic and non-linguistic domains. In M. B. (Ed.), *The emergence of language* (pp. 359–380). Mahwah, NJ: Lawrence Erlbaum Associates.

Astheimer, L. B., & Sanders, L. D. (2009). Listeners modulate temporally selective attention during natural speech processing. *Biological Psychology*, *80*(1), 23–34. doi: 10.1016/j.biopsycho.2008.01.015

Astheimer, L. B., & Sanders, L. D. (2011). Predictability affects early perceptual processing of word onsets in contin-

uous speech. *Neuropsychologia*, *49*(12), 3512–3516. doi: 10.1016/j.neuropsychologia.2011.08.014

Austin, G. (1966). *Chironomia: or, a treatise on rhethorical delivery*. Carbondale, IL: Southern Illinois University Press.

Bacon, F. (2005). *The advancement of learning*. Sioux Falls, SD: NuVision Publications.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. doi: 10.1146/annurev-psych-120710-100422

Bates, E., Camaioni, L., & Volterra, V. (1975). Acquisition of performatives prior to speech. *Merrill-Palmer Quarterly Journal of Developmental Psychology*, *21*(3), 205–226.

Bavelas, J., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: independent effects of dialogue and visibility. *Journal of Memory and Language*, *58*(2), 495–520. doi: 10.1016/j.jml.2007.02.004

Bavelas, J., Kenwood, C., Johnson, T., & Phillips, B. (2002). An experimental study of when and how speakers use gestures to communicate. *Gesture*, *2*(1), 1–17. doi: 10.1075/gest.2.1.02bav

Beattie, G., & Shovelton, H. (1999). Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *Journal of Language and Social Psychology*, *18*(4), 438–462. doi: 10.1177/0261927X99018004005

Beattie, G., & Shovelton, H. (2005). Why the spontaneous images created by the hands during talk can help make TV advertisements more effective. *British Journal of Psychology*, *96*, 21–37. doi: 10.1348/000712605X103500

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, *8*(6), 492–499. doi: 10.1111/j.1467-7687.2005.00440.x

Bellugi, U., van Hoek, K., Lillo-Martin, D. C., & O'Grady, L. (1988). The acquisition of syntax and space in young deaf signers. In B. D. & M. K. (Eds.), *Language development in exceptional circumstances* (pp. 132–149). Edinburgh: Churchill Livingstone.

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: a review. *International Journal of Psychophysiology*, *83*(2), 120–131. doi: 10.1016/j.ijpsycho.2011.08.003

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate:

a practicable and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, *29*(4), 1165–1188.

Bentin, S., Kutas, M., & Hillyard, S. A. (1995). Semantic processing and memory for attended and unattended words in dichotic listening: behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 54–67. doi: 10.1037/0096-1523.21.1.54

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*(4), 343–355. doi: 10.1016/0013-4694(85)90008-2

Bernardis, P., Salillas, E., & Caramelli, N. (2008). Behavioural and neurophysiological evidence of semantic interaction between iconic gestures and words. *Cogn Neuropsychol*, *25*(7-8), 1114–1128. doi: 10.1080/02643290801921707

Bishop, M. (2010). Happen can't hear: an analysis of code-blends in hearing, native signers of American Sign Language. *Sign Language Studies*, *11*(2), 205–240. doi: 10.1353/sls.2010.0007

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. *Cognition*, *104*(3), 437–458. doi: 10.1016/j.cognition.2006.07.003

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*(1), 55–73. doi: 10.1016/j.brainresrev.2008.05.003

Broadbent, D. E. (1966). *Perception and communication* (2nd ed.). Oxford: Pergamon Press.

Broaders, S. C., & Goldin-Meadow, S. (2010). Truth is at hand: how gesture adds information during investigative interviews. *Psychological Science*, *21*(5), 623–628. doi: 10.1177/0956797610366082

Bronkhorst, A. W. (2015). The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Attention Perception & Psychophysics*, *77*(5), 1465–1487. doi: 10.3758/s13414-015-0882-9

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in

language comprehension. *Brain Research*, *1446*, 127–143. doi: 10.1016/j.brainres.2012.01.055

Brouwer, H., & Hoeks, J. C. J. (2013). A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. *Frontiers in Human Neuroscience*, *7*, 1–12. doi: 10.3389/fnhum.2013.00758

Bullinger, A. F., Zimmermann, F., Kaminski, J., & Tomasello, M. (2011). Different social motives in the gestural communication of chimpanzees and human children. *Developmental Science*, *14*(1), 58–68. doi: 10.1111/j.1467-7687.2010.00952.x

Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: evidence from event-related brain potentials. *Brain and Language*, *98*(2), 159–168. doi: 10.1016/j.bandl.2006.04.005

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). Mahwah, NJ: Lawrence Erlbaum Associates.

Butterworth, G., & Morisette, P. (1996). Onset of pointing and the acquisition of language in infancy. *Journal of Reproductive and Infant Psychology*, *14*(3), 219–231. doi: 10.1080/02646839608404519

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22*(1), 77–126.

Carey, D., Mercure, E., Pizzioli, F., & Aydelott, J. (2014). Auditory semantic processing in dichotic listening: effects of competing speech, ear of presentation, and sentential bias on N400s to spoken words in context. *Neuropsychologia*, *65*, 102–112. doi: 10.1016/j.neuropsychologia.2014.10.016

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*(4), 1–143. doi: 10.2307/1166214

Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Research*, *51*(13), 1484–1525. doi: 10.1016/j.visres.2011.04.012

Cassell, J., McNeill, D., & McCullough, K.-E. (1999). Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics & Cognition*, *7*(1), 1–33. doi: 10.1075/pc.7.1.03cas

Chartrand, T. L., & Lakin, J. L. (2013). The antecedents and consequences of human behavioral mimicry. *Annual Review of Psychology*, *64*, 285–308. doi: 10.1146/annurev-psych-113011-143754

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and two ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979. doi: 10.1121/1.1907229

Church, R. B., Garber, P., & Rogalski, K. (2007). The role of gesture in memory and social communication. *Gesture*, *7*(2), 137–158. doi: 10.1075/gest.7.2.02bre

Clark, H. H. (1997). Dogmas of understanding. *Discourse processes*, *23*(3), 567–598.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*(1), 62–81. doi: 10.1016/j.jml.2003.08.004

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39.

Cohen, A. A., & Harrison, R. P. (1973). Intentionality in the use of hand illustrators in face-to-face communication situations. *Journal of Personality and Social Psychology*, *28*(2), 276–279. doi: 10.1037/h0035792

Coles, M. G. H., & Rugg, M. D. (1997). Event-related brain potentials: an introduction. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: event-related brain potentials and cognition* (pp. 1–26). Oxford: Oxford University Press.

Colonnesi, C., Stams, G. J. J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: a meta-analysis. *Developmental Review*, *30*(4), 352–366. doi: 10.1016/j.dr.2010.10.001

Cornejo, C., Simonetti, F., Ibáñez, A., Aldunate, N., Ceric, F., López, V., & Núñez, R. E. (2009). Gesture and metaphor comprehension: electrophysiological evidence of cross-modal coordination by audiovisual stimulation. *Brain and Cognition*, *70*(1), 42–52. doi: 10.1016/j.bandc.2008.12.005

Corteen, R. S., & Dunn, D. (1974). Shock-associated words in a nonattended message: a test for momentary awareness. *Journal of Experimental Psychology*, *102*(6), 1143–1144. doi: 10.1037/h0036362

Corteen, R. S., & Wood, B. (1972). Autonomic responses to shock-associated words in an unattended channel. *Journal of Expeirmental Psychology*, *94*(3), 308–313. doi: 10.1037/h0032759

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: evidence for lexical competition. *Language and Cognitive Processes*, *16*(5–6), 507–534.

Daneman, M., & Merikle, P. M. (1996). Working memory and language

comprehension: a meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433. doi: 10.3758/BF03214546

Daum, M. M., & Gredebäck, G. (2011a). The development of grasping comprehension in infancy: covert shifts of attention caused by referential actions. *Experimental Brain Research*, *208*(2), 297–307. doi: 10.1007/s00221-010-2479-9

Daum, M. M., & Gredebäck, G. (2011b). Spatial cueing by referential human gestures, arrows and mechanical devices. *International Journal of Mind, Brain and Cognition*, *2*(1–2), 113–126.

Daum, M. M., Ulber, J., & Gredebäck, G. (2013). The development of pointing perception in infancy: effects of communicative signals on covert shifts of attention. *Developmental Psychology*, *49*(10), 1898–1908. doi: 10.1037/a0031111

de Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory and Cognition*, *31*(3), 353–359.

Deutsch, J. A., & Deutsch, D. (1963). Attention – some theoretical considerations. *Psychological Review*, *70*(1), 80–90. doi: 10.1037/h0039515

Domjan, M. (2005). Pavlovian conditioning: a functional perspective. *Annual Review of Psychology*, *56*, 179–206. doi: 10.1146/annurev.psych.55.090902.141409

Donchin, E., Ritter, W., & McCallum, C. (1978). Cognitive psychophysiology: the endogenous components of the ERP. In E. Callaway, P. Tueting, & S. H. Koslow (Eds.), *Brain event-related potentials in man* (pp. 349–411). New York, NY: Academic Press.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*, 53–78. doi: 10.1348/000712601162103

Driver, J., & Spence, C. (1998). Crossmodal attention. *Current Opinion in Neurobiology*, *8*(2), 245–253. doi: 10.1016/S0959-4388(98)80147-5

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, *87*(3), 272–300. doi: 10.1037/0033-295X.87.3.272

Dunlap, A. S., & Stephens, D. W. (2014). Experimental evolution of prepared learning. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(32), 11750–11755. doi: 10.1073/pnas.1404176111

Dyer, F. N. (1971). Duration of word meaning responses: Stroop interference for different preexposures of word. *Psychonomic Science*, *25*(4),

229–231.

Efron, D. (1941). *Gesture and environment*. New York, NY: King's Crown Press.

Eimer, M., & Coles, M. G. H. (2003). The lateralized readiness potential. In M. Jahanshahi & M. Hallett (Eds.), *The Bereitschaftspotential: movement-related cortical potentials* (pp. 229–248). New York, NY: Kluwer Academic/Plenum Publishers.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*(4), 384–392. doi: 10.1037//0003-066X.48.4.384

Ekman, P., & Friesen, W. (1969). The repertoire of nonverbal behavior: categories, origins, usage and coding. *Semiotica*, *1*(1), 49–98.

Emmorey, K., Borinstein, H. B., Thompson, R., & Gollan, T. H. (2008). Bimodal bilingualism. *Bilingualism: Language and Cognition*, *11*(1), 43–61. doi: 10.1017/S1366728907003203

Emmorey, K., Corina, D., & Bellugi, U. (1995). Differential processing of topographic and referential functions of space. In K. Emmorey & J. S. Reilly (Eds.), *Language, gesture, and space* (pp. 43–62). Hillsdale, NJ: Lawrence Erlbaum Associates.

Emmorey, K., Petrich, J., & Gollan, T. H. (2012). Bilingual processing of ASL-English code-blends: The consequences of accessing two lexical representations simultaneously. *Journal of Memory and Language*, *67*(1), 199–210. doi: 10.1016/j.jml.2012.04.005

Emmorey, K., Petrich, J. A. F., & Gollan, T. H. (2013). Bimodal bilingualism and the frequency-lag hypothesis. *Journal of Deaf Studies and Deaf Education*, *18*(1), 1–11. doi: 10.1093/deafed/ens034

Enfield, N. J. (2001). 'lip-pointing': A discussion of form and function with reference to data from Laos. *Gesture*, *1*(2), 185–211. doi: 10.1075/gest.1.2.06enf.

Enfield, N. J., Kita, S., & de Ruiter, J. P. (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics*, *39*(10), 1722–1741. doi: 10.1016/j.pragma.2007.03.001

Engel, J. J. (1968). *Ideen zu einer Mimik*. Hildesheim: Olms.

Federmeier, K. D., & Kutas, M. (1999). A rose by another name: long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*(4), 469–495. doi: 10.1006/jmla.1999.2660

Fiebach, C. F., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German WH-questions. *Journal of Memory and Language*, *47*(2), 250–272. doi: 10.1016/S0749-596X(02)00004-9

Frick-Horburry, D., & Guttentag, R. E. (1998). The effects of restricting hand gesture production on lexical retrieval and free recall. *American Journal of Psychology*, *111*(1), 43–62. doi: 10.2307/1423536

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews*, *91*(4), 1357–1392. doi: 10.1152/physrev.00006.2011

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1219–1248. doi: 10.1037//0278-7393.22.5.1219

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to terms: quantifying the benefits of linguistic coordination. *Psychological Science*, *23*(8), 931–939. doi: 10.1177/0956797612436816

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": an electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience*, *8*(2), 89–106. doi: 10.1162/jocn.1996.8.2.89

Garcia, J., Kimeldorf, D. J., & Koelling, R. A. (1955, Jul). Conditioned aversion to saccharin resulting from exposure to gamma radiation. *Science*, *122*(3160), 157–158.

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science*, *4*(1), 123–124. doi: 10.3758/BF03342209

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition*, *27*(2), 181–218. doi: 10.1016/0010-0277(87)90018-7

Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage*, *35*(4), 1674–1684. doi: 10.1016/j.neuroimage.2007.02.003

Gifford, R. (1994). A lens-mapping framework for understanding the encoding and decoding of interpersonal dispositions in nonverbal behavior. *Journal of Personality and Social Psychology*, *66*(2), 398–412. doi: 10.1037/0022-3514.66.2.398

Giordano, A. M., McElree, B., & Carrasco, M. (2009). On the automaticity and flexibility of covert attention: a speed-accuracy trade-off analysis. *J Vis*, *9*(3), 30.1–3010. doi: 10.1167/9.3.30

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop

phenomenon. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(6), 875–894. doi: 10.1037//0096-1523.8.6.875

Goodwin, C. (1986). Between and within: alternative sequential treatments of continuers and assessments. *Human Studies*, *9*(2–3), 205–217. doi: 10.1007/BF00148127

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, *25*(2), 149–188. doi: 10.1080/01690960902965951

Graham, J. A., & Argyle, M. (1975). A cross-cultural study of the communication of extra-verbal meaning by gestures. *International Journal of Psychology*, *10*(1), 57–67. doi: 10.1080/00207597508247319

Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *The Quarterly Journal of Experimental Psychology*, *39A*(2), 295–319.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112. doi: 10.1007/BF02289823

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Gullberg, M. (2006). Handling discourse: gestures, reference tracking, and communication strategies in early l2. *Language Learning*, *56*(1), 155–196. doi: 10.1111/j.0023-8333.2006.00344.x

Gullberg, M., & Holmqvist, K. (2006). What speakers do and what addressees look at: Visual attention to gestures in human interaction live and on video. *Pragmatics & Cognition*, *14*(1), 53–82. doi: 10.1075/pc.14.1.05gul

Gunter, T. C., & Bach, P. (2004). Communicating hands: ERPs elicited by meaningful symbolic hand postures. *Neuroscience Letters*, *372*(1–2), 52–56. doi: 10.1016/j.neulet.2004.09.011

Gunter, T. C., Weinbrenner, J. E. D., & Holle, H. (2015). Inconsistent use of gesture space during abstract pointing impairs language comprehension. *Frontiers in Psychology*, *6*, 1–10. doi: 10.3389/fpsyg.2015.00080

Habets, B., Kita, S., Shao, Z., Özyürek, A., & Hagoort, P. (2011). The role of synchrony and ambiguity in speech-gesture integration during comprehension. *Journal of Cognitive Neuroscience*, *23*(8), 1845–1854. doi: 10.1162/jocn.2010.21462

Hadar, U., & Pinchas-Zamir, L. (2004). The semantic specificity of gesture: implications for gesture classification and function. *Journal of Language and Social Psychology*, *23*(2), 204–214. doi:

10.1177/0261927X04263825

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), 439–483. doi: 10.1080/01690969308407585

Hahne, A., & Friederici, A. D. (2002). Diferential task effects on semantic and syntactic processes as reavealed by erps. *Cognitive Brain Research*, *13*(3), 339–356. doi: 10.1016/S0926-6410(01)00127-6

Hansen, J. C., Dickstein, P. W., Berka, C., & Hillyard, S. A. (1983). Event-related potentials during selective attention to speech sounds. *Biological Psychology*, *16*(3–4), 211–224. doi: 10.1016/0301-0511(83)90025-X

Hare, B., & Tomasello, M. (2005). Human-like social skills in dogs? *Trends in Cognitive Sciences*, *9*(9), 439–444. doi: 10.1016/j.tics.2005.07.003

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177–180. doi: 10.1126/science.182.4108.177

Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: the interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, *19*(1), 59–73. doi: 10.1016/j.cogbrainres.2003.10.022

Holcomb, P. J. (1988). Automatic and attentional processing: an event-related brain potential analysis of semantic priming. *Brain and Language*, *35*(1), 66–85. doi: 10.1016/0093-934X(88)90101-0

Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: an electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(3), 721–742. doi: 10.1037/0278-7393.25.3.721

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. *Language and Cognitive Processes*, *5*(4), 281–312. doi: 10.1080/01690969008407065

Holle, H., & Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: erp evidence. *Journal of Cognitive Neuroscience*, *19*(7), 1175–1192. doi: 10.1162/jocn.2007.19.7.1175

Holle, H., Gunter, T. C., Rüschemeyer, S.-A., Hennenlotter, A., & Iacoboni, M. (2008). Neural correlates of the processing of co-speech gestures. *Neuroimage*, *39*(4), 2010–2024. doi: 10.1016/j.neuroimage.2007.10.055

