Indirect request comprehension in different contexts

Cover design: Alex Andrews
ISBN: 978-90-76203-93-5
Printed and bound by Ipskamp Drukkers b.v.

The research reported in this thesis was supported by an IMPRS fellowship funded by the Max Planck Gesellschaft, München, Germany.

Indirect request comprehension in different contexts

Proefschrift
ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 18 juni 2018
om 12.30 uur precies

door
Johanne Tromp
geboren op 23 juni 1989
te Ermelo

Promotoren:

Prof. dr. A.S. Meyer

Prof. dr. P. Hagoort

Manuscriptcommissie:

Prof. dr. I. Toni

dr. R.M. Willems

Prof. dr. M.M. Louwerse (Tilburg University)

Table of Contents

Chapter 1

General introduction

General introduction

Imagine yourself in a restaurant having some food and a glass of wine with your friends. At some point the waiter comes by to ask if you need anything. You see that the bottle of wine on the table is empty and you say to the waiter "Our bottle of wine is empty". Most likely the waiter will understand this as a request for a new bottle and he will go and get you one. He has understood the speaker meaning (i.e., what the speaker intended to communicate; Grice, 1989) even though the coded meaning (i.e., the linguistic code itself) did not specify a request at all. How does he do this? How does he know that with your remark about the state of the bottle you are actually requesting him to bring you a new one? Linguistic utterances are often underspecified with respect to the meaning they convey and understanding language goes beyond the mere coding and decoding of linguistic utterances (Hagoort, 2017). It is the communicative intention of the speaker that drives the listener's behavior, rather than his or her actual words, and listeners use the context in which linguistic utterances are provided to understand this intention (Levinson, 2014). The aim of this thesis is twofold; to shed light on the processes involved in understanding speaker meaning and to explore a new method, namely virtual reality, to investigate these processes in a naturalistic but controlled experimental setting.

Non-conventional indirect requests

This thesis focuses on non-conventional indirect request to investigate the processes involved in bridging the gap between coded meaning and speaker meaning. Requests are very often performed indirectly (Ervin-Tripp, 1976) and indirect requests differ from direct requests (e.g., "bring me a new bottle") in that they involve the performance of an indirect speech act via the performance of a direct act (e.g., *requesting* someone to bring a new bottle by *asserting* "Our bottle of wine is empty"; Holtgraves, 2008). Non-conventional indirect requests are different from more conventional indirect requests (e.g., "Could you get me a new bottle of wine?") as they are more ambiguous, they do not contain the request-based propositional content (e.g., "Get me a new bottle of wine" in "Could you get me a new bottle of wine?"), and they do not assert or question the conditions that underlie requests, for example the listener's willingness or ability to perform an action (e.g., "Would/Could you" in "Would/Could you bring us a new bottle of wine?"). As in the case of ironic sentences and indirect replies, the coded meaning of non-conventional indirect requests is very different from the speaker meaning, which makes them a good object of study for this thesis. One of the most common forms that non-conventional indirect requests (henceforth "indirect requests") take is a negative state remark (Holtgraves, 1994; 2008). Such a remark is based on the principle that a speaker can perform a request by questioning or asserting the existence of a negative state (or a state of which the listener can infer that it is negative in context). This is most successful if there is some action (e.g., get a new bottle of wine) that the listener can perform to alter the negative state (Holtgraves, 2002). In the present work, indirect requests also take the form of a negative state remark.

Indirect request processing

Previous research has shown that the neural infrastructure necessary to understand speaker meaning extends beyond classical language areas (e.g., Hagoort, 2017; see Hagoort & Levinson, 2014 for an overview). In short, these networks include the Theory of Mind network (e.g., Bašnáková, Weber, Petersson, Van Berkum, & Hagoort, 2013; Jang et al., 2013; Van Ackeren, Casasanto, Bekeering, Hagoort, & Ruschemeyer, 2012) and for indirect requests in particular, action-related regions (Van Ackeren et al., 2012; Van Ackeren, Smaragdi, & Rueschemeyer, 2016). From these fMRI studies it has been concluded that understanding speaker meaning depends on theory of mind, and in particular on reasoning about the other's intentions and purposes, and on updating what we think the other knows and what we think the other can infer (Hagoort & Levinson, 2014). However, less is known about the cognitive effort and time-sensitive processes involved in understanding speaker meaning. The first aim of this thesis is to contribute to this knowledge by investigating indirect request processing using different methods, pupillometry and electroencephalography (EEG), and in different contexts, a more classical experimental setting and in virtual reality (VR).

The chapters in this thesis are embedded in two related lines of theoretical and empirical research. The first is concerned with intention recognition and pragmatic inferences in conversation (Grice, 1957; Levinson, 2017; Searle, 1969; Sperber & Wilson, 1995). The second focuses more directly to the processing steps involved in understanding non-literal and indirect language (Gibbs, 1994; Giora, 2003; Grice, 1975). Both of these lines originated from Grice's (1957) proposal that intentions play a central role in communication and that these intentions are not always present in the lexical content of an utterance (i.e., the idea of speaker meaning and coded meaning).

To bridge the gap between speaker meaning and coded meaning, Grice proposed an inferential process guided by several 'principles of cooperation'. For example, the principle that interlocutors should be 'as informative as required and not more informative that is required' (i.e., the 'Maxims of Quantitiy'; as cited in Noveck & Reboul, 2008). It has been argued that Grice's theory was not designed to make explicit experimental predictions (Noveck & Reboul, 2008), however intention recognition and inferences also play an important role in more recent theories on pragmatics processing. Most theories agree that some amount of inference is necessary to close the gap between speaker meaning and coded meaning (Gibbs, 2003; Searle, 1969 Sperber & Wilson, 1995). Yet, there is some disagreement regarding the nature of this process, specifically regarding the automaticity and processing costs involved (Noveck & Reboul, 2008; Hagoort & Levinson, 2014). Relevance theory in particular claims that pragmatic inferences are associated with increased processing time and processing effort (Sperber & Wilson, 1995). On the other hand, Levinson (2000) proposes that inferences can be made fast and automatically. These proposals are mostly tested with respect to scalar implicatures, yet the question of whether pragmatic inferences are linked to processing effort is also relevant for indirect request processing. Another point of discussion in this line of research relates to the nature of the intention recognition (Holtgraves, 2008). One way to conceptualize the intention of the speaker is in terms of speech acts (Searle, 1969). A speech act conveys the sense in which utterances in conversation do not just carry meaning, but rather they perform actions (Levinson, 2017). In speech act theory (Austin, 1962; Searle 1969) utterances are considered to be actions with particular goals, such as apologizing, requesting or warning. According to Searle (1979), recognizing the speech act (or illocutionary force) of an utterance is crucial for utterance comprehension. Also in relevance theory intention recognition plays an important role, however in this theory utterance comprehension does not generally require illocutionary force recognition (Sperber & Wilson, 1995; see Holtgraves, 2008 for a discussion). Much of the experimental work on the nature of speech act recognition has focused on its automaticity and time-course. For example, Holtgraves (2008) presented participants with spoken and written sentences that performed specific speech acts (e.g., *complain* in "I don't know why we have to do so many experiments"). Overall, the results were most consistent with speech act theory; participants demonstrated automatic activation of the speech act for both spoken and the written sentences.

A second (related) line of research relevant to the current thesis is concerned with the processing steps involved in understanding indirect and non-literal language. Generally, theories are divided on whether the meaning of indirect utterances can be access directly – the direct access view- (Gibbs, 1983) or whether the listener first has to compute the literal meaning and only thereafter the indirect meaning can be understood (e.g., Grice, 1975; Searle 1975). According to the direct access view, the listener can by-pass the literal meaning of an utterance, if uttered in an appropriate context, and compute the indirect meaning immediately. Contextual information can very rapidly interact with lexical processes, so that the non-literal meaning can be understood early on during sentence comprehension (Gibbs, 1994; 2002). In contrast, the standard pragmatic view (or model), derived from the work by Grice (1975) and Searle (1975; 1993) proposes that access to the non-literal meaning is indirect. It postulates that listeners first have to compute the literal meaning of an utterance and then, if this meaning is inappropriate in the context, infer the speaker's communicative intention. Thus, because the literal meaning is first computed and only then the non-literal meaning is derived, some sort of processing cost is expected for the access to the non-literal meaning as compared to the literal meaning. Somewhat in between indirect and direct access is the graded salience hypothesis (Giora, 2003). According to this proposal, the most salient reading of the sentence will be the initial reading of a sentence (Giora, 2003). For example, for highly frequent conventional indirect request (e.g., "Can you get me a new bottle of wine?"), the initial reading will be the request reading (i.e. get me a new bottle of wine) rather than a question about the physical ability of the waiter to get a new bottle of wine (Coulson & Lovett, 2010). Studies have tested these theories, especially for irony (e.g., Spotorno, Cheylus, Van Der Henst, & Noveck, 2013), idioms (e.g., Canal, Pesciarelly, Vespignani, Mallinaro, & Cacciari, 2015) and metaphors (e.g., Bambini, 2016), and they obtained mixed results (for an overview, see Bambini, 2016). Most relevant for the present work is a study by Coulson & Lovett (2010) in which these theories were tested for indirect requests (i.e. non-conventional indirect requests). This study is discussed in more detail in Chapter 5, but in short, participants were presented with scenarios in which the final sentence could be interpreted as a request or a statement, depending on the preceding discourse. Based on the neurophysiological brain responses to the target sentence, the authors concluded that contextual cues can have an early influence during sentence processing, suggesting that the standard pragmatic model is not sufficient to describe indirect request processing. However, transient processing costs for indirect requests were also observed. The authors suggested that in the case of indirect requests participants had to draw on information from the context, world-knowledge and social conventions (e.g., speaker status) to interpret the indirect requests.