Holle, H., Obermeier, C., Schmidt-Kassow, M., Friederici, A. D., Ward, J., & Gunter, T. C. (2012). Gesture facilitates the syntactic analysis of speech. *Frontiers in Psychology*, *3*, 1–12. doi: 10.3389/fpsyg.2012.00074

Holler, J., Kokal, I., Toni, I., Hagoort, P., Kelly, S. D., & Özyürek, A. (2015). Eye'm talking to you: speakers' gaze direction modulates co-speech gesture processing in the right MTG. *Social Cognitive & Affective Neuroscience*, *10*(2), 255–261. doi: 10.1093/scan/nsu047

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*(1), 4–27. doi: 10.1177/0261927X06296428

Holler, J., & Wilkin, K. (2011). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*, *35*(2), 133–153. doi: 10.1007/s10919-011-0105-6

Hollis, K. L. (1997). Contemporary research on pavlovian conditioning: A "new" functional analysis. *American Psychologist*, *52*(9), 956–965.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Hommel, B. (1993). The relationship between stimulus processing and response selection in the Simon task: evidence for a temporal overlap. *Psychological Research*, *55*(4), 280–290. doi: 10.1007/BF00419688

Hommel, B. (2002). Responding to object files: automatic integration of spatial information revealed by stimulus-response compatibility effects. *Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *55*(2), 567–580. doi: 10.1080/02724980143000361

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychologica*, *136*(2), 189–202. doi: 10.1016/j.actpsy.2010.04.011

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*(2), 297–315. doi: 10.1037/a0022128

Hostetter, A. B., & Alibali, M. W. (2007). Raise your hand if you're spatial: relations between verbal and spatial skills and gesture production. *Gesture*, *7*(1), 73–95. doi: 10.1075/gest.7.1.05hos

Hostetter, A. B., & Potthoff, A. L. (2012). Effects of personality and social situation on representational gesture production. *Gesture*, *12*(1), 62–83. doi: 10.1075/gest.12.1.04hos

Ibáñez, A., Manes, F., Escobar, J., Trujillo, N., Andreucci, P., & Hurtado, E. (2010). Gesture influences the processing of figurative language in non-native speakers: ERP evidence. *Neuroscience Letters*, *471*(1),

48–52. doi: 10.1016/j.neulet.2010.01.009

Ibáñez, A., Toro, P., Cornejo, C., Urquina, H., Hurquina, H., Manes, F., ... Schröder, J. (2011). High contextual sensitivity of metaphorical expressions and gesture blending: A video event-related potential design. *Psychiatry Research: Neuroimaging*, *191*(1), 68–75. doi: 10.1016/j.pscychresns.2010.08.008

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*(5), 367–371. doi: 10.1111/j.0956-7976.2005.01542.x

Jacobs, N., & Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*, *56*(2), 291–303. doi: 10.1016/j.jml.2006.07.011

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201.

Kappenman, E. S., & Luck, S. J. (2012). ERP components: the ups and downs of brainwave recordings. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 3–30). New York, NY: Oxford University Press.

Keller, J. (1998). *Aspekte der Raumnutzung in der deutschen Gebärdensprache*. Hamburg, Germany: Signum.

Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: the role of speech and gesture in comprehension and memory. *Journal of Memory and Language*, *40*(4), 577–592. doi: 10.1006/jmla.1999.2634

Kelly, S. D., Creigh, P., & Bartolotti, J. (2010). Integrating speech and iconic gestures in a stroop-like task: evidence for automatic processing. *Journal of Cognitive Neuroscience*, *22*(4), 683–694. doi: 10.1162/jocn.2009.21254

Kelly, S. D., & Goldsmith, L. H. (2004). Gesture and right hemisphere involvement in evaluating lecture material. *Gesture*, *4*(1), 25–42. doi: 10.1075/gest.4.1.03kel

Kelly, S. D., Kravitz, C., & Hopkins, M. (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and Language*, *89*(1), 253–260. doi: 10.1016/S0093-934X(03)00335-3

Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. *Psychological Science*, *21*(2), 260–267. doi: 10.1016/j.jml.2006.07.011

Kelly, S. D., Ward, S., Creigh, P., & Bartolotti, J. (2007). An inten-

tional stance modulates the integration of gesture and speech during comprehension. *Brain and Language*, *101*(3), 222–233. doi: 10.1016/j.bandl.2006.07.008

Kendon, A. (2004). *Gesture: visible action as utterance*. Cambridge: Cambridge University Press.

Kendon, A., & Versante, L. (2003). Pointing by hand in "Neapolitan". In S. Kita (Ed.), *Pointing: where language, culture, and cognition meet* (pp. 109–137). Mahwah, NJ: Lawrence Erlbaum Associates.

Keselman, H. J. (1998). Testing treatment effects in repeated measures designs: an update for psychophysiological researchers. *Psychophysiology*, *35*(4), 470–478. doi: 10.1017/S0048577298000237

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: evidence from event-related potentials. *Journal of Memory and Language*, *52*(2), 205–225. doi: 10.1016/j.jml.2004.10.002

Kimbara, I. (2008). Gesture form convergence in joint description. *Journal of Nonverbal Behavior*, *32*(2), 123–131. doi: 10.1007/s10919-007-0044-4

Klima, E. S., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.

Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. W. (2003). Structure and limited capacity in verbal working memory: a study with event-related potentials. *Brain and Language*, *85*(1), 1–36. doi: 10.1016/S0093-934X(02)00548-5

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility–a model and taxonomy. *Psychological Review*, *97*(2), 253–270. doi: 10.1037/0033-295X.97.2.253

Krauss, R. M., Dushay, R. A., Chen, Y., & Rauscher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology*, *31*(6), 533–552. doi: 10.1006/jesp.1995.1024

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology*, *61*(5), 743–754. doi: 10.1037/0022-3514.61.5.743

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Research*, *1146*, 23–49. doi: 10.1016/j.brainres.2006.12.063

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*(1), 117–129. doi:

10.1016/S0926-6410(03)00086-7

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Reviews Psychology*, *62*, 621–647. doi: 10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980a). Event-related potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*(2), 99–116. doi: 10.1016/0301-0511(80)90046-0

Kutas, M., & Hillyard, S. A. (1980b). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. doi: 10.1126/science.7350657

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163. doi: 10.1016/S1364-6613(00)01560-6

Lange, K., Rösler, F., & Röder, B. (2003). Early processing stages are modulated when auditory stimuli are presented at an attended moment in time: an event-related potential study. *Psychophysiology*, *40*(5), 806–817. doi: 10.1111/1469-8986.00081

Lattner, S., & Friederici, A. D. (2003). Talker's voice and gender stereotype in human auditory sentence processing – evidence from event-related brain potentials. *Neuroscience Letters*, *339*(3), 191–194. doi: 10.1016/S0304-3940(03)00027-2

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933. doi: 10.1038/nrn2532

Leavens, D. A., Hopkins, W. D., & Bard, K. A. (2005). Understanding the point of chimpanzee pointing: epigenesis and ecological validity. *Current Directions in Psychological Science*, *14*(4), 185–189. doi: 10.1111/j.0963-7214.2005.00361.x

Lelekov, T., Dominey, P. F., & Garcia-Larrea, L. (2000). Dissociable ERP profiles for processing rules vs instances in a cognitive sequencing task. *NeuroReport*, *11*(5), 1129–1132. doi: 10.1097/00001756-200004070-00043

Leung, E. H. L., & Rheingold, H. L. (1981). Development of pointing as a social gesture. *Developmental Psychology*, *17*(2), 215–220. doi: 10.1037//0012-1649.17.2.215

Liddell, S. K. (1980). *American Sign Language syntax*. The Hague: Mouton.

Lim, V. K., Wilson, A. J., Hamm, J. P., Phillips, N., Iwabuchi, S. J., Corballis, M. C., … Thomas, M. O. J. (2009). Semantic processing of mathematical gestures. *Brain and Cognition*, *71*(3), 306–312. doi: 10.1016/j.bandc.2009.07.004

Liszkowski, U., Carpenter, M., & Tomasello, M. (2007). Pointing out new news, old news, and absent referents at 12 months of age. *Developmental Science*, *10*(2), F1-F7. doi: 10.1111/j.1467-7687.2006.00552.x

Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, *108*(3), 732–739. doi: 10.1016/j.cognition.2008.06.013

Liszkowski, U., Schäfer, M., Carpenter, M., & Tomasello, M. (2009). Prelinguistic infants, but not chimpanzees, communicate about absent entities. *Psychological Science*, *20*(5), 654–660. doi: 10.1111/j.1467-9280.2009.02346.x

Lu, C. H., & Proctor, R. W. (2001). Influence of irrelevant information on human performance: effects of s-r association strength and relative timing. *Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *54*(1), 95–136. doi: 10.1080/02724980042000048

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin*, *109*(2), 163–203.

MacSweeney, M., Capek, C. M., Campbell, R., & Woll, B. (2008). The signing brain: the neurobiology of sign language. *Trends in Cognitive Sciences*, *12*(11), 432–440. doi: 10.1016/j.tics.2008.07.010

Maricchiolo, F., Gnisci, A., Bonaiuto, M., & Ficca, G. (2009). Effects of different types of hand gestures in persuasive speech on receivers's evaluations. *Language and Cognitive Processes*, *24*(2), 239–266. doi: 10.1080/01690960802159929

McCallum, W. C., Farmer, S. F., & Pocock, P. V. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology*, *59*(6), 477–488. doi: 10.1016/0168-5597(84)90006-6

McCrae, R. R., & Costa, P., Jr. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*(5), 509–516. doi: 10.1037/0003-066X.52.5.509

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. doi: 10.1038/264746a0

McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*.

Chicago, IL: The University of Chicago Press.
McNeill, D. (2003). Pointing and morality in Chicago. In S. Kita (Ed.), *Pointing: where language, culture, and cognition meet* (pp. 293–306). Mahwah, NJ: Lawrence Erlbaum Associates.
McNeill, D. (2005). *Gesture and thought*. Chicago, IL: The University of Chicago Press.
McNeill, D., Cassell, J., & Levy, E. T. (1993). Abstract deixis. *Semiotica*, *95*(1–2), 5–19. doi: 10.1515/semi.1993.95.1-2.5
McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology*, *36*(1), 53–65. doi: 10.1017/S0048577299971196
Melinder, A. M. D., Konijnenberg, C., Hermansen, T., Daum, M. M., & Gredebäck, G. (2015, Feb). The developmental trajectory of pointing perception in the first year of life. *Experimental Brain Research*, *233*(2), 641–647. doi: 10.1007/s00221-014-4143-2
Miklósi, Á., & Soproni, K. (2006). A comparative analysis of animals' understanding of the human pointing gesture. *Animal Cognition*, *9*(2), 81–93. doi: 10.1007/s10071-005-0008-1
Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception and Psychophysics*, *40*(5), 331–343.
Mitchell, C. J., de Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behav Brain Sci*, *32*(2), 183–98; discussion 198-246. doi: 10.1017/S0140525X09000855
Moray, N. (1959). Attention in dichotic listening: affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *11*(1), 56–60. doi: 10.1080/17470215908416289
Morford, J. P., & Carlson, M. L. (2011). Sign perception and recognition in non-native signers of asl. *Language Learning and Development*, *7*(2), 149–168. doi: 10.1080/15475441.2011.543393
Mulcahy, N. J., & Hedge, V. (2012). Are great apes tested with an object-choice task? *Animal Behaviour*, *83*(2), 313–321. doi: 10.1016/j.anbehav.2011.11.019
Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(2), 315–330. doi: 10.1037/0096-1523.15.2.315
Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. G. (2000). *The syntax of American Sign Language*. Cambridge, MA: The MIT Press.
Neville, H. J., Coffey, S. S., Lawson, D. S., Fischer, A., Emmorey, K., &

Bellugi, U. (1997). Neural systems mediating american sign language: effects of sensory experience and age of acquisition. *Brain and Language*, *57*(3), 285–308. doi: 10.1006/brln.1997.1739

Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3*(2), 151–165. doi: 10.1162/jocn.1991.3.2.151

Nielsen, S. L., & Sarason, I. G. (1981). Emotion, personality, and selective attention. *Journal of Personality and Social Psychology*, *41*(5), 945–960. doi: 10.1037//0022-3514.41.5.945

Nùñez-Peña, M. I., & Honrubia-Serrano, M. L. (2004). P600 related to rule violation in an arithmetic task. *Cognitive Brain Research*, *18*(2), 130–141. doi: 10.1016/j.cogbrainres.2003.09.010

Näätänen, R., Kujala, T., & Winkler, I. (2011). Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology*, *48*(1), 4–22. doi: 10.1111/j.1469-8986.2010.01114.x

Obermeier, C., Dolk, T., & Gunter, T. C. (2012). The benefit of gestures during communication: Evidence from hearing and hearing-impaired individuals. *Cortex*, *48*(7), 857–870. doi: 10.1016/j.cortex.2011.02.007

Obermeier, C., Holle, H., & Gunter, T. C. (2011). What iconic gesture fragments reveal about gesture-speech integration: when synchrony is lost, memory can help. *Journal of Cognitive Neuroscience*, *23*(7), 1648–1663. doi: 10.1162/jocn.2010.21498

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, *108*(3), 483–522. doi: 10.1037//0033-295X.108.3.483

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. doi: 10.1016/0028-3932(71)90067-4

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806. doi: 10.1016/0749-596X(92)90039-Z

Özyürek, A., Willems, R. M., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: insights from event-related brain potentials. *Journal of Cognitive Neuroscience*, *19*(4), 605–616. doi: 10.1162/jocn.2007.19.4.605

Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: The MIT Press.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733. doi: 10.1162/089892998563121

Pavlov, I. P. (1940). *Conditioned reflexes. An investigation of the physiological activity of the cerebral cortex* (G. V. Anrep, Trans.). Oxford: Oxford University Press. (Original work published 1927)

Pellicano, A., Vu, K.-P. L., Proctor, R. W., Nicoletti, R., & Umiltà, C. (2008). Effects of stimulus-stimulus short-term memory associations in a Simon-like task. *European Journal of Cognitive Psychology*, *20*(5), 893–912. doi: 10.1080/09541440701676343

Phillips, C., Kazanina, N., & Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, *22*(3), 407–428. doi: 10.1016/j.cogbrainres.2004.09.012

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: a critical review. *Psychological Bulletin*, *134*(3), 427–459. doi: 10.1037/0033-2909.134.3.427

Poizner, H., Klima, E. S., & Bellugi, U. (1987). *What the hands reveal about the brain*. Cambridge, MA: The MIT Press.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25. doi: 10.1080/00335558008248231

Quintilian, M. F. (1995). *Ausbildung des Redners* (3rd ed.). Darmstadt: Wissenschaftliche Buchgesellschaft.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*, 574–590.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: the role of lexical movements in speech production. *Psychological Science*, *7*(4), 226–231. doi: 10.1111/j.1467-9280.1996.tb00364.x

Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience*, *23*(2), 277–293. doi: 10.1162/jocn.2010.21411

Rescorla, R. A. (1988). Behavioral studies of pavlovian conditioning. *Annu Rev Neurosci*, *11*, 329–352. doi: 10.1146/annurev.ne.11.030188.001553

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian condi-

tioning: variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. E. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Rimé, B. (1982). The elimination of visible behaviour from social interactions: effects on verbal, nonverbal and interpersonal variables. *European Journal of Social Psychology*, *12*(2), 113–129. doi: 10.1002/ejsp.2420120201

Riseborough, M. G. (1981). Physiographic gestures as decoding facilitators: three experiments exploring a neglected facet of communication. *Journal of Nonverbal Behavior*, *5*(3), 172–183. doi: 10.1007/BF00986134

Ro, T., Russell, C., & Lavie, N. (2001). Changing faces: a detection advantage in the flicker paradigm. *Psychological Science*, *12*(1), 94–99. doi: 10.1111/1467-9280.00317

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simples systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735. doi: 10.1353/lan.1974.0010

Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing. I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research*, *15*(3), 228–240. doi: 10.1016/S0926-6410(02)00195-7

Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, *5*(7), 700–703. doi: 10.1038/nn873

Sanford, A. J., Leuthold, H., Bohan, J., & Sanford, A. J. S. (2011). Anomalies at the borderline of awareness: an ERP study. *Journal of Cognitive Neuroscience*, *23*(3), 514–523. doi: 10.1162/jocn.2009.21370

Schröger, E., & Widmann, A. (1998). Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology*, *35*(6), 755–759.