This thesis aims to contribute to these lines of research by investigating the cognitive effort involved in intention recognition for indirect requests (Chapters 2 and 3) and by examining the behavioral and neural correlates of indirect request processing in a more realistic experimental setting (Chapter 5). In addition, since in other fields of pragmatics it has been shown that pragmatic abilities can play a role in pragmatic processing (e.g., Nieuwland, Ditmar, & Kuperberg, 2010), the relation between individual differences in pragmatic abilities and cognitive effort during indirect request processing is investigated (Chapter 3).

Pupils and potentials

In this thesis behavioral measures, pupillometry and electroencephalography (EEG) were used to investigate indirect request processing. Pupillometry is the study of changes in pupil diameter as a function of cognitive processing (Sirios & Brisson, 2014). It dates back to the studies by Hess and Polt (1964), which showed that the pupil diameter of participants increased as a function of the difficulty of math problems that they were presented with. In their study, increased cognitive load resulted in an increase in pupil diameter (Hess & Polt, 1964). Since then pupillometry has been used in different domains of psychology, including language processing (e.g., Just & Carpenter, 1993; Engelhardt, Ferreira, & Patsenko, 2010; see Sirios & Brisson, 2014 for an overview). Pupillometry is useful to study indirect request processing, since it can shed light on the question of whether indirect requests require more cognitive effort to be processed than statements. Furthermore, it can provide insight into the conditions that can modulate the amount of cognitive effort necessary for someone to understand an indirect request (e.g., individual differences in pragmatic abilities). In Chapters 2 and 3 pupillometry was used to investigate cognitive effort and individual differences in indirect request comprehension.

EEG is a method to record electrical activity from the brain by means of electrodes placed on the scalp. In the studies in this thesis the EEG signal was time-locked to the onset of the critical stimulus and averaged across trials to create event-related potentials (ERPs). ERPs are used to extract relevant stable components from the relatively noisy EEG signal. These components are usually labeled according to their polarity and latency and they have been linked to specific aspects of processing. The N400, for example, is a negativity that peaks around 400 milliseconds after the onset of a critical stimulus and it has been linked to semantic processing (for an overview see Kutas & Federmeier, 2011). In Chapters 4 and 5, ERPs were used to investigate audio-visual integration and indirect request processing, respectively.

Virtual reality

In addition to investigating the processes that underlie indirect request processing, this thesis aims to test the possibilities of VR to increase the ecological validity of studies in experimental- and neuropragmatics. VR is a digitally created space in which people can interact with objects, people and environments, bound only by the limits of human imagination (Fox, Arena, & Bailenson, 2009). In the early nineties, scientists began to consider ways in which VR could be used to study psychological phenomena and social interaction (e.g., Loomis, Blascovich, & Beall, 1999). This involves placing participants in a virtual environment (VE) in which their movements are tracked and their surroundings are digitally displayed in accordance with those movements. One of the biggest advantages of VR is that it provides researchers with the possibility to increase the ecological validity of their studies, while maintaining experimental control. Ecological validity refers to the extent to which observations can be generalized beyond constrained laboratory settings to natural behavior in the world (e.g., Brewer, 2000; Parsons, 2015). VR makes it possible to study behavior in different environments, without interference from uncontrollable cues, and it allows for the manipulation of variables that were previously hard to replicate in an experimental setting (Blascovich & Bailenson, 2011; Fox, Arena, & Bailenson, 2009).

VR has successfully been used even in combination with neurophysiological measures, to study spatial navigation (Bischof & Boulanger, 2003), driving behavior (e.g., Bayliss & Ballard, 2000), and spatial presence (Baumgartner, Valko, Esslen, & Jänke, 2006).

However, the use of VR is much less common in the study of language behavior, especially in combination with electrophysiological measures. Although recently several behavioral studies have used VR to investigate language behavior (e.g., Heyselaar, Hagoort & Segeart, 2017), I am not aware of any studies that have used VR in combination with electrophysiological brain measures to study language processes. Thus, it is necessary to provide a proof of concept of the feasibility of the combined use of VR and EEG to study language comprehension. To this end, audio-visual semantic integration was investigated using EEG in VR (Chapter 4). Previous research has shown that mismatches between auditory and visual information during sentence processing consistently elicit an N400 effect (e.g., Peeters, Hagoort, & Özyürek, 2015; Willems, Özyürek, & Hagoort, 2008), thus to allow for a reliable comparison with results obtained outside VR, this manipulation was used as a first step toward combining VR and EEG in language research (Chapter 4).

VR allows for the creation of rich experimental contexts and more realistic 'interactional' settings and therefore it might be especially useful for the study of pragmatic language phenomena, in which the context is especially crucial. For example, it makes it possible to include a visible (virtual) speaker with an established social relation to the listener in the experiment (Hoeks & Brouwer, 2010) and it allows for the inclusion of visual information that we use in everyday life to interpret what we hear (Knoeferle, 2015). Consider the context that was sketched at the beginning of this introduction. There was an exchange of utterances, a restaurant, a bottle of wine, and a relationship and between the two people in the exchange (i.e. a waiter and a guest). One could argue that to study how the waiter understood the indirect requests, it is important to create a setting that closely resembles this everyday environment. In Chapter 5 of this thesis, VR was used to create a virtual restaurant in which indirect request processing was investigated with behavioral and neurophysiological measures.

Thesis outline

In this thesis four empirical studies are presented. As the chapters were written as individual articles there is some overlap, especially in the introductory sections.

Chapter 2 is concerned with the use of pupillometry to study the cognitive effort involved in processing indirect requests as compared to (direct) statements. In two experiments, pupil diameter was used as an index of cognitive effort during sentence comprehension. Participants were presented with combinations of pictures and spoken sentences that could be interpreted as either indirect requests or statements. If processing indirect requests requires more cognitive effort than processing statements, this should be reflected in an increase in pupil diameter. Furthermore, this would demonstrate that pupillometry could be a useful tool for the field of experimental pragmatics.

In Chapter 3 the role of pragmatic abilities during indirect request processing was explored. Individuals vary in how and when they apply pragmatic knowledge and in other fields of pragmatics these differences have been shown to influence pragmatic processing (e.g., Nieuwland, Ditman & Kuperberg, 2010). In this study, participants were presented with scene-sentence combinations that could be interpreted as either indirect requests or statements. The communication subscale of the Autism Quotient was used to quantify a participant's pragmatic abilities. The aim of this study was to characterize the relation between individual differences in pragmatic abilities and the cognitive effort involved in indirect request processing.

In Chapter 4, EEG and VR were combined to examine the neurophysiological underpinnings of language processing in a VR environment. Participants were immersed in a virtual restaurant in which virtual agents (i.e. restaurant guests) produced sentences that related to a food or drink item on the table in front of them. The noun in the utterance (e.g., "I just ordered this salmon") could either match ("salmon") or mismatch ("pasta") with the item on the table (e.g., a plate with salmon), resulting in a possible mismatch between visual and auditory information. If similar ERP effects, namely an N400 effect, would be observed as in other non-VR studies with similar paradigms, this would serve as a proof of concept that VR and EEG can reliably be combined to study language processing in a more contextually rich naturalistic context.

Finally, in Chapter 5, VR and EEG were combined to study the behavioral and neurophysiological signatures of indirect request processing. In two experiments, participants were immersed in a virtual restaurant in which they were asked to listen to and respond to utterances from virtual guests in the restaurant. They were presented with possible requests (e.g., "My soup is cold") and statements (e.g., "My soup is nice"). By means of instruction and a virtual mirror in the restaurant, participants were assigned a role before the start of the experiment, namely to be a waiter or a restaurant critic. We predicted that 'waiters' would interpret the possible request as requests and the statements as statements, while 'restaurant critics' would interpret both sentence types as statements. In Experiment 2, the same design as in Experiment 1 was used with the addition that the EEG was measured. The prediction was to find a difference in the ERPs between possible requests and statements for the waiters, but not for the critics. On a methodological level, this study explored whether the behavioral and neural correlates of pragmatic language processing can be measured in a more naturalistic (virtual) environment.

Chapter 6 provides a summary and discussion of the four experimental chapters. In addition, several suggestions are provided for future research on indirect request processing and the use of VR in the fields of language comprehension and pragmatics.

References

Austin, J. L. (1962). *How to Do Things with Words.* Oxford: Oxford University Press.

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016). Disentangling metaphor from context: an ERP study. *Frontiers in psychology*, *7*.

Baumgartner, T., Valko, L., Esslen, M., & Jäncke, L. (2006). Neural correlates of spatial presence in an arousing and noninteractive virtual reality: an EEG and psychophysiology study. *CyberPsychology & Behavior*, *9*, 30-45.

Bayliss, J. D., & Ballard, D. H. (2000). Single trial P3 epoch recognition in a virtual environment. *Neurocomputing*, *32*, 637-642.

Bischof, W. F., & Boulanger, P. (2003). Spatial navigation in virtual reality environments: an EEG analysis. *CyberPsychology & Behavior*, *6*(5), 487-495.

Blascovich, J., & Bailenson, J. (2011). *Infinite reality: Avatars, eternal life, new worlds, and the dawn of the virtual revolution.* New York, NY: William Morrow & Co.

Bašnáková, J., Weber, K., Petersson, K. M., Van Berkum, J., Hagoort, P. (2013). Beyond the language given: The neural correlates of inferring speaker meaning. *Cerebral Cortex, 24*(10), 2572–2578.

Canal, P., Pesciarelli, F., Vespignani, F., Molinaro, N., & Cacciari, C. (2015). Electrophysiological correlates idioms comprehension: semantic composition does not follow lexical retrieval. In V. Pirrelli, C. Marzi, & M. Ferro (Eds.), *Proceedings of the NetWordS Final Conference on Word Knowledge and Word Usage: Representations and Processes in the Mental Lexicon,* Pisa, Italy, March 30–April 1, pp. 98–101.

Coulson, S., & Lovett, C. (2010). Comprehension of non-conventional indirect requests: An event-related brain potential study. *Italian Journal of Linguistics, 22*(1), 107-124.

Egorova, N., Shtyrov, Y., & Pulvermüller, F. (2013). Early and parallel processing of pragmatic and semantic information in speech acts: neurophysiological evidence. *Frontiers in Human Neuroscience,* 7.

Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *Quarterly Journal of Experimental Psychology*, *63*(4), 639-645.

Ervin-Tripp, S. (1976). Is Sybil there? The structure of some American English directives. *Language in Society, 5*(1), 25-66.

Fox, J., Arena, D., & Bailenson, J. N. (2009). Virtual reality: A survival guide for the social scientist. *Journal of Media Psychology, 21*(3), 95-113.

Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding.* Cambridge: Cambridge University Press.

Giora, R. (2003). *On our mind: Salience, context, and figurative language.* Oxford: Oxford University Press.

Gisladottir, R. S., Chwilla, D. J., & Levinson, S. C. (2015). Conversation electrified: ERP correlates of speech act recognition in underspecified utterances. *PloS one, 10*(3), e0120068.

Grice, H. P. (1957). Meaning. *The philosophical review, 66*(3), 377-388.

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

Hagoort, P., & Levinson, S. C. (2014). Neuropragmatics. In M.S. Gazzaniga, & G.R. Mangun (Eds.), *The Cognitive Neurosciences* (pp. 667–674). Cambridge, MA: MIT Press.

Hagoort, P. (2017). The core and beyond in the language-ready brain. *Neuroscience & Biobehavioral Reviews*, 81, 194-204.

Heyselaar, E., Hagoort, P., & Segaert, K. (2017). In dialogue with an avatar, language behavior is identical to dialogue with a human partner. *Behavior Research Methods, 49*(1), 46-60.

Hoeks, J. C., & Brouwer, H. (2014). Electrophysiological Research on Conversation and Discourse. Holtgraves, T. (Ed.), *Oxford Handbook of Language and Social Psychology*, pp 365-386. Oxford: Oxford University press.

Holtgraves, T. (1994). Communication in context: Effects of speaker status on the comprehension of indirect requests. *Journal of Experimental Psychology: Learning Memory and Cognition*, *20*(5), 1215–1208.

Holtgraves, T. (2008). Automatic intention recognition in conversation processing. *Journal of Memory and Language, 58*(3), 627-645.

Holtgraves, T. (2012). The role of the right hemisphere in speech act comprehension. *Brain and Language, 121*(1), 58-64.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology-revue Canadienne De Psychologie Experimentale*, *47*, 310-339.

Knoeferle, P. (2015). Language comprehension in rich non-linguistic contexts: combining eye tracking and event-related brain potentials. In R.M. Willems (Ed), *Cognitive Neuroscience of Natural Language Use* (pp. 77-100). Cambridge: Cambridge University Press.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology, 62*, 621.

Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers, 31*(4), 557-564.

Levinson, S. C. (2013). Action formation and ascription. In T. Stivers, & J. Sidnell (Eds.), *The handbook of conversation analysis* (pp. 103-130). Hoboken, NJ: Wiley-Blackwell.

Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language, 63*(3), 324-346.

Noveck, I. A., & Reboul, A. (2008). Experimental pragmatics: A Gricean turn in the study of language. *Trends in Cognitive Sciences, 12*(11), 425-431.

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science, 5(*6), 679-692.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd edition). Oxford: UK Blackwell Publishers.

Spotorno, N., Cheylus, A., Van Der Henst, J. B., & Noveck, I. A. (2013). What's behind a P600? Integration operations during irony processing. *Plos One, 8*(6), e66839.

Van Ackeren, M. J., Casasanto, D., Bekkering, H., Hagoort, P., & Rueschemeyer, S. A. (2012). Pragmatics in action: Indirect requests engage theory of mind areas and the cortical motor network. *Journal of Cognitive Neuroscience, 24*(11), 2237–2247.

Van Ackeren, M. J., Smaragdi, A., & Rueschemeyer, S. A. (2016). Neuronal interactions between mentalising and action systems during indirect request processing. *Social Cognitive and Affective Neuroscience, 11*(9), 1402–1410.

Chapter 2

Pupillometry reveals increased pupil size during indirect request comprehension

Abstract

Fluctuations in pupil size have been shown to reflect variations in processing demands during lexical and syntactic processing in language comprehension. An issue that has not received attention is whether pupil size also varies due to pragmatic manipulations. In two pupillometry experiments, we investigated whether pupil diameter was sensitive to increased processing demands as a result of comprehending an indirect request versus a direct statement. Adult participants were presented with 120 picture-sentence combinations that could either be interpreted as an indirect request (a picture of a window with the sentence “it's very hot here”) or as a statement (a picture of a window with the sentence “it's very nice here”). Based on the hypothesis that understanding indirect utterances requires additional inferences to be made on the part of the listener, we predicted a larger pupil diameter for indirect requests than statements. The results of both experiments are consistent with this expectation. We suggest that the increase in pupil size reflects additional processing demands for the comprehension of indirect requests as compared to statements. This research demonstrates the usefulness of pupillometry as a tool for experimental research in pragmatics.

Tromp, J., Hagoort, P., & Meyer, A. S. (2016). Pupillometry reveals increased pupil size during indirect request comprehension. *The Quarterly Journal of Experimental Psychology, 69*(6), 1093-1108.

Pupillometry reveals increased pupil size during indirect request comprehension

Fluctuations in pupil size have been shown to reflect variations in processing demands during a number of cognitive tasks (Kahneman, 1973, for an overview see: Beatty, 1982; Sirios & Brisson, 2014). For example, Hess and Polt (1964) recorded the pupil size of participants solving math problems that varied in difficulty. Their findings indicated that solving more difficult problems was accompanied by larger pupil diameters. Pupil size has also been shown to vary with processing effort during visual search and counting tasks (Porter, Troscianko, & Gilchrist, 2007), digit list recall (e.g., Piquado, Isaacowitz, & Wingfield, 2010) and with working memory load (e.g., Attar, Schneps, & Pomplun, 2013).

More recently, pupillometry has been used to study language processing, and pupillary responses have been taken as an index of increases in processing demands during sentence comprehension (e.g., Just, Carpenter, & Miyake, 2003). These demands can arise in a number of ways. For instance, processing negative sentences as compared to affirmative sentences resulted in an increase in pupil size (Beatty, 1982). Also, Schluroff (1982) observed a larger pupil diameter during the processing of grammatically complex sentences as compared to their simpler counterparts. He suggested that pupil size may be of considerable use as an online monitor of cognitive load imposed by grammatical complexity. Similarly, Just and Carpenter (1993) observed a larger change in pupil diameter for object relative sentences (e.g., *The senator that the reporter attacked admitted the error*) as compared to subject relative sentences (e.g., *The senator that attacked the reporter admitted the error*), within 1.2 seconds after the critical verb.

In addition to grammatical complexity, pupil diameter is sensitive to lexical and syntactical ambiguity (Ben-Nun, 1986; Schluroff et al., 1986) and to prosody manipulations (e.g., Zellin et al., 2011). Engelhardt, Ferreira, and Patsenko (2010) investigated the processing of syntactically ambiguous sentences in relation to prosody and visual context. In their first experiment, participants were presented with garden-path sentences (e.g., “While the woman cleaned the dog that was big and brown stood in the yard.”) accompanied by correct prosody (i.e., a prosodic break between “cleaned” and “the dog”) or conflicting prosody (i.e., no prosodic break). They predicted that, if prosody influences online processing of the sentence, an increase in pupil diameter should be observed for the incongruent sentence-prosody condition, indicating more cognitive effort. This prediction was confirmed, as pupil size reliably increased when prosodic structure was inconsistent with syntactic structure. In their second experiment, a visual context was added to the prosody manipulation. The visual context could either be congruent or incongruent with the correct interpretation of the sentence. The results indicated an interaction between prosody and visual context. When visual context was consistent with the correct interpretation of the sentence, prosody had little effect on processing effort. In contrast, when the visual context was inconsistent with the correct interpretation, prosody had an effect on processing effort. This suggests that, in addition to prosody, visual context affected online processing load as measured by pupil diameter change. In sum, there is strong evidence that pupil diameter during sentence comprehension is sensitive to differences in cognitive load resulting from increases in sentence complexity or ambiguity.