Seaman, M. A., Levin, J. R., & Serlin, R. C. (1991). New developments in pairwise multiple comparisons: some powerful and practicable procedures. *Psychological Bulletin*, *110*(3), 577–586. doi: 10.1037/0033-2909.110.3.577

Sekine, K., & Kita, S. (2015). Development of multimodal discourse comprehension: cohesive use of space by gestures. *Language, Cognition and Neuroscience*, *30*(10), 1245–1258. doi: 10.1080/23273798.2015.1053814

Seligman, M. E. P. (1970). On the generality of the laws of learning. *Psychological Review*, *77*(5), 406–418. doi: 10.1037/h0029790

Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy*, *2*(3), 307–320. doi: 10.1016/S0005-7894(71)80064-3

Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences*, *13*(3), 127–134. doi: 10.1016/j.tics.2008.11.009

Shaffer, J. P. (1986). Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association*, *81*(395), 826–831. doi: 10.2307/2289016

Simon, J. R., & Small, A., Jr. (1969). Processing auditory information: interference from an irrelevant cue. *Journal of Applied Psychology*, *53*(5), 433–435. doi: 10.1037/h0028034

Sitnikova, T., Kuperberg, G., & Holcomb, P. J. (2003). Semantic integration in videos of real-world events: An electrophysiological investigation. *Psychophysiology*, *40*(1), 160–164. doi: 10.1111/1469-8986.00016

Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2007). Speech-associated gestures, broca's area, and the human mirror system. *Brain and Language*, *101*(3), 260–277. doi: 10.1016/j.bandl.2007.02.008

So, W. C., Coppola, M., Licciardello, V., & Goldin-Meadow, S. (2005). The seeds of spatial grammar in in the manual modality. *Cognitive Science*, *29*(6), 1029–1043. doi: 10.1207/s15516709cog0000_38

So, W. C., Kita, S., & Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: gesture and speech go hand-in-hand. *Cognitive Science*, *33*(1), 115–125. doi: 10.1111/j.1551-6709.2008.01006.x

Stec, K., & Huiskes, M. (2014). Co-constructing referential space in multimodal narratives. *Cognitive Semiotics*, *7*(1), 31–59. doi: 10.1515/cogsem-2014-0003

Straube, E. R., & Germer, C. K. (1979). Dichotic shadowing and selective attention to word meaning in schizophrenia. *Journal of Abnormal Psychology*, *88*(4), 346–353. doi: 10.1037/0021-843X.88.4.346

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. doi: 10.1037/0096-3445.121.1.15

Tabensky, A. (2001). Gesture and speech rephrasings in conversation. *Gesture*, *1*(2), 213–236. doi: 10.1075/gest.1.2.07tab

Tagliabue, M., Zorzi, M., Umiltà, C., & Bassignani, F. (2000). The role

of long-term-memory and short-term-memory links in the Simon effect. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 648–670. doi: 10.1037//0096-1523.26.2.648

Thomas, M. S. C., & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language*, *43*(1), 44–66. doi: 10.1006/jmla.1999.2700

Tlauka, M., & McKenna, F. P. (1998). Mental imagery yields stimulus-response compatibility. *Acta Psychologica*, *98*(1), 67–79.

Tomasello, M. (2005). *Origins of human communication*. Cambridge, MA: The MIT Press.

Tomasello, M., & Call, J. (2007). Ape gestures and the origins of language. In J. Call & M. Tomasello (Eds.), *The gestural communication of apes and monkeys* (pp. 221–239). Mahwah, NJ: Lawrence Erlbaum Associates.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, *78*(3), 705–722. doi: 10.1111/j.1467-8624.2007.01025.x

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological Review*, *76*(3), 282–299. doi: 10.1037/h0027242

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory and Cognition*, *22*(1), 111–126. doi: 10.3758/BF03202766

Valenza, E., Simion, F., Cassia, V. M., & Umiltà, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 892–903. doi: 10.1037/0096-1523.22.4.892

van Berkum, J. J. A. (2009). The neuropragmatics of 'simple' utterance comprehension: an ERP review. In U. Sauerland & K. Yatsushiro (Eds.), *Semantics and pragmatics: from experiment to theory* (pp. 276–316). Basingstoke: Palgrave Macmillan.

van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: evidence from the N400. *Journal of Cognitive Neurosciene*, *11*(6), 657–671.

van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580–591. doi: 10.1162/jocn.2008.20054

van de Meerendonk, N., Kolk, H. H. J., Chwilla, D. J., & Vissers,

C. T. W. M. (2009). Monitoring in language perception. *Language and Linguistics Compass*, *3*(5), 1211–1224. doi: 10.1162/jocn.2008.21170

van de Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwilla, D. J. (2010). Monitoring in language perception: mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience*, *22*(1), 67–82. doi: 10.1162/jocn.2008.21170

van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, *18*(7), 1181–1197. doi: 10.1162/jocn.2006.18.7.1181

van Herten, M., Kolk, H. H., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22*(2), 241–255. doi: 10.1016/j.cogbrainres.2004.09.002

van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 394–417. doi: 10.1037/0278-7393.25.2.394

von Wright, J. M., Anderson, K., & Stenman, U. (1975). Generalization of conditioned GSRs in dichotic listening. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance* (pp. 194–204). London: Academic Press.

Vu, K.-P. L. (2007). Influences on the simon effect of prior practice with spatially incompatible mappings: transfer within and between horizontal and vertical dimensions. *Memory and Cognition*, *35*(6), 1463–1471. doi: 10.3758/BF03193616

Weisberg, J., McCullough, S., & Emmorey, K. (2015). Simultaneous perception of a spoken and a signed language: The brain basis of ASL-English code-blends. *Brain and Language*, *147*, 96–106. doi: 10.1016/j.bandl.2015.05.006

West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research*, *13*(3), 363–375. doi: 10.1016/S0926-6410(01)00129-X

Wilkins, D. (2003). Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In S. Kita (Ed.), *Pointing: where language, culture, and cognition meet* (pp. 171–215). Mahwah, NJ: Lawrence Erlbaum Associates.

Willems, R. M., Ozyürek, A., & Hagoort, P. (2007). When language meets action: the neural integration of gesture and speech. *Cerebral Cortex*, *17*(10), 2322–2333. doi: 10.1093/cercor/bhl141

Wood, N. L., & Cowan, N. (1995). The cocktail party phenomenon revisited: attention and memory in the classic selective listening procedure of cherry (1953). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 255–260. doi: 10.1037/0096-3445.124.3.243

Wu, Y. C., & Coulson, S. (2005). Meaningful gestures: electrophysiological indices of iconic gesture comprehension. *Psychophysiology*, *42*(6), 654–667. doi: 10.1111/j.1469-8986.2005.00356.x

Wu, Y. C., & Coulson, S. (2007a). How iconic gestures enhance communication: an ERP study. *Brain and Language*, *101*(3), 234–245. doi: 10.1016/j.bandl.2006.12.003

Wu, Y. C., & Coulson, S. (2007b). Iconic gestures prime related concepts: an ERP study. *Psychonomic Bulletin & Review*, *14*(1), 57–63. doi: 10.3758/BF03194028

Wu, Y. C., & Coulson, S. (2010). Gestures modulate speech processing early in utterances. *Neuroreport*, *21*(7), 522–526. doi: 10.1097/WNR.0b013e32833904bb

Wu, Y. C., & Coulson, S. (2011). Are depictive gestures like pictures? Commonalities and differences in semantic processing. *Brain and Language*, *119*(3), 184–195. doi: 10.1016/j.bandl.2011.07.002

Wu, Y. C., & Coulson, S. (2014). Co-speech iconic gestures and visuo-spatial working memory. *Acta Psychologica*, *153*, 39–50. doi: 10.1016/j.actpsy.2014.09.002

Wundt, W. (1911). Die Sprache. In *Völkerpsychologie: eine Untersuchung der Entwicklungsgesetze von Sprache, Mythus und Sitte* (3rd ed.). Leipzig: Kröner.

Yang, J., Andric, M., & Mathew, M. M. (2015). The neural basis of hand gesture comprehension: A meta-analysis of functional magnetic resonance imaging studies. *Neuroscience and Biobehavioral Reviews*, *57*, 88–104. doi: 10.1016/j.neubiorev.2015.08.006

Zhang, H., & Kornblum, S. (1998). The effects of stimulus-response mapping and irrelevant stimulus-response and stimulus-stimulus overlap in four-choice Stroop tasks with single-carrier stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 3–19.

Zhang, J. X., & Johnson, M. K. (2004). A memory-based, Simon-like, spatial congruence effect: evidence for persisting spatial codes. *Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *57*(3), 419–436. doi: 10.1080/02724980343000404

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185. doi: 10.1037/0033-2909.123.2.162

A Examplar interview topics

The following section provides transcripts of six exemplar topics. At the beginning of the interview, the interviewee was introduced to the audience as Sabine. The language of the topics is originally German, but for the purpose of this presentation they were translated. Each topic ends with an arbitrarily chosen version of the critical response. Note, that the topics were not presented in sequence, but are taken from different timepoints of the interview.

spruce and oak (2 establishments):
Interviewer: At least for parquet floor you need a lot of wood. Wood is gained from trees. Which is the most frequent broadleef tree in Germany and which is the most frequent needle beam?
Sabine: Er ...
Interviewer: Great transition.
Sabine: Oak? Spruce? But as a matter of fact, I don't really know.
Interviewer: Hm, okay, well, I guess, at least the spruce ... well, I would have guessed spruce, too. Okay, let's move forward ... yes, and from which one are there more? Spruce? Oak?
Sabine: Spruce.
Interviewer: You have an oak forest in the garden of your home and a spruce forest ...
Sabine: [looks irritated]
Interviewer: Well, just imagine that you have a big garden.
Sabine: Oh, so I am the forester's daughter!
Interviewer: Exactly! Through which forest do you prefer to stroll?
Sabine: Through the oak forest.
Interviewer: Hm. For what purposes are both types of wood utilized?
Sabine: For furniture?
Interviewer: Yes. Now you can pick a piece of furniture made out of one of those wood types. What do you choose?
Sabine: Then I choose this wood – fits better to the rest, because I have already quite a lot made out of spruce at home.

vodka and tequila (2 establishments):
Interviewer: Aha. Now we switch to hard liquor like tequila and vodka. Where do these drinks come from?
Sabine: Tequila from Mexico and vodka is produced in Russia and Poland, I think.
Interviewer: Er, how do you usually drink these liquors?
Sabine: Tequila is downed in one and vodka is also downed in one, but I think it is served cold.
Interviewer: Mhm. And which of these firewaters do you like more?
Sabine: Well, if really necessary then this drink, although you cannot say that I really enjoy tequila.

cannabis and nicotine (3 establishments):
Interviewer: With this topic we have basically already drifted into the criminal environment. And in this environment people tend to consume drugs, unfortunately. And my first question in this regard is, which drug is in your opinion socially more accepted – cannabis or nicotine?
Sabine: Oh, oh, oh, oh! Of course, nicotine is socially accepted here. Well, it is rare that somebody takes offense at that. Er, but in the Netherlands it is maybe the cannabis, I don't know.
Interviewer: Hm, and although some drugs might be socially accepted, we would like to emphasize at this point, that nobody should use drugs, okay? The MPI takes no other stance! So, for what kind of drug are there more consumers? Or what is smoked more often?
Sabine: Oh, yes, hm, I think that nicotine is more frequently used.
Interviewer: Okay, is it possible to use these drugs as medicine, too?
Sabine: No, that is something that I have only heard of with regard to cannabis.
Interviewer: Hm, how about the costs for these drugs? As of today, this question makes almost sense ... considering the costs for cigarettes.
Sabine: Yes, but I don't have a clue about that [smiles]. However, I guess that nicotine is still less pricey. Er, regarding cannabis ... don't know. No, I don't know.
Interviewer: Both is very often consumed in form of cigarettes. Admittedly, in case of cannabis you call that a joint, but the principle idea is not different. Er, as a consequence ... or a frequent question regarding this matter is, what is more harmful for the lung. What is your opinion on that?
Sabine: Hm, I think that this drug is worse in this case, because the really

bad thing about cannabis is that a joint is inhaled much more deeply than a cigarette.

Goethe and Shakespeare (3 establishments):
Interviewer: Dear Sabine, we will talk now ... or I would like to ... of course, you have the choice, okay, but I would really really like to talk with you about literature.
Sabine: Mhm.
Interviewer: The first question is, when talking about authors, which true giants, classics come to your mind?
Sabine: Oh, hm, Goethe and Shakespeare maybe.
Interviewer: When did ...
Sabine: [coughs] Just wait a second [coughs] Are you gone, little frog? Okay, sorry.
Interviewer: Fine again? Okay. When did they live approximately?
Sabine: Oh, well, Goethe roughly in the 18th century and Shakespeare, I think, 16th, 17th century.
Interviewer: I will simply believe you in this matter. And where do they come from?
Sabine: Well, Shakespeare is from England and Goethe was a German.
Interviewer: Yes. Well, and now a decision question again. So, if you had to pick your most favorite book, from which one would that be?
Sabine: Then this classic would win, because I've rarely read something as beautiful as Shakespeare's *Romeo and Juliet*.

Dresden and Leipzig (4 establishments):
Interviewer: Ok, so in eastern Germany you could visit Dresden and Leipzig, what exactly could you visit there?
Sabine: In Dresden, you could start with the Blue Wonder, which is quite a famous bridge. Then there is the Dresdner Zwinger, the Semperoper and, of course, the Frauenkirche. And in Leipzig there is the Monument of the Battle of the Nations, the St. Thomas Church, the St. Nicholas Church and, well, yeah, the fairground itself is worth a visit.
Interviewer: Yes.
Sabine: Yes, and the fairs happening over there, of course.
Interviewer: Ok. Both cities tend to be compared with international cities. What comes to your mind?
Sabine: Dresden is "Florence of the Elbe" and Leipzig is called "Little Paris" or "Athens of the Pleiße".

Interviewer: Mhm. Which famous people do you link to those cities.
Sabine: In case of Leipzig, that would be definitely Goethe. After all, he let his Faust take a seat in Auerbach's Cellar. And for Dresden, well, I think of Erich Kästner.
Interviewer: Mhm.
Sabine: I can also think of musicians. For Leipzig, this would be Bach.
Interviewer: Yes.
Sabine: And for Dresden, well, ok, this is kind of cheating, Robert Schumann. He was in Dresden, but in Leipzig as well.
Interviewer: Mhm, well done.
Sabine: And in Heidelberg. He also was in Heidelberg.
Interviewer: A much-traveled man... Which one is the capitol of Saxony?
Sabine: Dresden.
Interviewer: Mhm, and which city do you rather link to the Reunification?
Sabine: Leipzig.
Interviewer: Mhm. What do you think, which city has more residents?
Sabine: Well, this city maybe? I hope, the people in the audience don't know better and are in consequence fine with my vote for Dresden.

Schwarzenegger and Stallone (4 establishments):
Interviewer: Well, now, oh yes, we come to the fine arts. Well, it'll be, it'll be great! Please, give me the names of actors, who became famous through action movies.
Sabine: But you said "fine arts"?
Interviewer: Correct!
Sabine: Hm, er ... the viewers will start laughing. Arnold Schwarzenegger, Sylvester Stallone.
Interviewer: Well, okay, that is already ...
Sabine: You said action movies!
Interviewer: That's alright. Yes, I would definitely count them to this genre. What are probably their most famous movies?
Sabine: Schwarzenegger – *Terminator* and Stallone – *Rocky*.
Interviewer: Okay, and where do these actors originally come from?
Sabine: Stallone is born in the USA, but I think he has Italian roots. And Schwarzenegger is from Austria.
Interviewer: Right. One of them has started a political career in the meantime. Which one?
Sabine: Schwarzenegger. He is governor of California.

Interviewer: Mhm. And what do you think, which one of them has shot more movies?
Sabine: Don't know. Stallone maybe?
Interviewer: Mhm. Yes. Who do you think is the better actor?
Sabine: Well, then this muscleman, because at least once I watched a real good movie with Stallone. I just can't remember the title.

B Interview topics in Experiment 1, 2 and 3

Below is a list of all topics that were presented in Experiment 1 except for the demo trials. Topics marked with an [a] were excluded for Experiment 2. Topics marked with a [b] were selected for presentation in Experiment 3. The referent that was established on the left side of gesture space is mentioned first. Each topic's critical response is noted up until the verbal disambiguation, where the ERPs were taken. Both referents are provided as verbal disambiguation, because there were two versions of each topic – one where the interviewee refers to the left referent, one where she refers to the right referent. Metatopics are seperated by an empty line. Original language was German, the English translation is written in italics. Note, that due to different word order, the English translation gets sometimes disambiguated before the arrival of the designated verbal disambiguation.

Meta topic 1

Topic 1:

Homöopathie – Schulmedizin

Im Endeffekt glaube ich, dass die Behandlungsmethode verträglicher ist, denn das Gute an der Homöopathie/Schulmedizin...

homeopathy – conventional medicine

In the long run, I think that this treatment is better tolerated, because the good thing about homeopathy/conventional medicine...

Topic 2:

Private Krankenversicherung – Gesetzliche Krankenversicherung

Dann würde ich die Versicherung wählen – für mich ist der entscheidende Vorteil, dass die Privaten/Gesetzlichen...

private health insurance – compulsory health insurance

Under these circumstances I would choose this insurance – for me, the most important advantage is that the private health insurance/compulsory health insurance...

Topic 3:

Tattoo – Piercing

Da kann ich mich doch eher mit dem Körperschmuck anfreunden, denn der Vorteil daran ist ja, dass man für so ein Tattoo/Piercing...

tattoo – piercing

I guess I could rather live with this kind of body modification, because its advantage is, that for such a tattoo/piercing...

Topic 4^b:

Nikotin – Cannabis

Ich glaube, da ist die Droge schlimmer, denn das wirklich blöde bei Nikotin/Cannabis...

nicotine – cannabis

I think this drug is worse, because the bad thing about nicotine/cannabis...

Topic 5:

Pfeife – Zigarre

Keine Frage, ich mag den Rauch lieber. Wahrscheinlich bin ich da geprägt, weil mein Vater Pfeife/Zigarre...

pipe – cigar

No doubts, I like this smoke more. Probably I am shaped in this regard by my father, because he used to smoke the pipe/cigars.

Meta topic 2

Topic 6:

Kunstfaserschlafsack – Daunenschlafsack

Wahrscheinlich würde ich den Schlafsack mitnehmen. Ich würde also auf den Vorzug setzen, dass der Kunstfaserschlafsack/Daunenschlafsack...

synthetic fill sleeping bag – down fill sleeping bag

Probably, I would take this sleeping bag. So, I would put the emphasis on the advantage that the synthetic fill sleeping bag/down fill sleeping bag...

Topic 7:

Gaskocher – Benzinkocher

Nein, ich habe nur den Kocher. Ich find's nämlich viel besser, dass Du mit dem Gaskocher/Benzinkocher...

gas stove – liquid fuel stove

No, I just have this type of stove. I just like it way better that it's possible with the gas stove/liquid fuel stove...

Topic 8[b]:

Digitalkamera – Analogkamera

Ich bin ein Freund von den Kameras, alles in allem ist es meines Erachtens viel angenehmer mit einer Digitalkamera/Analogkamera…

digital camera – film camera

I'm a friend of this camera type, all in all it's in my opinion much more pleasant to take pictures with a digital camera/film camera.

Topic 9:

Fotos – Videos

Ich glaube, die Variante ist beliebter, ist doch schön, dass so ein Foto/Video…

photograph – video

I believe that this alternative is more popular. I mean, isn't it nice, that such a photograph/video…

Topic 10[b]:

Fernseher – Buch Na, dann lieber die Variante – ab aufs Sofa, unter die Decke kuscheln, 'n Fernseher/Buch…

TV set – book Well, under these cirumstances I would prefer this alternative – on the couch, under the blanket, turning on the TV/taking a book…

Meta topic 3

Topic 11:

Studium – Lehre

Ja. Der will mit der Ausbildung anfangen, aber im Moment genießt er noch die freie Zeit, die er bis dahin hat. Er hat schon gesagt, dass er das Wort "Studium"/"Lehre"…

studies – apprenticeship

Yes. He wants to start this training, but at the moment he enjoys the spare time he's got left. He already said that he doesn't want to hear the word "studies"/"apprenticeship"…

Topic 12:

Diplom – Bachelor

Na ja, zu dem Abschluss kann man noch sagen, dass der Nachteil am Diplom/Bachelor…

German university degree – bachelor degree

Well, concerning this degree you could add that the disadvantage of the Diplom/ Bachelor…

Topic 13:
Uni – FH
Ich nehme an, an der Hochschule gibt es dann mehr Studenten. Die werden halt alle davon angelockt, dass das Studium an der Uni/FH...
university – university of applied sciences
I think that there are more students going to this kind of institution. I guess they are attracted by the fact, that the studies at the university/university of applied sciences...
Topic 14[a]:
Einzimmerwohnung – Wohngemeinschaft
Ich tippe auf die Wohnform. Zumindest habe ich als Studentin länger in der Einzimmerwohnung/Wohngemeinschaft...
studio apartment – flat share
I have to guess, but I take this type of housing. At least when I was a student, I used to live for a longer period in a studio apartment/flat share.
Topic 15:
Brockhaus – Wikipedia
Q: Vielleicht das Lexikon... Doch, ich bin ganz zuversichtlich und denke, dass Brockhaus/Wikipedia...
German encyclopedia – Wikipedia
Perhaps this encyclopedia... Yes, I am quite confident and think that Brockhaus/Wikipedia...

Meta topic 4
Topic 16:
Fahrrad – Auto
Ich bin häufiger damit unterwegs. Du weißt ja, wo ich wohne, und da ist es in den meisten Situationen mit dem Fahrrad/Auto...
bicycle – car
Mostly I use that. I mean, you know where I live and over there it usually much more practical to use the bicycle/car...
Topic 17:
Bahn – Flugzeug
Ganz klar die Variante; da ist es doch viel schöner per Bahn/Flugzeug...
train – plane
Without a doubt this alternative; isn't it much nicer to travel by train/plane...
Topic 18:
Transrapid – ICE
Eigentlich der Zug, denn ich sehe den Vorteil vom Transrapid/ICE...

German maglev – German conventional train
In fact, I think this train, because the advantage of the Transrapid/ICE…

Meta topic 5
Topic 19:
Goethe – Shakespeare
Da würde der Klassiker gewinnen, denn so etwas schönes wie bei Goethes/Shakespeares…
an author – an author
Then this classic would win, because I've rarely read something as beautiful as Goethe's/Shakespeare's…
Topic 20:
Hesse – Kafka
An dem Schriftsteller find' ich mehr Gefallen – zugegeben, man muss sich erstmal an den Stil von Hesse/Kafka…
an author – an author
I favor this author – admittedly, you have to get used to Hesse's/Kafka's…
Topic 21:
Micky – Donald
Die Figur vielleicht. Ich glaub', ich hab' mal gelesen, dass es die höhere Auflage bei Mickey/Donald…
a cartoon character – a cartoon character
Perhaps this character. I think I've read once that the higher print runs come with Mickey/Donald…
Topic 22[b]:
Asterix – Lucky Luke
Ich glaube, dass es die Serie heute schwerer hat, denn leider merkt man es den neuen Asterix-Alben/Lucky-Luke-Alben…
a comic book character – a comic book character
I think, the present situation is tougher for this title, because unfortunately you notice when reading the latest issues of Asterix/Lucky Luke…
Topic 23[b]:
Stern – Spiegel
Häufiger das Magazin. Mir gefällt es eben besser, dass der Spiegel/Stern…
German news magazine – German news magazine
More often this magazine. I simply prefer that Spiegel/Stern…

Topic 24:
FAZ – Bild
Gute Frage… Ich meine, die Zeitung, denn man muss ganz wertneutral anerkennen, dass die FAZ/Bild…
German newspaper – German newspaper
Good question… I guess this newspaper, because you simply have to acknowledge that FAZ/Bild…
Topic 25:
Private Fernsehsender – Öffentlich-rechtliche Fernsehsender
Ich denke, die Sender, denn meines Erachtens spricht für die Öffentlich-rechtlichen/Privaten…
commercial broadcasting – public service broadcasting
I assume this type of broadcasting, because in my opinion a crucial aspect in favor of commercial broadcasting/public service broadcasting…

Meta topic 6
Topic 26:
Land – Stadt
Mir persönlich hat es die Umgebung mehr angetan; ich bin also eher der Landmensch/Stadtmensch…
countryside – city
Personally, I prefer this kind of surrounding; so I am rather the country/city type…
Topic 27:
Eigenheim – Mietwohnung
Ich ganz persönlich glaube, dass die Wohnform die angenehmere ist, denn – wie gesagt – das schöne an Eigenheimen/Mietwohnungen…
owner-occupied house – rented apartment
Personally, I believe that this kind of living is more enjoyable, because – as I mentioned before – the beauty about living in an owner-occupied house/a rented apartment…
Topic 28[a]:
Neubau – Altbau
Unser Haus ist so ein Bau, der hat auch – wie die meisten Neubauten/Altbauten –…
new building – old building
We live in such a house, which has – as most new/old buildings –…

Topic 29[b]:

Dachgeschoss – Erdgeschoss

Klar, ich kenn' sogar beides. Länger habe ich allerdings in dem Geschoss gewohnt, weswegen ich auch ziemlich gut mit allen Vor- und Nachteilen des Dachgeschosses/Erdgeschosses…

attic floor – first floor

Of course, I even know both variants. However, I have lived for a longer period on this kind of floor. That's the reason why I am pretty familiar with all the advantages and disadvantages of the attic/first floor.

Meta topic 7

Topic 30:

Joggen – Schwimmen

Also, wenn du's in einem vernünftigem Rahmen betreibst, ist beides gesund. Ich persönlich würde mich aber trotzdem immer für den Sport entscheiden, ich find's nämlich gut, dass du beim Joggen/Schwimmen…

jogging – swimming

Well, when practicing on a sound level, both sports are healthy. Personally, I would always choose this sport, as I think it's a nice aspect of jogging/swimming…

Topic 31:

Freibad – See

Ich dreh' lieber in dem Wasser meine Runden, denn lieber bin ich im Freibad/See…

open air pool – lake

I prefer to swim in this water, because I'm rather tormented in the open air pool/lake by…

Topic 32:

Ski – Snowboard

Früher war ich in dem Sport aktiv, aber das ist jetzt schon ein Weilchen her, dass wir skifahren/snowboarden…

ski – snowboard

Back in the days, I used to practice this sport, but it's been quite a while that we went skiing/snowboarding.

Topic 33:

Maske – Beckenbauer

Was ist denn das für eine Frage? Ich tippe auf den Sportler… oder doch der andere… Nein, ich bleibe dabei, dass der Maske/Beckenbauer…

German boxer – German football player
What kind of question is that? Maybe this athlete… or the other one… No, I stay with my first choice and think that Maske/Beckenbauer…

Meta topic 8
Topic 34:
Grillen – Raclette
Ich finde die Variante irgendwie netter, weil das tolle am Grillen/Raclette…
barbecue – raclette
I find this variant nicer, because the cool thing about the barbecue/raclette…
Topic 35:
McDonald's – Subway
Die bevorzugen die Fast-Food-Kette. Aber bei mir im Bekanntenkreis scheint das ein allgemeines Phänomen zu sein, die mögen alle McDonald's/Subway…
a fast food restaurant – a fast food restaurant
They prefer this fast food chain. Actually, this seems to be a more general phenomenon regarding the peoply I know, because they all prefer McDonald's/Subway.
Topic 36:
Pizza – Döner
Na, das war das Gericht… oder lass mich überlegen… doch, da hab' ich Pizza/Döner…
pizza – doner kebab
Well, that would be this dish… wait, let me reconsider… yes, it was pizza/doner kebab…
Topic 37[a]:
Vegetarier – Veganer
Ja, ich esse nach den Regeln, lebe also streng nach den Grundsätzen des Vegetarismus/Veganismus, …
vegetarian – vegan
Yes, I eat according to these rules, so I strictly adhere to the principles of vegetarianism/veganism, …
Topic 38:
Rewe – Aldi
Ich denke, mehr gehen zu dem Supermarkt. Die werden hauptsächlich den Vorteil sehen, dass man bei Rewe/Aldi…

German supermarket – German supermarket
I guess that more people go to this supermarket. Most of them will focus on the Rewe's/Aldi's advantage...

Topic 39:
Standardgemüse – Biogemüse
Ich hole meistens das Gemüse, manchmal auch das andere, aber meistens wird es doch das Standardgemüse/Biogemüse, ...
standard vegetables – organic vegetables
I buy mostly this kind of vegetables, sometimes also the other one, but mostly I stick to standard vegetables/organic vegetables, ...

Topic 40:
Pfandflasche – Tetra Pak
Ich tippe mal auf die Verpackung, ich glaube nämlich, dass schon irrsinnige Mengen von Pfandflaschen/Tetra Paks...
deposit bottle – Tetra Pak
I would guess this kind of packaging, as I think that already a ridiculuous amount of deposit bottles/Tetra Paks...

Meta topic 9

Topic 41:
Osten – Westen
Ganz ursprünglich kommen wir aus der Region und das ist auch heute noch so – ich zum Beispiel bin auch im Osten/Westen...
Eastern Germany – Western Germany
Originally, we come from this region and it's still the same today – I, for example, am also born in Eastern Germany/Western Germany.

Topic 42:
Dresden – Leipzig
Na, vielleicht die Stadt? Ich hoffe, die Zuschauer wissen das auch nicht besser und sind entsprechend mit meiner Entscheidung für Dresden/Leipzig...
German city – German city
Well, this city maybe? I hope, the people in the audience don't know better and are in consequence fine with my vote for Dresden/Leipzig.

Topic 43:
Thomaskirche – Nikolaikirche
Da mag ich die Kirche mehr; mir persönlich gefällt es halt, dass sie es bei der Thomaskirche/Nikolaikirche...

church in Leipzig – church in Leipzig
In this case, I prefer that church; I simply like the fact that they designed the Thomaskirche/Nikolaikirche...

Topic 44:
Protestantismus – Katholizismus
Ich denke, da überwiegt die Konfession, oder? Ja, ich leg' mich jetzt mal auf den Protestantismus/Katholizismus...
Protestantism – Catholicism
I guess that this church predominates, right? Yes, I stick to Protestantism/Catholicism...

Topic 45:
Hinduismus – Islam
Also, das ist ja jetzt übertrieben. Und wenn ich mich überhaupt ein bisschen auskenne, dann in der Religion. Das kommt daher, dass in meinem Freundeskreis einige dem Hinduismus/Islam...
Hinduism – Islam
Well, I think you are exaggerating. And if I know a tiny bit, then it is about this religion. The reason for this is that some of my friends belong to Hinduism/Islam.

Meta topic 10

Topic 46:
Tagesmutter – Kinderkrippe
Den habe ich dort untergebracht, aber das ist ja nicht nur bei ihm so, bei mir sind ja alle Kinder zur Tagesmutter/Kinderkrippe...
family child care – day nursery
He's going there, but this is not only for him the case, all my children have been going to family child care/day nursery.

Topic 47:
Ganztagsschule – Halbtagsschule
Ich finde die Schule besser. Also, lass es mich präzisieren: Ich finde die Ganztagsschule/Halbtagsschule...
all-day school – half-day school
I think this school type is better. Well, let be more precise: I think that the all-day school/half-day school...

Topic 48[b]:
Französisch – Latein
Ich würde ihm die Sprache empfehlen. In meinen Augen ist der große Vorteil von Französisch/Latein...

French – Latin
I would recommend that language to him. In my eyes, the big advantage of French/Latin…

Topic 49:
Wehrdienst – Zivildienst
Ich tippe auf den Dienst. Verbürgen kann ich mich dafür nicht, aber ich habe vor Jahren mal 'ne Statistik dazu gelesen und da hatte der Wehrdienst/Zivildienst…
military service – civilian service
I guess they opt for this service. I can't guarantee that this is the right answer, but a couple of years ago I saw a statistic, which showed that the military service/civilian service…

Meta topic 11

Topic 50:
Teppich – Parkett
Oh, ich glaube, dass die Antwort gar nicht so einfach ist, aber ich denke, dass der Bodenbelag schlechter geeignet ist, denn das blöde beim Teppich/Parkett…
carpeted floor – parquet floor
Oh, I assume it's not that easy to find the correct answer, but I think that this floor is less suited, because the bad thing about carpeted floor/parquet floor…

Topic 51[b]:
Fichte – Eiche
Da wähle ich das Holz. Passt auch besser zum Rest, ich hab' nämlich schon einiges aus Fichte/Eiche…
spruce – oak
Then I choose this kind of wood. Fits better with the rest, as I already have quite a lot made out of spruce/oak.

Topic 52:
Gasherd – Elektroherd
Definitiv mit dem Herd. Deswegen war ich auch ganz schön froh, dass in unserer Küche schon ein Gasherd/E-Herd…
gas kitchen stove – electric kitchen stove
Definitively with this stove. That's also why I was quite happy, that there was already a gas kitchen stove/electric kitchen stove…

Topic 53[a]:
Handy – Festnetz
Eher auf das Telefon, so schlimm stell' ich mir das Leben ohne Han-

dy/Festnetz...
cell phone – landline phone
Rather without this phone, I don't think that a live without cell phone/landline phone...

Topic 54:
Hund – Katze
Da würde ich mir das Tier anschaffen und das ist tatsächlich schon geschehen, denn wir wohnen in der Stadt und haben zwei Hunde/Katzen...
dog – cat
Then I would get this pet and that's what actually has already happened, because we live in the city and we have two dogs/cats...

Topic 55:
Terrarium – Aquarium
Oh, hm, vielleicht das Ding? Ehrlich gesagt, weiß ich's nicht, aber 'n Freund von mir, der hat so 'n Terrarium/Aquarium...
terrarium – aquarium
Oh, um, maybe this thing? Truth to be told, I don't know, but a friend of mine has such a terrarium/aquarium...

Meta topic 12

Topic 56:
Notebook – Desktop
Wir haben so'n Computer und meines Erachtens ist es auch viel praktischer mit so 'nem Notebook/Desktop...
notebook – desktop computer
We've got such a computer and in my opinion life is way more practical with such a notebook/desktop computer...

Topic 57:
CRT – LCD
Wenn er bei der Arbeit ist nutzt er lieber die Monitore. Gut, wie schon erwähnt, die blöde Eigenschaft bei CRTs/LCDs...
CRT – LCD
When being at work her prefers to use these monitors. Admittedly and as mentioned the disadvantage of CRTs/LCDs...

Topic 58[a]:
PC – Mac
Dem würd' ich den Computer empfehlen, denn das große Plus beim PC/Mac...

PC – Mac
I would recommend him this computer, because the big advantage of PCs/ Macs...

Topic 59:

Windows – Linux
Wenn Du mich so fragst, ist mir das System sympathischer. Ich find's einfach besser, dass du mit Windows/Linux...
an operating system – an operating system
Under these circumstances, I pick this system. I just find it better that you have less compatibility problems with Windows/you have a free and user developed system with Linux.

Topic 60:

Word – LaTeX
Meistens das Programm. Das ist bei mir aber keine Glaubenssache, in meinem speziellen Bereich ist es nur naheliegender Word/LaTeX...
a word processor – a word processor
Most of the times this program. This isn't a matter of faith, however, in my specific area of work it's just more plausible to use Word/LaTeX.

Meta topic 13

Topic 61:

Rasierer – Wachs
Die meisten Frauen werden die Variante benutzen. Ich persönlich finde es auch praktischer mit dem Rasierer/Wachs...
razor – wax
Most women will use this variant. Me, personally, I also find it more practical to work with a razor/with wax.

Topic 62:

Nassrasur – Trockenrasur
Keine Ahnung, so oft frag' ich die Männer nicht danach. Vielleicht nutzt der Großteil die Methode. Immerhin ist der Vorteil bei der Nassrasur/ Trockenrasur...
wet shaving – electric shaving
I have no clue, I mean that's usually not what I ask the men I meet. Perhaps the majority uses this method. After all, the advantage of wet shaving/electric shaving...

Topic 63:

Kondom – Pille
Also, ich glaube, da hat man es bei der Verhütungsmethode schwerer,

also bei der Verhütung per Kondom/Pille, ...
condom – birth-control pill
Well, I think that you can run into more difficulties with this method of contraception, so with contraception via condom/birth-control pill, ...