An issue that has not received any attention is whether pupil size is also sensitive to pragmatic manipulations. In the current study we investigated whether pupil diameter was sensitive to increased processing demands for non-conventional indirect requests compared to direct statements. During natural conversation, communication is often indirect. We might hint at what we want rather than expressing it directly. For example, in an appropriate context “It's cold in here” may be a request to shut the window, rather than a statement about the room temperature (Holtgraves, 1994). The way in which we comprehend the intended meaning of indirect speech acts, such as the indirect request in the example above, has been a topic of much debate (Holtgraves, 2002). However, most theories agree that understanding this type of request requires some form of intention recognition (Austin, 1962; Holtgraves, 1994; Levinson, 2000; Searle, 1975, 1979; Sperber & Wilson, 1986, 1995). The listener has to infer that the speaker intents to request something in order to interpret the utterance correctly. Indirect requests vary in their conventionality (e.g., Gibbs, 1986; Holtgraves, 1994, 2002). For example, “Can you pass the salt?” is an indirect request, but it is conventional. It has a literal meaning (“I ask if you are able to pass the salt”) and an indirect meaning (“I request you to pass the salt”). Generally, when a speech act is performed by conventional means, it suggests that the literal meaning is not to be taken seriously (Holtgraves, 2002). Usually this type of request can be performed by asserting or questioning the felicity conditions that underlie requests (e.g., question the hearer's ability). Also, it contains the request-based propositional content (e.g., “pass the salt”) and it allows the preverbal insertion of "please" (e.g., "Can you please pass the salt?"; Holtgraves, 2002). Research suggested that these conventional indirect requests are recognized fast (Gibbs, 1981, 1983) and immediately (Clark, 1979). Gibbs (1983) proposed that people do not always have to retrieve the literal meaning of conventional indirect requests first, but rather they can compute the indirect meaning automatically (see also Holtgraves, 1994). Thus, even though these types of sentences have two meanings, retrieving the indirect one does not seem to require additional processing effort on the part of the listener. Indirect requests that cannot be characterized by the features mentioned above are categorized as nonconventional (Holtgraves, 2002) and they are the focus of the current research. One common form of this type is a *negative state remark*, where the speaker asserts or questions a negative state (e.g., "It's cold in here"), which can be eliminated or lessened by the hearer (e.g., by closing the window; Holtgraves, 1994). Research showed that participants took longer to comprehend this type of request as compared to the conventional type (e.g., Gibbs, 1981; Holtgraves, 1994) and it has been suggested that comprehending these requests involves an inference process (Holtgraves 1994, 2002). During this process, listeners take into account information from other sources than the linguistic code, for example contextual factors and prior knowledge of the conversational partner's intentions (Holtgraves, 1994). That an additional inference process is necessary during comprehension of certain types of indirect speech acts is also supported by two recent neuroimaging studies. Bašnáková, Weber, Petersson, Van Berkum, and Hagoort (2013) conducted a functional magnetic resonance imaging (fMRI) study to investigate the neural underpinnings of inferring speaker meaning (i.e., the message of the speaker). Participants listened to sentences (e.g., "It's hard to throw a good party") that had different meanings depending on the dialogue and final question that preceded it. For example, the sentence mentioned above is a direct reply to the question "How hard is it to throw a party?", but it can also be an indirect reply to the question "Will you throw a party for your graduation?" Furthermore, it can be indirect reply to the question "Did you enjoy yourself at my party?". In the latter case, the motivation of the speaker for using an indirect reply is "face saving", or to mutually protect on another's public self (e.g., Brown & Levinson, 1987). For indirect replies as compared to direct replies, increased activation was found in and areas relevant for discourse-level processing (bilateral prefrontal cortex and right temporal regions) and areas involved in mentalizing and empathy (medial frontal cortex, MFC; right temporoparietal junction, TPJ; and anterior insula). For face saving replies, there was additional activation in regions involved in affective and social cognitive processing, such as insula and anterior cingulate cortex (ACC). Bašnáková and colleagues concluded, based on the activation pattern for the indirect replies, that when inferring speaker meaning, listeners take the speaker's perspective on both cognitive (theory of mind) and affective levels. Thus, comprehending indirect replies seems to rely on inferences made by the listener.

Most relevant for the present purposes is a study by Van Ackeren, Casasanto, Bekkering, Hagoort, and Rueschemeyer (2012). They investigated the neural correlates of indirect request (IR) comprehension. Participants were presented with picture-sentence combinations as shown in Figure 1. In each item set, two sentences were combined with each of the two pictures, such that in one combination the utterance could be interpreted as an indirect request, whereas in the remaining combinations it was a statement. For example, “It's very hot here” in combination with a picture of a window may be interpreted as an indirect request to open the door, while the sentence “It's very nice here” with the same picture would most likely be interpreted as a mere statement. First, Van Ackeren et al. (2012) observed increased activation in cortical motor areas for indirect requests as compared to statements. In addition, they expected increased activation in theory of mind (ToM) areas, such as medial prefrontal cortex (mPFC) and TPJ (Gallagher & Frith, 2003), for indirect requests as compared to statements, since making inferences about mental states of others has often been associated with having ToM. This prediction was confirmed: Both mPFC and left TPJ were sensitive to indirect requests versus statements. The authors concluded that, quite probably, these regions were crucial for making inferences about the communicative intent of speaker during IR comprehension.

In the present study we used a subset of the stimuli created by Van Ackeren et al. (2012) and recorded the participants' pupil size while they listened to the sentences and viewed the pictures. One-quarter of the scene-sentence combinations could be interpreted as indirect requests [e.g., a picture of a window (scene) and the sentence “It is very hot here”]. The other combinations served as controls for the indirect requests and could only be interpreted as statements [e.g., a picture of a window (scene) and the sentence “It is very nice here”]. In Experiment 1, following each picture-sentence pair, the participant had to indicate whether or not the utterance was an indirect request. In Experiment 2, a control experiment, participants were asked to make an affirmative response when they heard a direct statement.

Our first aim was to test whether pupil size would be sensitive to this difference in the implied meaning of the utterances. Since pupillometry is a relatively cheap noninvasive tool, it would be useful to demonstrate its applicability for studies of pragmatics. Our second aim was to investigate the cognitive effort involved in IR comprehension. Although the abovementioned fMRI studies are informative regarding the neural infrastructure supporting the comprehension of indirect utterances, they provide little information about the processing costs involved in understanding them. As noted above, there is strong evidence that pupil size is a good indicator of mental effort (e.g., Beatty, 1982; Engelhardt et al., 2010; Piquado et al., 2010). Thus, if deriving the meaning of indirect requests involves cognitive effort beyond the effort entailed in understanding mere statements, we should see this reflected in the participants' pupil size, which should be larger for the indirect requests than for the control combinations.

Experiment 1

Method

Participants

Forty-nine native speakers of Dutch participated in the study (nine men, mean age = 20.8 years, range = 18 - 26 years). All participants had normal hearing, normal or corrected-to-normal vision and no history of language disorders. All but one were right-handed. Informed consent was obtained from all participants. They were paid for taking part in the experiment. Ethical approval for the study was granted by the ethics board of the Social Sciences Faculty of Radboud University.

Materials and design

Materials consisted of a subset[1] of the materials used by Van Ackeren et al. (2012),

---

[1] Four item sets from the original study by Van Ackeren and colleagues (2012) were removed because in these sets the critical word was repeated, which could influence pupil dilation (e.g., Otero, Weekes, & Hutton, 2011). We randomly selected three of these four item sets to use in the practice block.

namely 120 images of visual scenes and 120 spoken sentences. The visual scenes were collected from publically available online search engines (e.g., flickr.com) and the sentences were recordings of a native speaker of Dutch. The stimuli were divided into 60 item sets. Each set (see Figure 1) consisted of two pictures and two sentences. Pictures were labeled “action picture” (AP) when they could appear in the IR (action) condition or “no-action picture” (NP) when they could only be in the statement conditions. The same was done for the two utterance types (AU = “action utterance”, NU = “no-action utterance”). The pictures and sentences could be combined in four different ways, which resulted in four experimental conditions; indirect request (AP/AU), picture control (AP/NU), Utterance control (NP/AU) and picture-utterance control (NP/NU). In a pretest, Van Ackeren et al. (2012) confirmed that the indirect request (AP/AU) sentence-scene combinations were interpreted as indirect requests more often than items in the other conditions. The control conditions were included to control for the unique effects of picture and utterance separately.

Picture	Utterance	Condition
	"It is very hot here" *Action Utterance (AU)*	Indirect request (AP/AU) *Action Picture / Action Utterance*
Action Picture (AP)	"It is very nice here" *No-action Utterance (NU)*	Picture control (AP/NU) *Action Picture / No-action Utterance*
	"It is very hot here" *Action Utterance (AU)*	Utterance control (NP/AU) *No-action Picture / Action Utterance*
No-action Picture (NP)	"It is very nice here" *No-action Utterance (NU)*	Picture-utterance control (NP/NU) *No-action Picture / No-action Utterance*

Figure 1. Illustration of the design with a single item set.

Since pupil size is sensitive to luminosity, luminosity values of the pictures were adjusted so that all pictures had values between 140 and 160. Luminosity was measured using the luminosity tool in Adobe Photoshop, version 11.0.2. Picture size was kept relatively small (250 x 250 pixels), so that the larger part of the computer screen was white. Each participant saw two scene-sentence combinations from each item set. For example, from the set in Figure 1, participant 1 would see the AP/AU (indirect request) scene-sentence combination and the NA/NU (picture-utterance control) combination. Participant 2 would see the remaining combinations. Thus, two lists were created, and items were never repeated within participants. Each participant viewed 30 combinations per condition, resulting in a total of 120 trials. The sentence-scene combinations were distributed over four blocks (30 items per block) and they were pseudo-randomized so that combinations from the same condition were never presented more than two times in a row. Between blocks, the participant was encouraged to take short break. Before the experimental blocks, participants completed 12 practice trials.

Apparatus and procedure

Participants were seated in a medium-lit sound-proof booth. The lighting was kept constant for all participants. Stimuli were presented using Experiment Builder version 1.10.1025 (SR Research Ltd., Mississauga, Canada). The sentences were presented through Sennheiser HD201 Lightweight over-ear binaural headphones. The pupil diameter of each participant's right eye was measured with an Eyelink 1000 Tower Mount eye-tracker (SR Research Ltd., Mississauga, Canada). In Eyelink 1000, pupil size is measured in arbitrary units which have a linear relation to the recorded pupil diameter (see Eyelink user manual; Einhäuser, Stout, Koch, & Carter, 2008). Before the start of the experiment, randomized target order calibration and validation routines were performed using EyeLink 1000 software (SR Research Ltd., Mississauga, Canada). Button presses were recorded by means of a button box. Each trial started with a fixation cross that was presented for 1000 ms, after which the visual scene appeared on the screen. Two-hundred ms later the sentence was presented through the headphones. Then, the fixation cross appeared again for 2500 ms followed by the statement: “The person made a request”. Participants then indicated whether or not they thought the statement was true (right button press) or false (left button press). After participants made their choice the fixation cross appeared again for 2500 ms to give the pupil enough time to return to baseline before the next trial (see Figure 2).