Meta topic 14
Topic 64:
Wein – Bier
Mir sagt das Getränk mehr zu – ein schöner Abend mit Freunden und dazu etwas Wein/Bier...
wine – beer
I tend to prefer this beverage – a nice evening with friends, drinking some wine/beer...
Topic 65:
Wodka – Tequila
Na, wenn's sein muss, dann der Drink, aber dann kann man trotzdem nicht davon sprechen, dass ich Wodka/Tequila...
vodka – tequila
Well, if it's really necessary, then this drink, but that doesn't mean at all that I truly enjoy drinking Wodka/Tequila.
Topic 66[b]:
Whisky – Rum
Also, da muss ich gestehen, dass ich von dem Getränk schon ein paar Flaschen gekauft habe. Vor ein paar Jahren hab' ich Whisky/Rum...
whisky – rum
Well, I have to admit, that I've already bought a couple of bottles of this drink. Some years ago I got to know whisky/rum...
Topic 67:
Kaffee – Tee
Hm, ja, aber lieber noch das Getränk, schmeckt mir einfach besser. Dann tu' ich noch ein bisschen Milch in den Kaffee/Tee...
coffee – tea
Um, yeah, but I prefer this beverage, simply tastes better in my opinion. Then I add a pinch of milk to my coffee/tea...

Meta topic 15
Topic 68[b]:
Italien – Frankreich
Ich finde beides schön, aber wenn ich mich entscheiden soll, dann

würde ich das Land nehmen. Allein wegen der guten Küche könnt' ich immer wieder nach Italien/Frankreich…
Italy – France
I like them both, but if I have to make a choice, I pick this country. Just because of its good cuisine I could travel over and over again to Italy/France.

Topic 69:
Norwegen – Spanien
Leider nicht, ich kenne nur das Land von innen. Das kenn' ich dafür aber ziemlich gut, mit der Familie waren wir nämlich schon etliche Male in Norwegen/Spanien, …
Norway – Spain
Sadly no, I just know this country from the inside. And I know it truly well, as we have already been a couple of times to Norway/Spain, …

Topic 70:
Berge – Meer
Ich vermute, dass die meisten in die Region fahren. Immerhin verbindet man mit dem schönen Wort "Berge"/"Meer"…
the mountains – the sea
I assume that most people visit this region. After all, when hearing the beautiful word "mountains"/"sea"…

Topic 71:
Pauschalreise – Individualreise
Da werden mein Mann und ich die Reiseform wählen, weil wir finden's einfach angenehmer, dass Du bei der Pauschalreise/Individualreise…
package holiday – individually organized holiday
My husband and me, we will go with this way of travelling, because we simply find it more pleasant that the package holiday/individually organized holiday…

Topic 72:
Koffer – Rucksack
Für meine Belange ist das Gepächstück definitiv praktischer. Für mich macht es einfach mehr Sinn mit einem Koffer/Rucksack…
suitcase – backpack
Considering my needs this piece of luggage comes definitively more handy. In my case it just makes more sense to travel with a suitcase/backpack, …

Meta topic 16

Topic 73[b]:

Schwarzenegger – Stallone

Also, wenn dann den Muskelprotz, denn zumindest einmal hab' ich 'n wirklich guten Film mit Schwarzenegger/Stallone...

an actor – an actor

Well, then this muscleman, because at least once I watched a real good movie with Schwarzenegger/Stallone. ...

Topic 74:

Hitchcock – Spielberg

Ich freu' mich immer wieder, wenn von dem Regisseur 'n Film im Fernseher kommt. Allein, was Hitchcock/Spielberg...

a director – a director

I feel always happy, when they show a movie by this director on TV. I mean, just recall what Hitchcock/Spielberg...

Topic 75:

Theater – Kino

Ich glaube die Variante, also, ich weiß es ja nicht wirklich, aber ich glaube das Besondere am Theater/Kino...

theater – cinema

I think this variant, well, as a matter of fact I don't know for sure, but I think the exceptional thing about theater/cinema...

Topic 76[a]:

Platte – CD

Ich nutze lieber das Medium. Für mich ist der entscheidende Vorzug, dass die Platte/CD...

vinyl record – compact disc

I prefer to use this medium. In my opinion, the crucial advantage of the vinyl record/compact disk...

Topic 77:

Raumschiff Enterprise – Krieg der Sterne

Ganz klar die Science-Fiction-Saga. Komm', gib's zu, du bist doch auch der größere Fan von Raumschiff Enterprise/Krieg der Sterne...

Star Trek – Star Wars

Clearly this science fiction saga. Come on, admit it, you are also a bigger fan of Star Trek/Star Wars...

Topic 78:

Freud – Einstein

Na gut, da wähl' ich mal den Wissenschaftler, das war – glaub' ich –

ziemlich einflussreich, was Freud/Einstein...
famous psychoanalyst – famous physicist
Well, I think I pick this scientist. I believe that it was really influential what Freud/Einstein...

Meta topic 17

Topic 79:
Füller – Kugelschreiber
Nein, mit dem Schreibgerät kann ich mich nicht so anfreunden, der Nachteil am Füller/Kuli...
fountain pen – ball pen
No, I cannot get used with this writing utensil, I mean isn't it the big disadvantageof the fountain pen/ball pen...

Topic 80:
Digitaluhr – Analoguhr
Ich glaube mehr solche Uhren, wobei ein Großteil davon wahrscheinlich aus den Digitaluhren/Analoguhren...
digital watch – analogue watch
I guess more of these watches. The biggest part of that are probably those digital watches/analogue watches...

Topic 81:
Kontaktlinsen – Brille
Nein, aber wenn ich eine bräuchte, dann würde ich mir die Sehhilfe anschaffen, ich find's nämlich viel besser, dass Du mit Kontaktlinsen/Brille
contact lenses – glasses
No, but if I needed one, then I would buy this optical aid, becuase I find it way better, that contact lenses/glasses...

Topic 82:
Schirm – Jacke
Am liebsten ist mir da der Regenschutz, es ist meines Erachtens nach viel praktischer mit Schirm/Jacke...
umbrella – jacket
I prefer this kind of rain cover. In my opinion it is way more practical to use an umbrella/jacket.

Topic 83:
Wollpullover – Fleecepullover
Ich kaufe meistens die Pullover, schließlich ist der angenehme Vorteil von Wollpullovern/Fleecepullovern...

wool sweater – fleece sweater
Most of the times I buy this kind of sweater. After all, the nice advantage of wool sweater/fleece sweaters...

Topic 84:
Rock – Hose
Meistens das Kleidungsstück, da ist sie also ganz die Mama, wenn möglich laufe ich nämlich auch mit Rock/Hose...
skirt – trousers
Most of the times this kind of garment. So, in this case she's just like her mother, because if possible I also wear a skirt/trousers.

Meta topic 18

Topic 85:
Neuwagen – Gebrauchtwagen
Vermutlich so'n Wagen, für uns spricht wie gesagt dafür, dass 'n Neuwagen/Gebrauchtwagen...
new car – used car
Proably such a car, as I already said it is an advantage in our eyes that a new car/used car...

Topic 86:
Diesel – Benziner
Wir haben den Wagentyp. Das hat den Grund, dass unser Auto als weniger störanfällig gilt in der Benzinervariante...
diesel fuel – gasoline
We have this kind of car. The reason is that our car is said to be less susceptible to break-downs in the diesel fuel version/gasoline version...

Topic 87:
Erdgasauto – Hybrid
Also, zu dem Auto könntest du mich nicht überreden, weil's mich doch sehr nervt, dass man beim Erdgauto/Hybridauto...
natural gas car – hybrid car
Well, you could talk me into this car, because I'm actually quite annoyed by the fact that the natural gas car/hybrid car...

Topic 88[b]:
Mini – Käfer
Da gewinnt der Wagen, das ist keine Frage für mich. Das beste am Mini/Käfer...
a car – a car
In this case this car wins, this is no question for me. The best part about the

Mini/Käfer…

Topic 89:

Kombi – SUV

Bei den Autos. Es entlockt mir wirklich keine Begeisterung, wenn ich so'n Kombi/SUV…

station wagon – sports utility vehicle

With these cars. I really don't feel any excitement, when seeing such a station wagon/sports utility vehicle…

Topic 90:

Motorrad – Auto

Mir ist prinzipiell das Fahrzeug sympathischer, schließlich fühlt man sich mit einem Motorrad/Auto…

motorcycle – car

As a matter of fact I like these vehicles more, after all with a motorcycle/car…

C Interview topics in Experiment 4

The following list contains all interview topics that were used for Experiment 4 except for the demo trials. The referent that was established on the left side of gesture space is always mentioned first. Aside from the referents, the question asked to the participants is also noted with the left referent variant being mentioned first again. Text in italics represents a translation. Meta topics are separated by an empty line.

Meta topic 1
Topic 1:
Homöopathie – Schulmedizin
Wer verschreibt Bachblüten/Betablocker?
homeopathy – conventional medicine
Who prescribes Bach flower remedies/beta blockers?
Topic 2:
Private Krankenversicherung – Gesetzliche Krankenversicherung
Bei welcher Krankenversicherung ist wohl ein Vielverdiener/Geringverdiener?
private health insurance – compulsory health insurance
To which insurance will a high-income earner/low-income earner go?
Topic 3:
Tattoo – Piercing
Was hält Sabine für verruchter/weniger verrucht?
tattoo – piercing
What thinks Sabine to be more/less disreputable?
Topic 4:
Nikotin – Cannabis
Welche Droge ist in Deutschland legal/illegal?
nicotine – cannabis
Which drug is in Germany legal/illegal?
Topic 5:
Pfeife – Zigarre
Wo sind laut Sabine die Schwierigkeiten für Anfänger größer/kleiner?

pipe – cigar
Where do beginners encounter more/less problems

Meta topic 2
Topic 6:
Kunstfaserschlafsack – Daunenschlafsack
Welchen Schlafsack verbindet Sabine nicht mit dem Wort "down"/Sabine mit dem Wort "down"?
synthetic fill sleeping bag – down fill sleeping bag
Which sleeping bag does Sabine not connect with the word "down"/Sabine connect with the word "down"?
Topic 7:
Gaskocher – Benzinkocher
Welcher Kochertyp versagt laut Sabine manchmal/nie?
gas stove – liquid fuel stove
Which type of stove fails according to Sabine sometimes/never?
Topic 8:
Digitalkamera – Analogkamera
Bei welcher Kamera braucht man eine Speicherkarte/einen Film?
digital camera – film camera
For which camera do you need a memory card/a film?
Topic 9:
Fotos – Videos
Die Nachbearbeitung wovon ist laut Sabine einfacher/aufwändiger?
photograph – video
The post-production of what medium is according to Sabine easier/more extensive?
Topic 10:
Fernseher – Buch
Wozu gehört ein Moderator/Autor?
TV set – book
To what would you associate a show host/an author?

Meta topic 3
Topic 11:
Studium – Lehre
Womit wird man ein Arzt/Frisör?
studies – apprenticeship
What do you need to become a medical doctor/a coiffeur?

Topic 12:
Diplom – Bachelor
Bei welchem Abschluss ist laut Sabine das Niveau höher/niedriger?
German university degree – bachelor degree
Which degree stands for a higher/lower level?
Topic 13:
Brockhaus – Wikipedia
Was findet man im Bücherregal/Internet?
German encyclopedia – Wikipedia
What can you find on a book shelve/the internet?

Meta topic 4
Topic 14:
Fahrrad – Auto
Was hat eine Klingel/einen Motor?
bicycle – car
What has a bell/motor?
Topic 15:
Bahn – Flugzeug
Welches Verkehrsmittel hat Waggons/Flügel?
train – plane
What means of transport has coaches/wings?
Topic 16:
Transrapid – ICE
Welcher Zug ist laut Sabine schneller/langsamer?
German maglev – German conventional train
Which train is according to Sabine faster/slower?

Meta topic 5
Topic 17:
Goethe – Shakespeare
Welcher Autor schrieb "Faust"/"Macbeth"?
an author – an author
Which author wrote "Faust"/"Macbeth"
Topic 18:
Hesse – Kafka
Welcher Schriftsteller lebte laut Sabine in Deutschland/Prag?
an author – an author
Which author lived according to Sabine in Germany/Prague?

Topic 19:
Micky – Donald
Welche Figur erschien laut Sabine früher/später?
a cartoon character – a cartoon character
Which character was presented according to Sabine earlier/later?
Topic 20:
Asterix – Lucky Luke
Welche Comicserie gefällt Sabines Kindern besser/weniger?
a comic book character – a comic book character
Which title do Sabines children like more/less?
Topic 21:
Stern – Spiegel
Beide Magazine berichten über Politik – wo ist laut Sabine der Anteil geringer/größer?
German news magazine – German news magazine
Both magazines report about politics – where's the share smaller/bigger according to Sabine?
Topic 22:
FAZ – Bild
Das Image welcher Zeitung ist besser/schlechter?
German newspaper – German newspaper
What magazine's image is better/worse?

Meta topic 6
Topic 23:
Land – Stadt
Wo findet man eher eine Kuh/Kino?
countryside – city
Where do you rather find a cow/cinema?
Topic 24:
Eigenheim – Mietwohnung
Was ist in der Regel größer/kleiner?
owner-occupied house – rented apartment
What is usually bigger/smaller?
Topic 25:
Dachgeschoss – Erdgeschoss
Welche Wohnung liegt im Haus oben/unten?
attic floor – first floor
Which apartment is located at the top/bottom of a building?

Meta topic 7

Topic 26:

Joggen – Schwimmen

Welche Sportart hält Sabine für beliebter/unbeliebter?

jogging – swimming

What sport does Sabine like more/less?

Topic 27:

Ski – Snowboard

In welcher Sportart hat man laut Sabine typischerweise Verletzungen am Knie/Handgelenk?

ski – snowboard

Doing what type of sport you would typically end up with a knee/wrist injury?

Topic 28:

Maske – Beckenbauer

In wessen Sportart gibt es den Ausdruck "K. o."/"Freistoß"?

German boxer – German football player

Which sport knows the term "KO"/"free kick"?

Meta topic 8

Topic 29:

Grillen – Raclette

Die Säuberung wovon ist laut Sabine schwieriger/leichter?

barbecue – raclette

What is more difficult/easier to clean according to Sabine?

Topic 30:

McDonald's – Subway

Was ist wohl bekannter/unbekannter?

a fast food restaurant – a fast food restaurant

What is probably more/less famous?

Topic 31:

Rewe – Aldi

Wo findet man laut Sabine Markenprodukte häufiger/seltener?

German supermarket – German supermarket

Where do you find more/less branded products?

Meta topic 9
Topic 32:
Osten – Westen
In welchem Teil Deutschlands liegt Chemnitz/Stuttgart?
Eastern Germany – Western Germany
In what part of Germany is Chemnitz/Stuttgart?
Topic 33:
Dresden – Leipzig
Welche Stadt vergleicht man laut Sabine gerne mit Florenz/Paris?
German city – German city
Which city is according to Sabine often compared with Florence/Paris?
Topic 34:
Thomaskirche – Nikolaikirche
Welche Kirche hält Sabine für jünger/älter?
church in Leipzig – church in Leipzig
Which church is younger/older in Sabine's opinion?
Topic 35:
Protestantismus – Katholizismus
Welche Kirche ist laut Sabine stärker vertreten im Norden/Süden?
Protestantism – Catholicism
Which of both churches is more strongly represented in the North/South according to Sabine?
Topic 36:
Hinduismus – Islam
Die Medienpräsenz welcher Religion ist momentan kleiner/größer?
Hinduism – Islam
Which religion receives less/more media coverage at the moment?

Meta topic 10
Topic 37:
Tagesmutter – Kinderkrippe
Wo ist laut Sabine das Zeitmanagement einfacher/schwieriger?
family child care – day nursery
Where is the time management easier/more difficult according to Sabine?
Topic 38:
Französisch – Latein
Welche Sprache verbindet man mit der Stadt Paris/Rom?
French – Latin
What language would you associate with the city of Paris/Rome?

Topic 39:
Wehrdienst – Zivildienst
Welchen Dienst verbindet man mit dem Wort "Waffe"/"Krankenhaus"?
military service – civilian service
Which service do you associate with the word "weapon"/"hospital"?
Topic 40:
Familie – Altersheim
Stell dir vor, die betreute Person ruft um Hilfe – bei welcher Betreuungsform kommt wohl die Tochter/Pflegerin?
family – retirement home
Imagine, that the cared for person calls for help – where will the daughter/a nurse show up?

Meta topic 11
Topic 41:
Teppich – Parkett
Welcher Bodenbelag besteht oft aus Wolle/besteht aus Holz?
carpeted floor – parquet floor
What kind of floor is often made out of wool/is made out of wood?
Topic 42:
Fichte – Eiche
Welche Baumart gibt es laut Sabine häufiger/seltener?
spruce – oak
Whas kind of tree is more frequent/rare?
Topic 43:
Handy – Festnetz
Womit ist laut Sabine das Telefonieren teurer/billiger?
cell phone – landline phone
Where's making calls more/less expensive?
Topic 44:
Hund – Katze
Welches Tier bellt/miaut?
dog – cat
What animal does bark/meow?
Topic 45:
Terrarium – Aquarium
Was gibt es laut Sabine in Deutschland seltener/häufiger?
terrarium – aquarium
What is according to Sabine more rare/frequent in Germany?