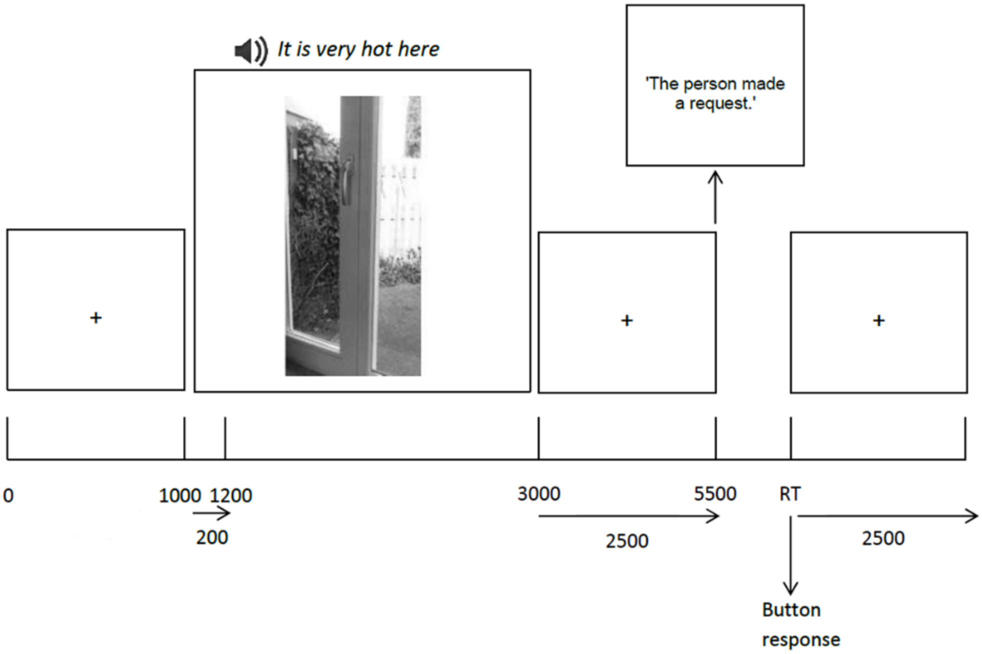

Figure 2. Example of the time-course of a single trial (time in ms).

Behavioral data analysis

A one-sample *t* test (test value = 0.5) was conducted on the correct responses to the AP/AU combinations to assess whether participants performed above chance in identifying the indirect requests. Further analysis of comprehension accuracy was conducted with logit mixed models in R (Jaeger, 2008). Predictors were mean-centered. The model included the fixed effects Picture (action, no-action), and Utterance (action, no-action) and the interaction. Also it contained random intercepts and slopes for Picture and Utterance by participant and a by-item (picture) random intercept for the effect of Utterance. This was the maximal random structure justified by the data leading to convergence (Barr, Levy, Scheepers, & Tily, 2013).

Reaction times were analyzed with linear mixed effects models in R (version 3.0.3; The R foundation for statistical computing; lme4 package, Bates et al., 2014). The model included the fixed effects Picture (action, no-action), and Utterance (action, no-action) and the interaction. The random structure of the model was the same as that for the comprehension accuracy analysis described above.

Pupillometry data analysis

Pupillometry data were pre-processed and analyzed in R (version 3.0.3). The R-scripts for the signal pre-processing procedure were developed by Gerakaki, Sjerps, and Meyer (in prep.). Pupil dilation was originally measured with a sampling rate of 500 Hz, for the analysis the signal was down-sampled to 50 Hz. To detect and remove outliers, the change in pupil diameter was assessed from sample to sample. Based on Piquado et al. (2010), all data point with a ratio that differed more than one standard deviation from the mean pupil change of the trial, were categorized as outliers. Outliers were treated as missing values and linear interpolation was used to replace them Trials were completely removed if more than 25% of values were missing (3.4% of the data). On a trial by trial basis, absolute pupil diameter was transformed to relative pupil diameter by means of baseline-correction and normalization. This was done to correct for tonic changes in pupil dilation and to allow for a comparison between participants (e.g., Van Rijn, Dalenberg, Borst, & Springer, 2012). First, the baseline pupil size of a given trial was subtracted from each sample in the trial. These values were then divided by the baseline to calculate the pupil size change. The baseline was defined as the average pupil size in the first 1000 ms of a trial. In this time window, a fixation cross was presented on the screen. To plot the task-evoked pupillary responses (TERPs), which represent the percentage of pupil diameter change (PDC) over time, the value of pupil diameter change was multiplied by 100. For the statistical analyses, each trial was partitioned into four parts; baseline (0 to 1000 ms), audio (1000 ms to critical word onset (M=2337)), critical (1500 ms window from critical word onset), end of trial (from end of critical window to trial offset). The choice for a critical window of 1500 ms from word onset was based on a study by Just and Carpenter (1993), which found the largest peak in pupil size ~1.2 seconds after the critical word offset in an ambiguous sentence. We took a time window of 1.5 seconds starting from critical word onset, adopting the 1.2 seconds window of Just and Carpenter (1993) plus 300 ms for word recognition (see Engelhardt et al., 2010, for a similar approach).

Mean pupil size was analyzed using linear mixed effects models in R (version 3.0.3; The R foundation for statistical computing; lme4 package, Bates et al., 2014) which allow for simultaneous inclusion of items and participants as random variables (Baayen, Davidson, & Bates, 2008). Statistical analyses were performed only for the critical time window. The predictors (Picture, Utterance) and the random structure were the same as in the logit mixed model for the comprehension accuracy data. To assess the effects of Utterance, Picture, and the interaction, a backwards elimination procedure was used in which models were compared using a likelihood ratio test. The same procedure was followed for the peak pupil size data. Peak pupil size was included as an additional dependent variable in this study, since this measure has been shown to be less dependent on the number of observations in the critical time window than mean pupil size (Beatty & Lucero-Wagoner, 2000). This is important, since in our study indirect requests occur less frequently than statements (ratio: 1:3). Although recent papers (e.g., Wierda, van Rijn, Taatgen, & Martens, 2012, for an overview see Sirios & Brisson, 2014) proposed more advanced analyses of the time-course of pupillary responses, the analyses reported here are sufficient for the purposes of the current research.

Results

Behavioral results

Utterances were categorized as requests more often in the IR condition as compared to the control conditions (see Table 1). In line with the results reported by Van Ackeren et al. (2012), a one-sampled t test (test value =.5) confirmed that participants were able to correctly identify the AP/AU combinations as indirect requests, $t(48) = 14.64$, $p < .001$ ($M = 76.97\%$, $SE = 1.12\%$).

Table 1. *Experiment 1: Percentage of IR responses ("yes" to the statement "The person made a request") per condition for the entire sample (top row, n = 49) and for the subset (bottom row, n = 22)*

		Indirect request (AP/AU)		*Picture Control (AP/NU)*		*Utterance control (NP/AU)*		*Picture-utterance control (NP/NU)*	
Condition	*n*	*%*	*SE*	*%*	*SE*	*%*	*SE*	*%*	*SE*
IR responses	49	76.97	1.12	27.22	1.19	16.56	0.99	14.62	0.93
IR responses	22	80.97	1.55	16.93	1.50	8.78	1.12	9.06	1.14

Note: AP = action picture; NP = no-action picture; AU = action utterance; NU = no-action utterance.

The logit mixed effect model for comprehension accuracy (n = 49) indicated a significant effect of Picture (β = -1.072, SE = 0.123, z = -8.750, p < .001). Accuracy was lower for the action pictures (M = 74.89%, SE = 0.82%) than the no-action pictures (M = 84.42%, SE = 0.68%). There was no effect of Utterance (β = 0.017, SE = 0.225, z = 0.076, p = .94), nor an interaction between Picture and Utterance (β = 0.411, SE = 0.278, z = 1.475, p = .14).

There was substantial variation in accuracy rates across participants. For the first analysis of the pupillometry data reported below, we selected only participants with accuracy rates of 70% or higher for each condition. This criterion was used to make sure that we captured IR comprehension in the AP/AU condition and to ensure that after removal of incorrect trials, participants still contributed similar numbers of data points to the analysis. For this group of 22 participants the pattern of comprehension accuracy across conditions was similar to the pattern for the entire sample (see bottom row of Table 1). In the statistical analysis of the comprehension accuracy data we again only found an effect of Picture (β = -0.789 SE = 0.198, z = -3.981, p < .001). Accuracy was lower for action pictures (M = 82.00%, SE = 1.08%) than no-action pictures (M = 91.08%, SE = 0.80%). There was no effect of Utterance (β = -0.069, SE = 0.245, z = -0.281, p = .78), nor an interaction between Picture and Utterance (β = -0.257, SE = 0.330, z = -0.778, p = .44). For the reaction times, the best fitting model included the interaction between Picture and Utterance (β = -2.134, SE = 36.918, t = -2.134). There was no evidence for a main effect of Picture (t < 1.4), nor of

Utterance (t < 1). Closer inspection of the interaction revealed a trend (p = .07) for shorter reaction times for the AP/AU (M = 524, SE = 16) as compared to AP/NU (M = 587, SE = 18) combinations. There was no difference for the no-action pictures--that is, between NP/AU (M = 599, SE = 18) and NP/NU (M = 568, SE = 13).

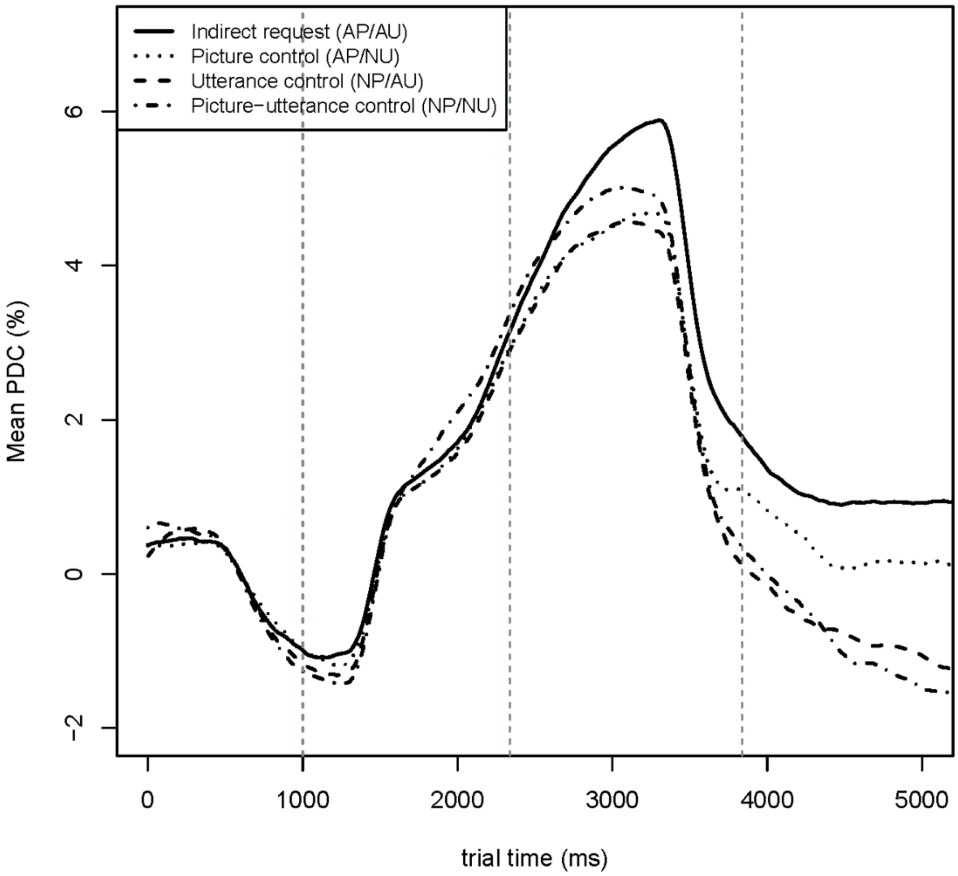

Figure 3. Average percentage of pupil diameter change (PDC) as a function of condition (IR, PC, UC, PUC). Trial time (ms) is represented on the x-axis and pupil diameter change (%) on the y-axis. The vertical lines represent the different time windows: baseline (0 to 1000 ms), audio (1000 ms to critical word onset, M=2337), critical (critical word onset + 1500-ms window after critical word onset) and end of trial (from end of the critical time window to end of trial).

Pupillometry results

The results reported here are based on the trials with correct responses. However, the same pattern is present when all trials are included in the analysis. Visual inspection of the TEPRs, which show the percentage of pupil diameter change (PDC) per condition (see Figure 3), suggested a larger pupil diameter for the AP/AU combinations (indirect requests) as compared to all other conditions in the critical time-window (1.5 s after critical word onset).

No difference between conditions was observed in the preceding time window (audio). In the last window (end of trial), a larger mean pupil was still observed for the AP/AU combinations as compared to all other conditions. Also, the AP/NU condition showed a slightly higher mean than the other control conditions.

The statistical analyses revealed that, for the critical time-window, the optimal model for the mean pupil size contained the interaction between Picture and Utterance (β = 0.012, SE = 0.005, t = 2.051). Including the interaction significantly improved model fit, (χ^2 (1) = 4.182, $p < .05$). There was no evidence for the fixed effects of picture and/or utterance (all $t < 1$). Further inspection of the interaction, based on planned comparisons, revealed an effect of utterance type (action versus no-action) for the action pictures, $t(21) = 2.78$, $p < .02$, but not for the no-action pictures, $t(21) = -1.234$, $p = 0.2308$. For the action pictures, mean pupil size was larger for action utterance--that is, indirect requests (M = 0.051, SE = .003)--than no-action utterances (M = 0.041, SE = .004).

In the model for the peak pupil size there was no evidence for a main effect of Picture (β = 0.004, SE = 0.004, t = 1.197), or Utterance (β = 0.002, SE = 0.005, t = .043). However, as for the mean pupil diameter, there was evidence for the interaction between Picture and Utterance (β = 0.014, SE = 0.006, t = 2.222). Including the interaction improved the fit of the model (χ^2 (1) = 4.960, $p < .03$). Further examination of the interaction revealed a difference between utterance types (action versus no-action) for the action pictures ($t(21)$ = 3.434, $p < .005$), but not for the no-action pictures ($t(21) = -1.131$, $p = 0.271$). In line with the results of the analysis of the mean pupil diameter, the peak for the AP/AU combinations (i.e. the indirect requests) was larger (M = .080, SE = 0.003) than the peak for the AP/NU combinations (M = 0.067, SE = 0.003). In sum, the results indicated that there was an effect of utterance type for the action pictures, but not for the no-action pictures, for both the mean and peak pupil diameter. Namely, indirect requests, or the unique combination of action utterances with action pictures, resulted in a larger mean and peak pupil diameter than control combinations.

In the above analyses, we compared the pupil sizes for correct responses over the four pre-defined experimental conditions, or, picture-sentence combinations. In the following analyses we compared the pupil sizes for the two response types (indirect request vs. statement) regardless of the stimulus condition. In other words, we compared trials where participants did versus did not indicate that they had heard an indirect request. All trials ("yes" response = 1923, "no" response = 3757) from all 49 participants were included.

The preprocessing procedure was the same as for the first analysis, except that the data were split according to response rather than experimental condition (see Figure 4). The mixed effects model thus contained only one predictor: Response (indirect request versus statement). The random structure of the model was the same as in all other analyses and models were again compared using a likelihood ratio test.

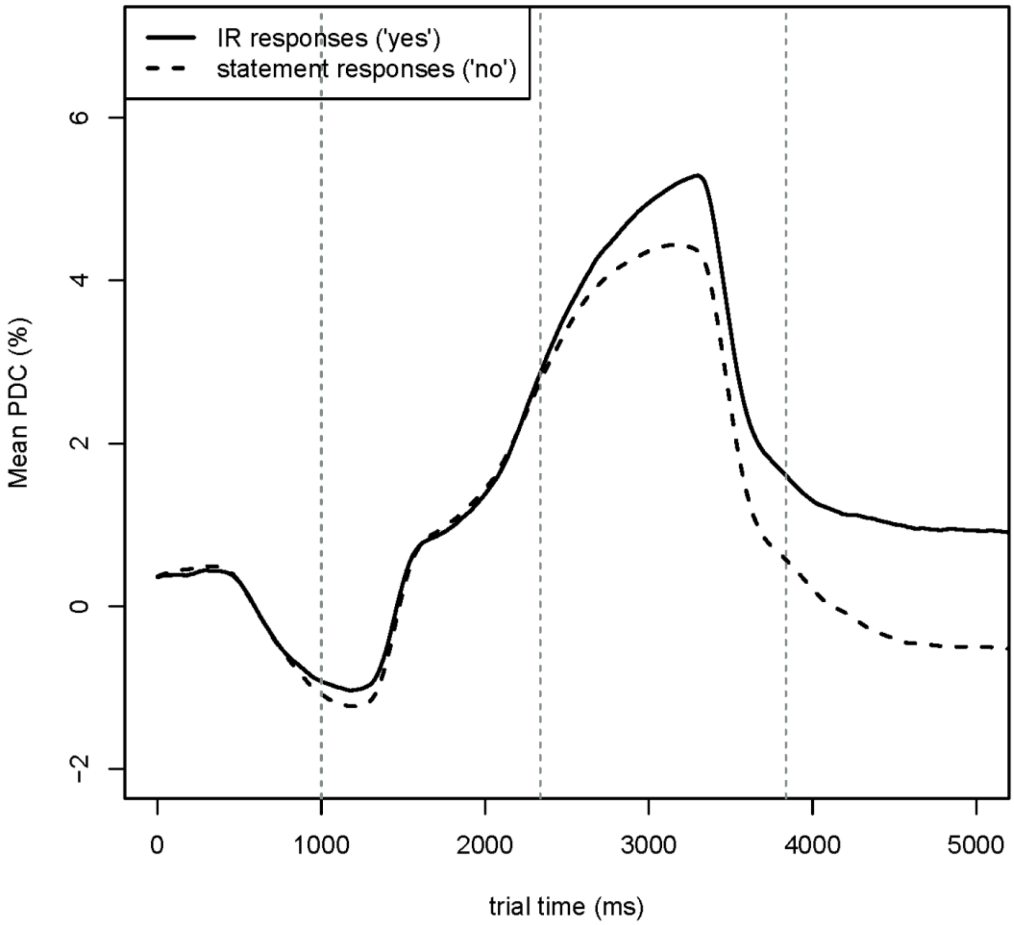

Figure 4. Experiment 1: Average percentage of pupil diameter change (PDC) as a function of response (indirect request versus statement), $n = 49$. Trial time (ms) is represented on the x-axis and pupil diameter change (%) on the y-axis. The vertical lines represent the different time windows: baseline (0 to 1000 ms), audio (1000 ms to critical word onset (M=2337)), critical (critical word onset + 1500 ms window after critical word onset) and end of trial (from end of the critical time window to end of trial).