Meta topic 12
Topic 46:
Notebook – Desktop
Welche Form des Computers ist leichter/schwerer?
notebook – desktop computer
What kind of computer is lighter/heavier?
Topic 47:
Windows – Linux
Welches Betriebssystem ist wohl bekannter/unbekannter?
an operating system – an operating system
Which operating system is probably better/less known?
Topic 48:
Word – LaTeX
Was benutzt man laut Sabine eher zum Schreiben eines Briefes/Buchs?
a word processor – a word processor
What do you rather use for the purpose of writing a letter/book according to Sabine?

Meta topic 13
Topic 49:
Rasierer – Wachs
Bei welcher Methode sind laut Sabine die Schmerzen geringer/größer?
razor – wax
With which method will the pain be smaller/greater according to Sabine?
Topic 50:
Nassrasur – Trockenrasur
Für welche Rasurmethode braucht man Wasser/Strom?
wet shaving – electric shaving
For which shaving method do you need water/electricity?
Topic 51:
Kondom – Pille
Welche Verhütungsmethode wird angewandt von Männern/Frauen?
condom – birth-control pill
What birth control technique is applied by the man/woman?

Meta topic 14
Topic 52:
Wein – Bier
Womit verbindet Sabine das Image “gepflegt”/“prollig”?

wine – beer
Which drink does Sabine associate with the image "decent"/"chavvy"?
Topic 53:
Wodka – Tequila
Für welches Getränk nennt Sabine das Herkunftsland Russland/Mexiko?
vodka – tequila
For what drink names Sabine Russia/Mexico as the country of origin?
Topic 54:
Kaffee – Tee
Worum handelt es sich bei einem Mokka/Darjeeling?
coffee – tea
What are we dealing with when talking about a Turkish coffee/Darjeeling?

Meta topic 15
Topic 55:
Italien – Frankreich
Bei welchem Land nennt Sabine die beliebte Region Toscana/Provence?
Italy – France
For what country names Sabine the popular region Toscana/Provence?
Topic 56:
Norwegen – Spanien
Die Hauptstadt welches Landes ist Oslo/Madrid?
Norway – Spain
To which country belongs the capital Oslo/Madrid?
Topic 57:
Berge – Meer
Was ist der Himalaya/Pazifik?
the mountains – the sea
What is the Himalaya/Pacific?
Topic 58:
Pauschalreise – Individualreise
Wo gibt es laut Sabine die Aussicht auf Preisvorteile/Geheimtipps?
package holiday – individually organized holiday
Where do you have the outlook for a price advantage/an insider tip?

Meta topic 16

Topic 59:

Schwarzenegger – Stallone

Beide Schauspieler haben laut Sabine Verbindungen nach Europa – welcher Darsteller nach Österreich/Italien?

an actor – an actor

Both actors have connections to Europe according to Sabine – who has one to Austria/Italy?

Topic 60:

Hitchcock – Spielberg

Welcher Regisseur drehte den Film "Psycho"/"E.T."?

a director – a director

Who directed the movie "Psycho"/"E.T."?

Topic 61:

Theater – Kino

Sabine bekommt Karten geschenkt – wo ist ihre Freude größer/kleiner?

theater – cinema

Sabine gets some tickets as a present – what would she enjoy more/less?

Topic 62:

Platte – CD

Bei welchem Medium braucht man zum Abspielen eine Nadel/einen Laser?

vinyl record - compact disc

For what medium do you need a needle/laser in order to play it?

Topic 63:

Beatles – Rolling Stones

Welche der Bands charakterisiert Sabine als liebenswert/verrucht?

a music band – a music band

Which group does Sabine characerize as likable/disreputable?

Topic 64:

Freud – Einstein

Wer wirkte laut Sabine früher/ist Sabine sympathischer?

famous psychoanalyst – famous physicist

Who made his contributions earlier according to Sabine/Whom does Sabine like more?

Meta topic 17
Topic 65:
Füller – Kugelschreiber
Welches Schreibgerät ist das traditioneller/modernere?
fountain pen – ball pen
Which of the pens is the more traditional/modern one?
Topic 66:
Digitaluhr – Analoguhr
Welche Uhren sind laut Sabine für gewöhnlich billiger/teurer?
digital watch – analogue watch
What kind of watch is usually less/more expensive?
Topic 67:
Schirm – Jacke
Welchen Gegenstand muss man aufspannen/anziehen?
umbrella – jacket
What do you have to open up/put on?
Topic 68:
Wollpullover – Fleecepullover
Welchen Pulli findet Sabine bei schlechtem Wetter schlechter/besser?
wool sweater – fleece sweater
What sweater does Sabine find less/better suited for bad weather?

Meta topic 18
Topic 69:
Diesel – Benziner
Eine poetische Frage an dich: Was reimt sich auf das Wort "Wiesel"/"Mediziner"?
diesel fuel – gasoline
A rather poetic question to you: What rhymes on "weasel"/"physician"?
Topic 70:
Mini – Käfer
Bei welchem Wagen ist die Anzahl an Kratzern wohl kleiner/größer?
a car – a car
Where will you probably find less/more scratches
Topic 71:
Kombi – SUV
Welches Auto empfiehlt Sabine einem Bassisten/Förster?
station wagon – sports utility vehicle
Which car does Sabine recommend to a bassist/ranger?

Topic 72:
Motorrad – Auto
Bei welchem Fahrzeug liegt die Räderanzahl bei zwei/vier?
motorcycle – car
What means of transport has two/four wheels?

Bibliographische Darstellung

J. E. Douglas Weinbrenner
Abstract Pointing. ERP and behavioral evidence for its role in reference tracking
Fakultät für Biowissenschaften, Pharmazie und Psychologie
Universität Leipzig
Dissertation
247 Seiten, 275 Literaturangaben, 22 Abbildungen, 8 Tabellen

The application of the gesture type *abstract pointing* for the task of *reference tracking* was explored. Assumedly, a speaker can use a pointing gesture to empty space to virtually position a discourse referent. For instance, when talking about literature, the speaker could point to the left when mentioning Goethe and to the right when mentioning Shakespeare. Later, a further pointing to the left could indicate on its own that the speaker refers to Goethe. Four experiments were carried out in order to explore this idea. In general, participants watched a video of an interview about dualistic topics like *Goethe and Shakespeare*. In each topic, the interviewee established a gesturing order (e.g. *Goethe – left and Shakespeare – right*), which was followed by the experimental manipulation. Effects were measured via event-related potentials or behaviorally. Experiment 1 showed that recipients process the gestural referent information, but this did not seem beneficial in moments of underspecified speech. Experiment 2 demonstrated that abstract pointing can be beneficial when being a reliable referent indicator. A main finding of Experiment 3 was that some recipients readily utilize the information from abstract pointing while others do not. Experiment 4 indicates that recipients inherit the gesturing order to such a degree that it affects their own contributions to communication. In sum, there is strong support for the idea that abstract pointing can be used for reference tracking.

Summary

Introduction

Co-speech gestures are those hand and arm movements we produce during everyday discourse. Sometimes it is easy to determine their communicative function. An utterance like "The keys are over there", for instance, would be useless without the accompanying pointing gesture that depicts the location of the desired object. In contrast, imagine a conversation on classic literature, where the speaker says that she enjoys reading Goethe and Shakespeare and conducts a pointing to the left on the word "Goethe" and one to the right on the word "Shakespeare". This pointing into the void is termed *abstract pointing* and its function is by no means clear. An intriguing speculation is that it could be used for *spatial reference tracking*, where recipients infer the referent from gesture location (McNeill, 2003). For example, based on the previously applied gesturing order a further pointing to the left during later discourse could signal to the recipient that the speaker talks once more about Goethe.

Spatial reference tracking is commonly used in sign language, but its application in spoken language has not been fully explored yet. In a pilot study preceding this dissertation, it was demonstrated that the recipient can indeed associate words with the locations of abstract pointing gestures (Gunter et al., 2015). Several questions, however, remained. For instance, it was not clear whether such an association would suffice to make the recipient pick a referent when speech is ambiguous. Formal as well as interactional characteristics of abstract pointing were also still unknown. The present dissertation was aimed at further exploring abstract pointing and starting to look into these matters.

The setup for all experiments was roughly identical. In particular, an interview in video format was created, which was supposed to look as natural as possible, while allowing for as much control as necessary in order to run true experiments with systematic and focussed manipulations of the independent variable. The interview contained several dualistic topics like *Goethe and Shakespeare* or *dogs and cats*. By conducting abstract pointings in a spatially consistent manner, the female interviewee estab-

lished for every topic a gesturing order, e.g. *Goethe – left and Shakespeare – right*. After two to four establishing gestures per side the critical phase of an interview topic arrived, where the experimental manipulation took place.

Experiment 1

The guiding question for Experiment 1 was whether abstract pointing can be used by the recipient to infer the referent when speech lacks clarity. To this end, the critical phase presented the participants with a verbal statement that was initially ambiguous and only later on explicit. An example is given in (1), which represents the interviewee's response to the question for her preferred author.

(1) Then this classic$_{[—/\text{Goethe}/\text{Shakespeare}]}$ would win, because I've rarely read something as beautiful as *Goethe's* Faust.

As indicated by subscript in (1), there was a baseline condition, where no gesture information was available for the participants. In a congruent condition, abstract pointing indicated at "this classic" the same referent as the later verbal disambiguation. In an incongruent condition, the opposing referent was indicated. The ERP data taken at the verbal disambiguation "Goethe's" showed that participants are sensitive to the incongruent condition, where abstract pointing is misleading about the referent. Specifically, there was a more positive P600 to this condition, which suggests that the cognitive system required more processing resources when dealing with a conflict about the discourse referent. In light of this finding, it was remarkable that there was no difference between the congruent and the baseline condition. Taken together, this pattern of results would suggest that incongruent referent information from abstract pointing can impede with language comprehension, whereas congruent referent information cannot facilitate it. A potential explanation could be that abstract pointing was not a reliable cue. In 50% of the cases, where ambiguous speech was presented together with an abstract pointing, the gesture was misleading, which might have led to unconventional processing.

Experiment 2

Consequently, abstract pointing was turned into a reliable cue in Experiment 2, which was accomplished by removing the incongruent condition

from the experimental design. After this modification, the ERPs taken at "Goethe's" depicted clear differences for the baseline as compared to the congruent condition. First, there was a more negative N400 indicating that participants had to initiate a lexical access regarding the referent. Second, there was a more positive P600 reflecting the process of implementing the referent into the mental representation of what is being communicated. Both processes had apparently already been carried out in the congruent condition based on the gesture information. Thus, when abstract pointing is reliable, it can render language comprehension easier.

The comparison of Experiment 1 and 2 is a challenging task. Obviously, the reduced reliability of abstract pointing regarding its referent indicating capacity in Experiment 1 did change how the gestures were processed. Several accounts for the exact nature of this altered processing are thinkable. It might be that fewer participants utilized abstract pointing to infer the referent or that they did so in fewer trials. Alternatively, participants processed abstract pointing maybe on a more shallow level, where congruent abstract pointing showed no observable impact, but incongruent abstract pointing did. This – so to say – preference for incongruent gestures may be rooted in the fact that a potential error in language comprehension, as indicated by a referent conflict, has to be treated with a higher priority than a situation, where everything runs smoothly as in the congruent condition. Finally, it could also be that the P600 for the incongruent condition reflects a monitoring process, where the reliability status of abstract pointing is updated based on the occurrences of incongruent gesturing. Just in general, reliability updating seems to be a mandatory step for all accounts, because without it, there would have been no modification of the processing mode between the experiments.

Irrespective of the nature of this altered processing, it remains that abstract pointing was continuously processed, even when repeatedly failing to be informative about the referent. The reason for this affinity could be that the recipient expects a communicative value from abstract pointing. As demonstrated with Experiment 2, the value could be the disambiguation of verbally ambiguous situations. Additionally, it might be that abstract pointing facilitates language comprehension even when it is completely redundant with speech. This would put it close to the phenomenon of code-blending, where people knowing a sign and a spoken language simultaneously sign and utter equivalent semantic content. First evidence indicates that such code-blends are beneficial for compre-

hension (Emmorey et al., 2012) and the same might be true for spatial reference tracking via abstract pointing.

Experiment 3

With Experiment 3, the spotlight was moved to the appearance of abstract pointing, in particular to whether spatial reference tracking is also possible with artificial stimuli that are not encountered in natural communication. The short answer is yes. Participants were presented with the interview topics until and including the verbal ambiguity together with abstract pointing, but excluding the verbal disambiguation. The speech stream of the videos was kept the same, but half of the participants did not see the original video stream, but a black screen with dots popping up on the left or right side as a replacement for abstract pointing. When explicitly asked for the discourse referent of the verbal ambiguity, accuracy was in essence the same for the gesture and the dot group. Hence, it appears the recipient is quite flexible regarding what kind of stimuli are accepted for spatial reference tracking. It is also notable that those participants, who engaged in spatial reference tracking, acquired the associations between spatial cues and referents quickly.

An unexpected finding was that, regardless of stimulus type, half of the participants did not engage in spatial reference tracking, i.e. they were not able to tell the momentary discourse referent based on the spatial cue. Three reasons could play a role in my opinion. First, it could be that the cues did not appear important enough during the critical phase. Specifically in case of the gestures, this might have to do with the fact that their amplitude was too small. Just in general, it might have to do with the fact that there was no variation in the appearance of the spatial cue, so that the critical cue did not signal in any way that it is important and should be regarded with sufficient attention. Second, the high amount of non-detectors might be rooted in the fact that the experimental design lacked interaction. Specifically, participants were not given feedback about their referent hypotheses. This way, evaluation of their task strategies was difficult for them potentially leading to random performance. Third, interindividual differences could contribute to the high amount of non-detectors of spatial reference tracking. They could be located in cognitive skills such as the ability to divide attention between the auditory (speech) and the visual (gesture) modality or in personality traits such as extraversion.

Experiment 4

The fourth experiment was intended to focus on interactional aspects, more specifically to see whether a recipient would still be influenced by abstract pointing and its gesturing order when becoming the producer during a conversation. This is indeed the case. In a Simon-like task, participants had to answer a dual-choice question at the end of each establishing phase with the main referents of the present interview topic representing the response alternatives; responses had to be indicated by lifting the left or right index. Crucially, the interviewee's gesturing order and the alignment of the response alternatives on the screen were pitted against each other. When both spatial orders did not correspond, participants' reaction time and accuracy decreased significantly. Thus, although the gesturing order was never of relevance for the task, participants reactivated it when they had to indicate a response. This is first evidence for the idea that people could have a tendency to continue on referent indication by the means of space that was started by the interlocutor.

Conclusion

Recipients are capable of inferring the referent from the location of abstract pointing and this can be beneficial for the communicational process. While it is not clear under what exact circumstances people engage in spatial reference tracking, they show flexibility regarding the type of spatial stimulus they accept and they can quickly learn the associations between spatial stimulus and speech. Furthermore, people observing a gesturing order show a tendency to align with this order beyond the utterance of the interactant, i.e. when it is their turn to respond. In sum, there is ample evidence that abstract pointing, these pointings into the void, can impact communication by their location in gesture space.

Zusammenfassung

Einführung

Als sprachbegleitende Gesten bezeichnet man solche Bewegungen der Hände und Arme, die wir während alltäglicher Gespräche ausführen. Ihre Funktion ist teils einfach zu bestimmen. Die Aussage: „Die Schlüssel sind dort drüben“, wäre zum Beispiel nutzlos ohne eine entsprechende Zeigegeste. Aber wenn eine Sprecherin anmerkt, sie lese sowohl Goethe als auch Shakespeare und dabei eine Zeigegeste nach links auf das Wort „Goethe“ ausführt und eine nach rechts auf „Shakespeare“, dann ist die Funktion nicht offensichtlich. Diese Zeigesten ins Leere werden als *abstrakte Zeigegesten* bezeichnet und es wird vermutet, dass sie für *räumliche Referenten-Nachverfolgung* genutzt werden (McNeill, 2003). Im späteren Gesprächsverlauf könnte eine weitere Zeigegeste nach links dem Rezipienten beispielsweise angeben, dass sich die Sprecherin erneut auf Goethe bezieht.

Räumliche Referenten-Nachverfolgung ist aus Gebärdensprachen bekannt, aber die Anwendung in Lautsprachen wurde bisher nicht ausführlich untersucht. In einer dieser Dissertation vorausgehenden Pilotstudie konnte nachgewiesen werden, dass Rezipienten tatsächlich Wörter mit den Orten von abstrakten Zeigegesten assoziieren (Gunter et al., 2015). Es war aber beispielsweise offen, ob solche Assoziationen ausreichen, um einen Referenten zu wählen, wenn die Sprache ambigue ist. Außerdem blieben formale und interaktionale Charakteristika dieses Gestentyps weiterhin ungeklärt. Ziel dieser Dissertation war es, dergleichen Aspekte weiter zu erforschen.

Das experimentelle Design aller Experimente folgte dem gleichen Muster. Die Versuchsteilnehmer sahen auf Video ein Interview, das einerseits vergleichsweise natürlich gestaltet war, andererseits aber genügend experimentelle Kontrolle zuließ, um Experimente mit systematischen Manipulationen der unabhängigen Variablen durchzuführen. In dem Interview wurden mehrere dualistische Themen wie *Goethe und Shakespeare* oder *Hunde und Katzen* behandelt. Durch konsistente Nutzung des Gestenraumes etablierte die interviewte Person für jedes Thema eine Gestenord-

nung wie *Goethe – links und Shakespeare – rechts*. Nach zwei bis vier etablierenden Gesten pro Referent erreichte das Thema die kritische Phase der jeweiligen experimentellen Manipulation.

Experiment 1

Die leitende Frage für Experiment 1 war, ob abstrakte Zeigegesten zur Identifikation des Referenten genutzt werden können, wenn die Sprache uneindeutig ist. In der kritischen Phase wurden die Versuchsteilnehmer mit einer Aussage konfrontiert, die zunächst ambigue war und erst später aufgelöst wurde. Ein Beispiel ist in (1) gegeben, wo die Interviewte auf die Frage nach dem von ihr präferierten Autor antwortet.