Pupillometry results: response-split

Visual inspection of the PDC (see Figure 4) suggested a higher mean and peak pupil size for indirect requests as compared to statements. This was confirmed by the analysis. The optimal model for the mean pupil size during the critical time-window included the fixed effect of Response (IR versus statement), $\beta = 0.006$, SE = 0.001, $t = 3.263$. Indirect requests resulted in a larger mean pupil diameter as compared to statements (IR: $M = 0.042$, $SE = 0.001$, statement: $M = 0.039$, $SE = 0.001$). Including the effect improved the fit of the model (χ^2 (1) = 10.492, $p < .01$).

The model for the peak pupil size was also optimal with the inclusion of Response ($\beta = 0.008$, SE = 0.002, $t = 3.917$). Including the effect improved model fit (χ^2 (1) = 13.953, $p < .001$). Peak pupil size was larger for IRs ($M = 0.067$, $SE = 0.002$) as compared to statements ($M = 0.065$, $SE = 0.001$). In sum, these results indicated that response type, reflecting whether participants thought they heard an indirect request or a statement, predicted mean and peak pupil size, regardless of the stimulus conditions. Mean and peak pupil size were larger for indirect request as compared to statement responses.

Interim discussion

Based on the hypothesis that understanding indirect utterances requires additional inferences by the listener, we predicted a larger mean pupil size for indirect requests compared to control items (Holtgraves, 1994, 2008; Searle, 1975, 1979). In the analyses of trials with correct responses, this prediction was confirmed by an interaction between utterance and picture for the mean and peak pupil size in the 1.5 s time window following the critical word. In other words, the unique combination of action pictures with action utterances in the IR condition resulted in a larger mean and peak pupil size. The hypothesis that this increase in pupil size was related to additional inferences leading to the decision that a request was intended is supported by the second set of analyses, where the data set was split depending on the response type rather than the condition; larger mean and peak pupil sizes were observed for utterances classified as indirect requests than for statements. Thus, pupil size appears to capture the effort leading to the decision that a request was made.

An interesting finding of the current experiment is the discrepancy between the pupil data and the accuracy scores. For the accuracy scores, we only observed an effect of picture (action versus no-action), but no interaction with utterance. Participants were less accurate on trials containing an action picture, regardless of the type of utterance these pictures were combined with. This comprehension accuracy pattern is plausible since the action-pictures were designed to be more ambiguous. The trials on which these pictures appeared should allow for an indirect request interpretation or a statement interpretation. In contrast, the no-action pictures did not have to contain this ambiguity as they were designed to only accompany a statement. Consequently, participants were less accurate for the action pictures compared to the no-action pictures. However, this difference in response accuracy due to picture type was not reflected in the pupil dilation. Rather, larger mean and peak pupil size was observed for the action pictures in combination with an action utterance (AP/AU), not for the action picture combined with a no-action utterance (AP/NU).

Most probably, the accuracy scores reflected uncertainty in response selection, whereas pupil size captured the effect of processing an indirect request online. Comprehending and classifying a scene-sentence combination as an indirect request required online inferences which entailed additional processing effort, reflected in the increase in pupil diameter.

This hypothesis - that the increase in pupil size reflects processing effort related to interpreting an utterance as an indirect request - allows for a prediction for the pupillometry data of the entire group of participants, including those who did not pass the threshold of 70% correct responses for each condition. The prediction is that regardless of being correct or not, all indirect request responses should be associated with an increase in pupil size. The results of the response-split analyses support this hypothesis. Regardless of whether or not the response was correct, sentences classified as indirect requests were associated with a larger mean and peak pupil diameter as compared to statements, suggesting that only when participants made an inference, pupil diameter increased.

A possible confound for this interpretation is that in Experiment 1, participants always pressed the right button, labelled “yes”, to indicate that they heard an indirect request and the left button, labelled “no”, to indicate that they heard a statement. It is conceivable that the observed differences in pupil size for the response-split analysis were not related to differences in the processing of the utterances, but rather to making positive or negative decisions or to choice of a response hand. This mapping of indirect request responses to the dominant hand can also explain the shorter reaction times for indirect requests observed in the behavioral data (see Van Ackeren et al., 2012 for a similar explanation). To rule out the possibility that affirmative responses with the right hand could explain the differences in pupil diameter observed in the response-split analysis, a control experiment was conducted where responses were reversed such that a right-hand "yes" response indicated a statement and a left-hand "no" response an indirect request.

Experiment 2

Method

Participants

Twelve native speakers of Dutch participated in the study (1 male, mean age = 21.4 years, range = 18-25 years). All participants were right-handed, had normal hearing, normal or corrected-to-normal vision, and no history of language disorders. Informed consent was obtained from all participants. Ethical approval was granted by the ethics board of the Social Sciences Faculty of Radboud University. Participants were paid for taking part in the experiment.

Materials and design

Materials and design were the same as for Experiment 1.

Apparatus and procedure

The same equipment and procedure were used as in Experiment 1 except that participants now saw the sentence "The person made a statement." following each item. They were asked to push the right button (dominant hand) if they thought this was true ("yes" responses) and the left button if they thought this was not true ("no" responses). The experimenter indicated during the instruction that sometimes the speaker "means something more" with his statement, for example, he might be asking the listener, in an indirect way, to perform an action. Thus, participants were asked to respond "no" if they thought this was the case.

Analyses

Comprehension accuracy scores were analyzed using a one-sampled *t*-test to confirm that participants were able to discriminate between statements and indirect requests. The analysis of the reaction time was the same as in experiment 1.

Mean pupil size and peak pupil size were analyzed using linear mixed effects models in R (Baayen, Davidson, & Bates, 2008). All models contained the same random structure as in Experiment 1. For this control experiment, we only examined whether or not pupil size was related to the participants' response type ("yes" (statement) versus "no" (indirect request)). To assess the effect of Response (statement versus IR), models were compared via a backwards elimination, using a likelihood ratio test. The predictor was mean-centered.

Results and Discussion

Behavioral results

Requests were identified more often in the IR condition as compared to the control conditions (see Table 2). A one-sampled *t*-test (test value =.5) confirmed that participants were able to identify the indirect requests correctly in the AP/AU condition, $t(11) = 4.83$, $p < .01$) ($M = 76.11\%$, $SE = 5$).

Table 2. *Experiment 2: Percentage of IR responses ("no" to the statement "The person made a request") per condition*

	Indirect request (AP/AU)		*Picture Control (AP/NU)*		*Utterance control (NP/AU)*		*Picture-utterance control (NP/NU)*	
Condition	*%*	*SE*	*%*	*SE*	*%*	*SE*	*%*	*SE*
IR responses	76.11	5.41	29.17	5.51	3.00	11.94	11.94	2.86

Comprehension accuracy scores were slightly higher than in Experiment 1, especially for the control conditions. This might be due to the fact that in the current experiment participants were not explicitly asked to identify indirect requests, although they knew that indirect requests were present in the experiment. Possibly, participants tried less hard to identify sentences as indirect requests, resulting in lower false alarm rates.

The model for the reaction times revealed no reliable effects and/or interactions (all *t*-values < 1.4).

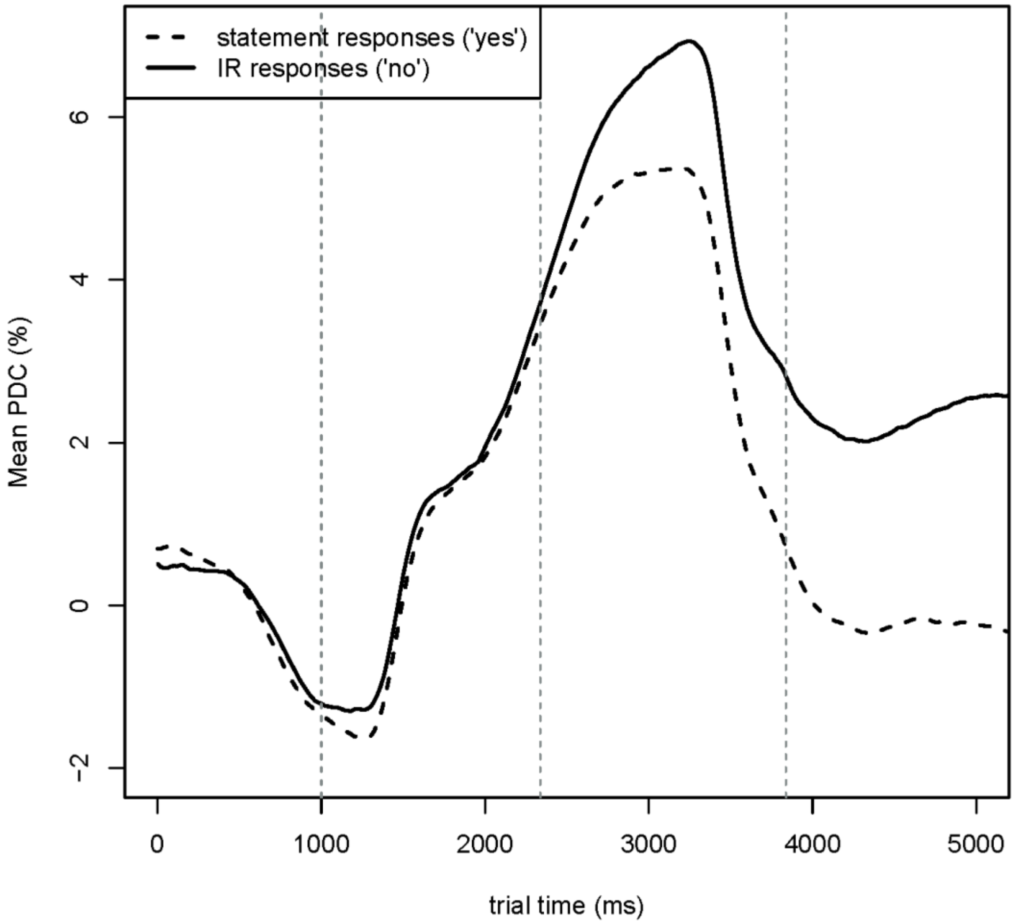

Figure 5. Experiment 2: Average percentage of pupil diameter change (PDC) as a function of response (statement versus indirect request), n =12. Trial time (ms) is represented on the x-axis and pupil diameter change (%) on the y-axis. The vertical lines represent the different time windows: baseline (0 to 1000 ms), audio (1000 ms to critical word onset (M=2337)), critical (critical word onset + 1500 ms window after critical word onset) and end of trial (from end of the critical time window to end of trial).