(1) Dann würde der Klassiker$_{[—/\text{Goethe}/\text{Shakespeare}}$ gewinnen, denn so etwas schönes wie bei *Goethes* Faust habe ich selten gelesen.

Wie dem tiefergestellten Text zu entnehmen ist, gab es eine Baseline-Bedingung, in der den Probanden keine Gesteninformation zur Verfügung stand. Daneben gab es eine kongruente Bedingung, in der die abstrakte Zeigeste auf „der Klassiker" denselben Referenten angab wie die spätere verbale Auflösung „Goethes". In einer inkongruenten Bedingung wurde der entgegengesetzte Referent durch die Geste angegeben. Ereigniskorrelierte Potentiale (EKPs), die auf den Beginn des Wortes „Goethes" gemessen wurden, ergaben eine abweichende Reaktion der Probanden in der inkongruenten Bedingung. Speziell war eine positivere P600-Komponente zu beobachten, was auf einen erhöhten Ressourcenverbrauch des kognitiven Systems im Falle eines Referentenkonflikts deutet. In Anbetracht dieses Ergebnisses war es allerdings bemerkenswert, dass kein Unterschied zwischen der kongruenten und der Baseline-Bedingung nachgewiesen wurde. Dies würde bedeuten, dass abstrakte Zeigegesten mit inkongruenter Referentinformation die Kommunikation erschweren können, wohingegen solche mit kongruenter Referenteninformation sie nicht erleichtern können. Eine mögliche Erklärung könnte sein, dass die abstrakten Zeigegesten in Experiment 1 keinen zuverlässigen Hinweisreiz darstellten. Konkret waren die Gesten in 50% der Fälle sprachlicher Ambiguität irreführend, was zu einer ungewöhnlichen Verarbeitung der Gesten geführt haben könnte.

Experiment 2

Für Experiment 2 wurde die inkongruente Bedingung entfernt, so dass die Gesten reliabel waren. Nun zeigten die EKPs auf das Wort „Goethes“ klare Abweichungen der Baseline- von der kongruenten Bedingung. Einerseits war eine negativere N400 zu beobachten, was darauf deutet, dass die Probanden einen lexikalischen Zugriff für den Referenten durchzuführen hatten. Andererseits zeigte sich die P600 positiver, was darauf schließen lässt, dass sie den Referenten in ihre mentale Repräsentation dessen, was kommuniziert werden soll, einbauen mussten. Beide Prozesse waren in der kongruenten Bedingung demzufolge bereits mit Hilfe der Gesteninformation durchgeführt worden. Wenn abstrakte Zeigegesten zuverlässig sind, können sie die Kommunikation also erleichtern.

Ein offensichtlicher Schluss aus dem Vergleich von Experiment 1 und 2 ist, dass die Reliabilität von abstrakten Zeigegesten einen Einfluss auf ihre Verarbeitung hatte, wobei verschiedene Szenarien denkbar sind. Eine Möglichkeit ist, dass durch die Unzuverlässigkeit in Experiment 1 weniger Probanden die Gesten zur Referenten-Nachverfolgung genutzt haben oder dass sie es in weniger Trials gemacht haben, so dass keine messbaren Effekte mehr zu beobachten waren. Es könnte auch sein, dass die Zeigegesten oberflächlicher prozessiert wurden, wobei nur noch inkongruente Gesteninformationen einen beobachtbaren Einfluss hatten. Die spezielle Beachtung von inkongruenten Gesten könnte dabei daher rühren, dass ein Referentenkonflikt vom kognitiven System mit höherer Priorität behandelt werden muss als eine kongruente, also unauffällige Situation. Letztlich könnte es auch sein, dass der P600-Effekt für die inkongruente Bedingung in Experiment 1 schlicht einen Überwachungsprozess zur Reliabilitätseinschätzung von abstrakten Zeigegesten darstellt. Solch einen Überwachungsprozess muss es grundsätzlich gegeben haben, da es sonst nie zu einer Änderung in der Gestenprozessierung gekommen wäre.

In jedem Fall ist festzuhalten, dass abstrakte Zeigegesten auch dann kontinuierlich prozessiert wurden, wenn sie wiederholt den falschen Referenten anzeigten. Der Grund dafür könnte sein, dass sich das kognitive System letztlich einen Vorteil davon erhofft. Dieser könnte wie in Experiment 2 gezeigt in der Auflösung verbal ambiguer Situationen liegen. Es könnte aber auch sein, dass abstrakte Zeigegesten die Kommunikation sogar dann unterstützen, wenn die Gesteninformation mit der Sprachinformation redundant ist. Dies würde abstrakte Zeigegesten in die Nähe von sogenanntem Code-Blending rücken: Menschen, die sowohl eine

Gebärden- als auch eine Lautsprache beherrschen, können den gleichen semantischen Inhalt simultan über Gebärden und Sprache ausdrücken. Erste Ergebnisse weisen darauf hin, dass Code-Blends die Sprachverarbeitung erleichtern können (Emmorey et al., 2012) und dasselbe könnte auch auf räumliche Referenten-Nachverfolgung per abstrakten Zeigegesten zutreffen.

Experiment 3

In Experiment 3 wurde der Fokus auf das Erscheinbungsbild der Gesten gesetzt und zwar darauf, ob räumliche Referenten-Nachverfolgung auch mit Hilfe von artifiziellen Stimuli möglich ist. Die kurze Antwort lautet: Ja. Eine Hälfte der Probanden sah das ursprüngliche Interview, wobei jedes Thema kurz vor der verbalen Auflösung gestoppt wurde. Die andere Hälfte wurde prinzipiell mit dem gleichen Stimulusmaterial konfrontiert, jedoch war die Videospur durch ein schwarzes Bild ersetzt, auf dem immer dann klar wahrnehmbare Punkte links oder rechts erschienen, wenn im Original eine abstrakte Zeigegeste zu sehen war. Befragt nach dem Referenten der verbalen Ambiguität stellte sich bei den Probanden der Anteil korrekter Antworten unabhängig von der Art des Hinweisreizes dar. Rezipienten sind also flexibel und akzeptieren verschiedene Stimuli zur räumlichen Referenten-Nachverfolgung. Daneben war zu beobachten, dass diejenigen Probanden, die räumliche Referenten-Nachverfolgung betrieben, die Assoziationen zwischen den Hinweisreizen und den Referenten rasch lernten.

Überraschenderweise machte circa die Hälfte der Versuchspersonen unabhängig von der Art des Hinweisreizes keinen Gebrauch von räumlicher Referenten-Nachverfolgung. Mindestens drei Ursachen sind denkbar. Erstens, dass die Hinweisreize bei der verbalen Ambiguität nicht wichtig genug erschienen. Speziell im Fall der Gesten könnte dies auf ihr geringes Bewegungsausmaß zurückzuführen sein und allgemein darauf, dass innerhalb der Gesten beziehungsweise innerhalb der Punkte keine bedeutsame Variation zwischen den einzelnen Reizen vorkam. Im Endeffekt erschien der kritische Reiz eventuell nicht relevant genug, weil er sich von den anderen nicht klar absetzte. Zweitens könnte mangelnde Interaktion für den hohen Anteil an Nicht-Nutzern verantwortlich sein. Beispielsweise erhielten die Probanden keine Rückmeldung, ob ihre Antwort korrekt ist. So war eine Evaluation ihres Vorgehens schwierig, was letztlich zu einem Ergebnis auf Zufallsniveau geführt haben könnte. Drittens könn-

ten interindividuelle Differenzen beeinflusst haben, ob jemand räumliche Referenten-Nachverfolgung nutzt. Kognitive Fähigkeiten wie die zur Aufmerksamkeitsteilung zwischen auditorischer und visueller Modalität könnten hier eine Rolle spielen oder auch Persönlichkeitseigenschaften wie Extraversion.

Experiment 4

In Experiment 4 sollten interaktionale Aspekte abstrakter Zeigegesten erforscht werden, speziell ob sich ein Rezipient weiterhin von der Gestenordnung beeinflusst zeigen würde, wenn er selber einen Beitrag zur Kommunikation leisten soll. Dies ist tatsächlich der Fall. In einer Aufgabe, die dem klassischen Simon-Paradigma ähnelt, mussten die Versuchspersonen eine Zweifachwahl-Aufgabe am Ende jeder Etablierungsphase bearbeiten. die Referenten des jeweiligen Interviewthemas stellten die beiden Antwortalternativen dar und die Antwort wurde durch Heben des linken oder rechten Zeigefingers gegeben. Entscheidend war, dass dabei die von der Interviewten etablierte Gestenordnung der Anordnung der Antwortalternativen auf dem Bildschirm gegenübergestellt wurde. Wenn die beiden Anordnungen nicht übereinstimmten, dann zeigten die Versuchspersonen eine erhöhte Fehlerzahl und eine erhöhte Reaktionszeit. Die Versuchspersonen reaktivierten also die Gestenordnung, wenn sie selber eine Antwort geben sollten, obwohl diese Ordnung während des Experiments niemals von Bedeutung war. Dies deutet auf eine fragliche Tendenz der Rezipienten, vom Interaktionspartner begonnenes abstraktes Zeigen fortzuführen.

Schlussfolgerung

Rezipienten können abstrakte Zeigegesten zum Erschließen des Referenten nutzen, was von zusätzlichem Vorteil sein kann. Es ist unklar, unter welchen genauen Umständen sie mit räumlicher Referenten-Nachverfolgung beginnen, aber sie sind flexibel bezüglich der Art des räumlichen Hinweisreizes und sie können die nötigen Assoziationen schnell lernen. Außerdem zeigen sie sich über die Äußerung des Gesprächspartners hinaus bei ihren eigenen Kommunikationsbeiträgen von der beobachteten Gestenordnung beeinflusst. Insgesamt sprechen die Ergebnisse dafür, dass abstrakte Zeigegesten über ihre Lokation einen Einfluss auf Kommunikation haben.

Curriculum Vitae

Personal Details

Name	J. E. Douglas Weinbrenner
Date of birth	March 9, 1978
Place of birth	Cologne

Education

since 2012	Postgraduate education in clinical neuropsychology Society for Neuropsychology, Germany
1999 - 2008	Diploma in psychology University of Würzburg
2003 - 2004	Studies of French Centre d'Approches Vivantes des Langues et des Médias, Vichy
2002 - 2003	International exchange student Behavioral Neuroscience Program, University of Texas at Austin
1998	University entrance diploma Dossenberger-Gymnasium, Günzburg

Professional Experience

since 2012	Clinical neuropsychologist Neurological Center Bad Segeberg
2008 - 2012	PhD student Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig

Publikationen und Vorträge

Publikationen

Gunter, T. C. & Weinbrenner, J. E. D. (2016). *When to take a gesture seriously: on how we use and prioritize communicative cues.* Manuscript submitted for publication.

Gunter, T. C., Weinbrenner, J. E. D., & Holle, H. (2015). Inconsistent use of gesture space during abstract pointing impairs language comprehension. *Frontiers in Psychology, 6,* 80.

Vorträge

Weinbrenner, J. E. D. (2012, Oct). *Investigating the comprehension and production of abstract pointing gestures with ERPs.* Talk presented at Donders Discussions 2012, Nijmegen.

Weinbrenner, J. E. D. (2011, Feb). *Behavioral and ERP experiments on how the consistent usage of gesture space is processed.* Talk presented at The Nijmegen Gesture Centre Lecture Series, Nijmegen.

Weinbrenner, J. E. D. (2010, Jul). *Pointing into the void – ERP evidence on how abstract pointing can affect speech processing.* Talk presented at 4th Conference of the International Society for Gesture Studies, Frankfurt/Oder.

Weinbrenner, J. E. D. (2010, Jul). *Pointing into the void – how abstract pointing can affect speech processing.* Talk presented at 2nd Workshop on the Neuroscience of Gesturing, Leipzig.

Selbstständigkeitserklärung

Ich versichere, dass ich die vorliegende Arbeit ohne unzulässige Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe und dass ich die aus fremden Quellen direkt oder indirekt übernommenen Gedanken als solche in der Arbeit kenntlich gemacht habe.

Ort, Datum	(Douglas Weinbrenner)

MPI Series in Human Cognitive and Brain Sciences:

1 Anja Hahne
Charakteristika syntaktischer und semantischer Prozesse bei der auditiven Sprachverarbeitung: Evidenz aus ereigniskorrelierten Potentialstudien

2 Ricarda Schubotz
Erinnern kurzer Zeitdauern: Behaviorale und neurophysiologische Korrelate einer Arbeitsgedächtnisfunktion

3 Volker Bosch
Das Halten von Information im Arbeitsgedächtnis: Dissoziationen langsamer corticaler Potentiale

4 Jorge Jovicich
An investigation of the use of Gradient- and Spin-Echo (GRASE) imaging for functional MRI of the human brain

5 Rosemary C. Dymond
Spatial Specificity and Temporal Accuracy in Functional Magnetic Resonance Investigations

6 Stefan Zysset
Eine experimentalpsychologische Studie zu Gedächtnisabrufprozessen unter Verwendung der funktionellen Magnetresonanztomographie

7 Ulrich Hartmann
Ein mechanisches Finite-Elemente-Modell des menschlichen Kopfes

8 Bertram Opitz
Funktionelle Neuroanatomie der Verarbeitung einfacher und komplexer akustischer Reize: Integration haemodynamischer und elektrophysiologischer Maße

9 Gisela Müller-Plath
Formale Modellierung visueller Suchstrategien mit Anwendungen bei der Lokalisation von Hirnfunktionen und in der Diagnostik von Aufmerksamkeitsstörungen

10 Thomas Jacobsen
Characteristics of processing morphological structural and inherent case in language comprehension

11 Stefan Kölsch
Brain and Music
A contribution to the investigation of central auditory processing with a new electrophysiological approach

12 Stefan Frisch
Verb-Argument-Struktur, Kasus und thematische Interpretation beim Sprachverstehen

13 Markus Ullsperger
The role of retrieval inhibition in directed forgetting – an event-related brain potential analysis

14 Martin Koch
Measurement of the Self-Diffusion Tensor of Water in the Human Brain

15 Axel Hutt
Methoden zur Untersuchung der Dynamik raumzeitlicher Signale

16 Frithjof Kruggel
Detektion und Quantifizierung von Hirnaktivität mit der funktionellen Magnetresonanztomographie

17 Anja Dove
Lokalisierung an internen Kontrollprozessen beteiligter Hirngebiete mithilfe des Aufgabenwechselparadigmas und der ereigniskorrelierten funktionellen Magnetresonanztomographie

18 Karsten Steinhauer
Hirnphysiologische Korrelate prosodischer Satzverarbeitung bei gesprochener und geschriebener Sprache

19 Silke Urban
Verbinformationen im Satzverstehen

20 Katja Werheid
Implizites Sequenzlernen bei Morbus Parkinson

21 Doreen Nessler
Is it Memory or Illusion? Electrophysiological Characteristics of True and False Recognition

22 Christoph Herrmann
Die Bedeutung von 40-Hz-Oszillationen für kognitive Prozesse

23 Christian Fiebach
Working Memory and Syntax during Sentence Processing.
A neurocognitive investigation with event-related brain potentials and functional magnetic resonance imaging

24 Grit Hein
Lokalisation von Doppelaufgabendefiziten bei gesunden älteren Personen und neurologischen Patienten

25 Monica de Filippis
Die visuelle Verarbeitung unbeachteter Wörter. Ein elektrophysiologischer Ansatz

26 Ulrich Müller
Die katecholaminerge Modulation präfrontaler kognitiver Funktionen beim Menschen

27 Kristina Uhl
Kontrollfunktion des Arbeitsgedächtnisses über interferierende Information

28 Ina Bornkessel
The Argument Dependency Model: A Neurocognitive Approach to Incremental Interpretation

29 Sonja Lattner
Neurophysiologische Untersuchungen zur auditorischen Verarbeitung von Stimminformationen

30 Christin Grünewald
Die Rolle motorischer Schemata bei der Objektrepräsentation: Untersuchungen mit funktioneller Magnetresonanztomographie

31 Annett Schirmer
Emotional Speech Perception: Electrophysiological Insights into the Processing of Emotional Prosody and Word Valence in Men and Women

32 André J. Szameitat
Die Funktionalität des lateral-präfrontalen Cortex für die Verarbeitung von Doppelaufgaben

33 Susanne Wagner
Verbales Arbeitsgedächtnis und die Verarbeitung ambiger Wörter in Wort- und Satzkontexten

34 Sophie Manthey
Hirn und Handlung: Untersuchung der Handlungsrepräsentation im ventralen prämotorischen Cortex mit Hilfe der funktionellen Magnet-Resonanz-Tomographie

35 Stefan Heim
Towards a Common Neural Network Model of Language Production and Comprehension: fMRI Evidence for the Processing of Phonological and Syntactic Information in Single Words

36 Claudia Friedrich
Prosody and spoken word recognition: Behavioral and ERP correlates

37 Ulrike Lex
Sprachlateralisierung bei Rechts- und Linkshändern mit funktioneller Magnetresonanztomographie

38 Thomas Arnold
Computergestützte Befundung klinischer Elektroenzephalogramme

39 Carsten H. Wolters
Influence of Tissue Conductivity Inhomogeneity and Anisotropy on EEG/MEG based Source Localization in the Human Brain

40 Ansgar Hantsch
Fisch oder Karpfen? Lexikale Aktivierung von Benennungsalternative bei der Objektbenennung

41 Peggy Bungert
Zentralnervöse Verarbeitung akustischer Informationen Signalidentifikation, Signallateralisation und zeitgebundene Informationsverarbeitung bei Patienten mit erworbenen Hirnschädigungen

42 Daniel Senkowski
Neuronal correlates of selective attention: An investigation of electrophysiological brain responses in the EEG and MEG

43 Gert Wollny
Analysis of Changes in Temporal Series of Medical Images

S 1 Markus Ullsperger & Michael Falkenstein
Errors, Conflicts, and the Brain Current Opinions on Performance Monitoring

44 Angelika Wolf
Sprachverstehen mit Cochlea-Implantat: EKP-Studien mit postlingual ertaubten erwachsenen CI-Trägern

45 Kirsten G. Volz
Brain correlates of uncertain decisions: Types and degrees of uncertainty

46 Hagen Huttner
Magnetresonanztomographische Untersuchungen über die anatomische Variabilität des Frontallappens des menschlichen Großhirns

47 Dirk Köster
Morphology and Spoken Word Comprehension: Electrophysiological Investigations of Internal Compound Structure

48 Claudia A. Hruska
Einflüsse kontextueller und prosodischer Informationen in der auditorischen Satzverarbeitung: Untersuchungen mit ereigniskorrelierten Hirnpotentialen

49 Hannes Ruge
Eine Analyse des raum-zeitlichen Musters neuronaler Aktivierung im Aufgabenwechselparadigma zur Untersuchung handlungssteuernder Prozesse

50 Ricarda I. Schubotz
Human premotor cortex: Beyond motor performance

51 Clemens von Zerssen
Bewusstes Erinnern und falsches Wiedererkennen: Eine funktionelle MRT Studie neuroanatomischer Gedächtniskorrelate

52 Christiane Weber
Rhythm is gonna get you.
Electrophysiological markers of rhythmic processing in infants with and without risk for Specific Language Impairment (SLI)

53 Marc Schönwiesner
Functional Mapping of Basic Acoustic Parameters in the Human Central Auditory System

54 Katja Fiehler
Temporospatial characteristics of error correction

55 Britta Stolterfoht
Processing Word Order Variations and Ellipses: The Interplay of Syntax and Information Structure during Sentence Comprehension

56 Claudia Danielmeier
Neuronale Grundlagen der Interferenz zwischen Handlung und visueller Wahrnehmung

57 Margret Hund-Georgiadis
Die Organisation von Sprache und ihre Reorganisation bei ausgewählten, neurologischen Erkrankungen gemessen mit funktioneller Magnetresonanztomographie – Einflüsse von Händigkeit, Läsion, Performanz und Perfusion

58 Jutta L. Mueller
Mechanisms of auditory sentence comprehension in first and second language: An electrophysiological miniature grammar study

59 Franziska Biedermann
Auditorische Diskriminationsleistungen nach unilateralen Läsionen im Di- und Telenzephalon

60 Shirley-Ann Rüschemeyer
The Processing of Lexical Semantic and Syntactic Information in Spoken Sentences: Neuroimaging and Behavioral Studies of Native and Non-Native Speakers

61 Kerstin Leuckefeld
The Development of Argument Processing Mechanisms in German. An Electrophysiological Investigation with School-Aged Children and Adults

62 Axel Christian Kühn
Bestimmung der Lateralisierung von Sprachprozessen unter besondere Berücksichtigung des temporalen Cortex, gemessen mit fMRT

63 Ann Pannekamp
Prosodische Informationsverarbeitung bei normalsprachlichem und deviantem Satzmaterial: Untersuchungen mit ereigniskorrelierten Hirnpotentialen

64 Jan Derrfuß
Functional specialization in the lateral frontal cortex: The role of the inferior frontal junction in cognitive control

65 Andrea Mona Philipp
The cognitive representation of tasks – Exploring the role of response modalities using the task-switching paradigm

66 Ulrike Toepel
Contrastive Topic and Focus Information in Discourse – Prosodic Realisation and Electrophysiological Brain Correlates

67 Karsten Müller
Die Anwendung von Spektral- und Waveletanalyse zur Untersuchung der Dynamik von BOLD-Zeitreihen verschiedener Hirnareale

68 Sonja A.Kotz
The role of the basal ganglia in auditory language processing: Evidence from ERP lesion studies and functional neuroimaging

69 Sonja Rossi
The role of proficiency in syntactic second language processing: Evidence from event-related brain potentials in German and Italian

70 Birte U. Forstmann
Behavioral and neural correlates of endogenous control processes in task switching

71 Silke Paulmann
Electrophysiological Evidence on the Processing of Emotional Prosody: Insights from Healthy and Patient Populations

72 Matthias L. Schroeter
Enlightening the Brain – Optical Imaging in Cognitive Neuroscience

73 Julia Reinholz
Interhemispheric interaction in object- and word-related visual areas

74 Evelyn C. Ferstl
The Functional Neuroanatomy of Text Comprehension

75 Miriam Gade
Aufgabeninhibition als Mechanismus der Konfliktreduktion zwischen Aufgabenrepräsentationen

76 Juliane Hofmann
Phonological, Morphological, and Semantic Aspects of Grammatical Gender Processing in German

77 Petra Augurzky
Attaching Relative Clauses in German – The Role of Implicit and Explicit Prosody in Sentence Processing

78 Uta Wolfensteller
Habituelle und arbiträre sensomotorische Verknüpfungen im lateralen prämotorischen Kortex des Menschen

79 Päivi Sivonen
Event-related brain activation in speech perception: From sensory to cognitive processes

80 Yun Nan
Music phrase structure perception: the neural basis, the effects of acculturation and musical training

81 Katrin Schulze
Neural Correlates of Working Memory for Verbal and Tonal Stimuli in Nonmusicians and Musicians With and Without Absolute Pitch

82 Korinna Eckstein
Interaktion von Syntax und Prosodie beim Sprachverstehen: Untersuchungen anhand ereigniskorrelierter Hirnpotentiale

83 Florian Th. Siebörger
Funktionelle Neuroanatomie des Textverstehens: Kohärenzbildung bei Witzen und anderen ungewöhnlichen Texten

84 Diana Böttger
Aktivität im Gamma-Frequenzbereich des EEG: Einfluss demographischer Faktoren und kognitiver Korrelate

85 Jörg Bahlmann
Neural correlates of the processing of linear and hierarchical artificial grammar rules: Electrophysiological and neuroimaging studies

86 Jan Zwickel
Specific Interference Effects Between Temporally Overlapping Action and Perception

87 Markus Ullsperger
Functional Neuroanatomy of Performance Monitoring: fMRI, ERP, and Patient Studies

88 Susanne Dietrich
Vom Brüllen zum Wort – MRT-Studien zur kognitiven Verarbeitung emotionaler Vokalisationen

89 Maren Schmidt-Kassow
What's Beat got to do with ist? The Influence of Meter on Syntactic Processing: ERP Evidence from Healthy and Patient populations

90 Monika Lück
Die Verarbeitung morphologisch komplexer Wörter bei Kindern im Schulalter: Neurophysiologische Korrelate der Entwicklung

91 Diana P. Szameitat
Perzeption und akustische Eigenschaften von Emotionen in menschlichem Lachen

92 Beate Sabisch
Mechanisms of auditory sentence comprehension in children with specific language impairment and children with developmental dyslexia: A neurophysiological investigation

93 Regine Oberecker
Grammatikverarbeitung im Kindesalter: EKP-Studien zum auditorischen Satzverstehen

94 Şükrü Barış Demiral
Incremental Argument Interpretation in Turkish Sentence Comprehension

95 Henning Holle
The Comprehension of Co-Speech Iconic Gestures: Behavioral, Electrophysiological and Neuroimaging Studies

96 Marcel Braß
Das inferior frontale Kreuzungsareal und seine Rolle bei der kognitiven Kontrolle unseres Verhaltens

97 Anna S. Hasting
Syntax in a blink: Early and automatic processing of syntactic rules as revealed by event-related brain potentials

98 Sebastian Jentschke
Neural Correlates of Processing Syntax in Music and Language – Influences of Development, Musical Training and Language Impairment

99 Amelie Mahlstedt
The Acquisition of Case marking Information as a Cue to Argument Interpretation in German
An Electrophysiological Investigation with Pre-school Children

100 Nikolaus Steinbeis
Investigating the meaning of music using EEG and fMRI

101 Tilmann A. Klein
Learning from errors: Genetic evidence for a central role of dopamine in human performance monitoring

102 Franziska Maria Korb
Die funktionelle Spezialisierung des lateralen präfrontalen Cortex: Untersuchungen mittels funktioneller Magnetresonanztomographie

103 Sonja Fleischhauer
Neuronale Verarbeitung emotionaler Prosodie und Syntax: die Rolle des verbalen Arbeitsgedächtnisses

104 Friederike Sophie Haupt
The component mapping problem: An investigation of grammatical function reanalysis in differing experimental contexts using eventrelated brain potentials

105 Jens Brauer
Functional development and structural maturation in the brain's neural network underlying language comprehension

106 Philipp Kanske
Exploring executive attention in emotion: ERP and fMRI evidence

107 Julia Grieser Painter
Music, meaning, and a semantic space for musical sounds

108 Daniela Sammler
The Neuroanatomical Overlap of Syntax Processing in Music and Language - Evidence from Lesion and Intracranial ERP Studies

109 Norbert Zmyj
Selective Imitation in One-Year-Olds: How a Model's Characteristics Influence Imitation

110 Thomas Fritz
Emotion investigated with music of variable valence – neurophysiology and cultural influence

111 Stefanie Regel
The comprehension of figurative language: Electrophysiological evidence on the processing of irony

112 Miriam Beisert
Transformation Rules in Tool Use

113 Veronika Krieghoff
Neural correlates of Intentional Actions

114 Andreja Bubić
Violation of expectations in sequence processing

115 Claudia Männel
Prosodic processing during language acquisition: Electrophysiological studies on intonational phrase processing

116 Konstanze Albrecht
Brain correlates of cognitive processes underlying intertemporal choice for self and other

117 Katrin Sakreida
Nicht-motorische Funktionen des prämotorischen Kortex: Patientenstudien und funktionelle Bildgebung

118 Susann Wolff
The interplay of free word order and pro-drop in incremental sentence processing: Neurophysiological evidence from Japanese

119 Tim Raettig
The Cortical Infrastructure of Language Processing: Evidence from Functional and Anatomical Neuroimaging

120 Maria Golde
Premotor cortex contributions to abstract and action-related relational processing

121 Daniel S. Margulies
Resting-State Functional Connectivity fMRI: A new approach for assessing functional neuroanatomy in humans with applications to neuroanatomical, developmental and clinical questions

122 Franziska Süß
The interplay between attention and syntactic processes in the adult and developing brain: ERP evidences

123 Stefan Bode
From stimuli to motor responses: Decoding rules and decision mechanisms in the human brain

124 Christiane Diefenbach
Interactions between sentence comprehension and concurrent action: The role of movement effects and timing

125 Moritz M. Daum
Mechanismen der frühkindlichen Entwicklung des Handlungsverständnisses

126 Jürgen Dukart
Contribution of FDG-PET and MRI to improve Understanding, Detection and Differentiation of Dementia

127 Kamal Kumar Choudhary
Incremental Argument Interpretation in a Split Ergative Language: Neurophysiological Evidence from Hindi

128 Peggy Sparenberg
Filling the Gap: Temporal and Motor Aspects of the Mental Simulation of Occluded Actions

129 Luming Wang
The Influence of Animacy and Context on Word Order Processing: Neurophysiological Evidence from Mandarin Chinese

130 Barbara Ettrich
Beeinträchtigung frontomedianer Funktionen bei Schädel-Hirn-Trauma

131 Sandra Dietrich
Coordination of Unimanual Continuous Movements with External Events

132 R. Muralikrishnan
An Electrophysiological Investigation Of Tamil Dative-Subject Constructions

133 Christian Obermeier
Exploring the significance of task, timing and background noise on gesture-speech integration

134 Björn Herrmann
Grammar and perception: Dissociation of early auditory processes in the brain

135 Eugenia Solano-Castiella
In vivo anatomical segmentation of the human amygdala and parcellation of emotional processing

136 Marco Taubert
Plastizität im sensomotorischen System – Lerninduzierte Veränderungen in der Struktur und Funktion des menschlichen Gehirns

137 Patricia Garrido Vásquez
Emotion Processing in Parkinson's Disease: The Role of Motor Symptom Asymmetry

138 Michael Schwartze
Adaptation to temporal structure

139 Christine S. Schipke
Processing Mechanisms of Argument Structure and Case-marking in Child Development: Neural Correlates and Behavioral Evidence

140 Sarah Jessen
Emotion Perception in the Multisensory Brain

141 Jane Neumann
Beyond activation detection: Advancing computational techniques for the analysis of functional MRI data

142 Franziska Knolle
Knowing what's next: The role of the cerebellum in generating predictions

143 Michael Skeide
Syntax and semantics networks in the developing brain

144 Sarah M. E. Gierhan
Brain networks for language
Anatomy and functional roles of neural pathways supporting language comprehension and repetition

145 Lars Meyer
The Working Memory of Argument-Verb Dependencies
Spatiotemporal Brain Dynamics during Sentence Processing

146 Benjamin Stahl
Treatment of Non-Fluent Aphasia through Melody, Rhythm and Formulaic Language

147 Kathrin Rothermich
The rhythm's gonna get you: ERP and fMRI evidence on the interaction of metric and semantic processing

148 Julia Merrill
Song and Speech Perception – Evidence from fMRI, Lesion Studies and Musical Disorder

149 Klaus-Martin Krönke
Learning by Doing?
Gesture-Based Word-Learning and its Neural Correlates in Healthy Volunteers and Patients with Residual Aphasia

150 Lisa Joana Knoll
When the hedgehog kisses the frog
A functional and structural investigation of syntactic processing in the developing brain

151 Nadine Diersch
Action prediction in the aging mind

152 Thomas Dolk
A Referential Coding Account for the Social Simon Effect

153 Mareike Bacha-Trams
Neurotransmitter receptor distribution in Broca's area and the posterior superior temporal gyrus

154 Andrea Michaela Walter
The role of goal representations in action control

155 Anne Keitel
Action perception in development: The role of experience

156 Iris Nikola Knierim
Rules don't come easy: Investigating feedback-based learning of phonotactic rules in language.

157 Jan Schreiber
Plausibility Tracking: A method to evaluate anatomical connectivity and microstructural properties along fiber pathways

158 Katja Macher
Die Beteiligung des Cerebellums am verbalen Arbeitsgedächtnis

159 Julia Erb
The neural dynamics of perceptual adaptation to degraded speech

160 Philipp Kanske
Neural bases of emotional processing in affective disorders

161 David Moreno-Dominguez
Whole-brain cortical parcellation: A hierarchical method based on dMRI tractography

162 Maria Christine van der Steen
Temporal adaptation and anticipation mechanisms in sensorimotor synchronization

163 Antje Strauß
Neural oscillatory dynamics of spoken word recognition

164 Jonas Obleser
The brain dynamics of comprehending degraded speech

165 Corinna E. Bonhage
Memory and Prediction in Sentence Processing

S 2 Tania Singer, Bethany E. Kok, Boris Bornemann, Matthias Bolz, and Christina A. Bochow
The Resource Project
Background, Design, Samples, and Measurements

166 Anna Wilsch
Neural oscillations in auditory working memory

167 Dominique Goltz
Sustained Spatial Attention in Touch: Underlying Brain Areas and Their Interaction

168 Juliane Dinse
A Model-Based Cortical Parcellation Scheme for High-Resolution 7 Tesla MRI Data

169 Gesa Schaadt
Visual, Auditory, and Visual-Auditory Speech Processing in School Children with Writing Difficulties

170 Laura Verga
Learning together or learning alone: Investigating the role of social interaction in second language word learning

171 Eva Maria Quinque
Brain, mood and cognition in hypothyroidism

172 Malte Wöstmann
Neural dynamics of selective attention to speech in noise

173 Charles-Étienne Benoit
Music-based gait rehabilitation in Parkinson's disease

174 Anja Fengler
How the Brain Attunes to Sentence Processing Relating Behavior, Structure, and Function

175 Emiliano Zaccarella
Breaking Down Complexity: The Neural Basis of the Syntactic Merge Mechanism in the Human Brain

S 2 Tania Singer, Bethany E. Kok, Boris Bornemann, Matthias Bolz, and Christina A. Bochow
2nd Edition The Resource Project
Background, Design, Samples, and Measurements

176 Manja Attig
Handlungsverständnis in den ersten Lebensjahren: retrospektive und prospektive Verarbeitung

177 Andrea Reiter
Out of control behaviors?
Investigating mechanisms of behavioral control in alcohol addition, binge eating disorder, and associated risc factors

178 Anna Strotseva-Feinschmidt
The processing of complex syntax in early childhood

179 Smadar Ovadia-Caro
Plasticity following stroke: the recovery of functional networks as measured by resting-state functional connectivity

180 Indra Kraft
Predicting developmental dyslexia at a preliterate age by combining behavioral assessment with structural MRI

181 Sabine Frenzel
How actors become attractors
A neurocognitive investigation of linguistic actorhood

182 Anja Dietrich
Food craving regulation in the brain: the role of weight status and associated personality aspects

183 Haakon G. Engen
On the Endogenous Generation of Emotion

184 Seung-Goo Kim
Myeloarchitecture and Intrinsic Functional Connectivity of Auditory Cortex in Musicians with Absolute Pitch

185 Yaqiong Xiao
Resting-state functional connectivity in the brain and its relation to language development in preschool children

186 Sofie Louise Valk
The Structure of the Social Brain:
Dissociating socio-affective and socio-cognitive networks through the study of individual differences, brain plasticity, and disease models