Pupillometry results: response-split

Visual inspection of Figure 5 shows that mean pupil diameter was larger for “no” responses (indirect requests) than “yes” responses (statements). This was confirmed by the analysis. The optimal model for the mean pupil size (n = 12) included the fixed effect of Response (statement versus IR), β = -0.010, SE = 0.014, t = 4.118. “No” responses (indirect requests) resulted in a larger mean pupil diameter as compared to “yes” responses (IR: M = 0.049, SE = 0.004, statement: M = 0.044, SE = 0.002). Including the effect improved the fit of the model (χ^2 (1) = 4.252, p < .05).

The model for the peak pupil size was also optimal with the inclusion of Response (β = -0.012, SE = 0.005, t = -2.308). Peak pupil size was larger for IRs (M = 0.079, SE = 0.004) as compared to statements (M = 0.072, SE = 0.003). The predictor improved model fit (χ^2 (1) = 5.012, p < .05).

In sum, as in Experiment 1, pupil size was larger when participants categorized the utterances as indirect requests compared to statements. This was true even though indirect request responses were now left-hand "no" responses. Together, the two response-split analyses indicated that the differences in pupil size between the indirect request condition and the control conditions can be related to the processing of the picture-utterance combinations rather than the choice of a response hand or the selection of an affirmative or negative response.

General discussion

Previous research has demonstrated that pupil diameter is sensitive to increases in processing demands as a consequence of, for example, higher memory load (e.g., Piquado, Isaacowitz, & Wingfield, 2010), syntactic anomalies (Schluroff, 1982), sentence complexity (Just & Carpenter, 1993), syntactic ambiguity (e.g., Engelhardt et al., 2010) and lexical ambiguity (Ben-Nun, 1986). To our knowledge the present study is the first to test whether pupil diameter is also sensitive to pragmatic factors, specifically the processing of indirect requests or direct statements. We presented participants with combinations of sentences and pictures, chosen such that in one out of four combinations the sentence could be interpreted as an indirect request, whereas in the remaining combinations the sentences were mere statements. In Experiment 1, participants were asked to make a yes-response with their right-hand when they thought they heard an indirect request and a no-response with their left-hand for statements. In Experiment 2 the response choices were reversed, such that right-hand affirmative responses were to be given to statements.

In a time window of 1.5 second following the critical word ("hot" or "nice" in "It is hot/nice here.") we observed a larger mean pupil diameter and a larger peak pupil size for indirect requests than for statements. As explained above, this was true when we compared the pupil size in the experimental conditions of Experiment 1 (indirect request versus control) and also, in both experiments, when we compared the pupil size for indirect request and statement responses regardless of the experimental condition. Thus, we observed an increase in pupil size whenever participants inferred that the utterance was an indirect request, which demonstrates the sensitivity of pupil size to our pragmatic manipulation. Measuring pupil diameter allowed us to observe a unique pattern for indirect requests compared to control statements. From a methodological point of view, this is an encouraging result, as it demonstrates that pupillometry can be used to study the processing of the pragmatic implications of utterances.

Evidently, pupillometry, like any other technique, has certain limitations. The most obvious ones are that participants have to wear eye-tracking equipment and that the visual environment must be carefully controlled in order to minimize changes in pupil sizes due to variations in luminosity (e.g., Janisse, 1977). In addition, the temporal resolution of pupil size measures may be seen as relatively poor, in comparison to, for instance, electroencephalography (EEG). Pupillary responses are relatively slow. In the present study the peak pupil size was reached on average 1000 ms after the onset of the critical word, and it took more than 3000 ms for the pupil diameter to return to baseline. However, although the pupil reaction was quite slow, the temporal resolution of the measure was sufficient for the purposes of the current study. Analyses of pupil diameter in the pre-defined critical window of 1.5 s after critical word onset allowed us to observe reliable differences as a result of our manipulation.

Our second aim was to investigate the cognitive effort involved in indirect request comprehension. In the current experimental setting, we observed an increase in mean and peak pupil diameter, reflecting an increase in processing demands for indirect requests as compared to control statements. This supports the view that identifying (and presumably understanding) non-conventional indirect requests is not an automatic process but requires processing effort beyond that needed to process mere statements. This conclusion is in line with behavioral studies on non-conventional indirect requests (e.g., Holtgraves, 1994) and with an Event-Related Potential (ERP) study on this type of indirect request. In this study, Coulson & Lovett (2010) observed transient processing costs for indirect requests in the form of a larger Late Positivity Component (LPC) on the final word for indirect requests as compared to literal statements. Based on our own findings and the findings by Van Ackeren et al. (2012), which were obtained with a similar stimulus set as used in the current experiment, we propose that the differences in pupil diameter observed for indirect requests compared to statements in our study reflect the cognitive effort involved in inference processes. Van Ackeren et al. (2012) observed increased activation in ToM areas (mPFC and left TPJ) for indirect requests compared to statements and proposed that these regions were important for making inferences about the communicative intents of the speaker. We suggest that the increases in pupil diameter in our data reflect this inference process as well. This interpretation is consistent with theories of pragmatic processing that postulate that drawing pragmatic inferences requires time and effort (e.g.,, Sperber & Wilson, 1995). However, it should be noted that in the paradigm used here, the requests were non-conventional (Holtgraves, 2002) and they were difficult to identify, as evidenced by the relatively low comprehension accuracy scores. Also, in natural conversations, interlocutors experience a shared context and have access to each other's speech, gesture, and facial expressions, which has been shown to influence language comprehension (e.g., Özyürek, Willems, Kita, & Hagoort, 2007; Van Berkum, Van den Brink, Tesink, Kos, & Hagoort, 2008). Moreover, Holtgraves (1994) demonstrated that information about the status of the speaker influenced indirect request comprehension. Thus, in everyday contexts identifying indirect requests may be easier or harder than in our laboratory setting, depending on the availability of more or less (non-linguistic) information. Finally, in the present study, participants were asked to provide explicit judgments concerning the implied meaning of the utterances. This is not the case in most everyday contexts and may have increased the processing load. Consequently, on the basis of the present data we cannot make claims about the processing costs incurred during indirect request processing in other contexts. However, we can say that in some situations identifying and understanding indirect requests entails processing effort beyond the effort needed to understand mere statements, and that this additional effort is reflected in the listeners' pupil size. An important direction for future research is to study the determinants of processing costs for different types of indirect utterances in different contexts. The current study has demonstrated that pupillometry is a useful tool in this endeavor.

References

Attar, N., Schneps, M., & Pomplun, M. (2013). Pupil Size as a Measure of Working Memory Load During a Complex Visual Search Task. *Journal of Vision*, *13*(9), 160-160.

Austin, J. L. (1962). *How to Do Things with Words.* Oxford: Oxford University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

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

Bašnáková, J., Weber, K., Petersson, K. M., Van Berkum, J., Hagoort, P. (2013). Beyond the language given: The neural correlates of inferring speaker meaning. *Cerebral Cortex*, *24*(10), 2572-2578.

Bates, D. Maechler, M., Bolker, B., Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*, 276-292.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of Psychophysiology*, *2*, 142-162.

Brown, P. & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge: Cambridge University Press.

Clark, H. H. (1979). Responding to indirect speech acts. *Cognitive Psychology*, *11*(4), 430-477.

Coulson, S., & Lovett, C. (2010). Comprehension of non-conventional indirect requests: An event related brain potential study. *Psychology*, *11*, 430-477.

Ben-Nun, Y. (1986). The use of pupillometry in the study of on-line verbal processing: Evidence for depths of processing. *Brain and Language*, *28*, 1-11.

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(5), 1704–9.

Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *The Quarterly Journal of Experimental Psychology*, *63*(4), 639-645.

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences*, *7*(2), 77-83.

Gibbs, R. W. (1981). Your wish is my command: Convention and context in interpreting indirect requests. *Journal of Verbal Learning and Verbal Behavior*,*20*(4), 431-444.

Gibbs, R. W. (1983). Do people always process the literal meanings of indirect requests?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(3), 524-533.

Gibbs Jr, R. W. (1986). What makes some indirect speech acts conventional?. *Journal of Memory and Language*, *25*(2), 181-196.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(3611), 1190-1192.

Holtgraves, T. (1994). Communication in context: Effects of speaker status on the comprehension of indirect requests. *Journal of Experimental Psychology: Learning Memory and Cognition*, *20*(5), 1215-1208.

Holtgraves, T. (2002). *Language as social action.* New Jersey: Lawrence Erlbaum Associates.

Holtgraves, T. (2008). Automatic intention recognition in conversation processing. *Journal of Memory and Language*, *58*(3), 627-645.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, *59*(4), 434-446.

Janisse, M. P. (1977). *Pupillometry: The psychology of the pupillary response*. Washington, DC: Hemisphere Publishing Corporation.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology-revue Canadienne De Psychologie Experimentale*, *47*, 310-339.

Just, M. A., Carpenter, P. A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science*, *4*(1), 56-88.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT press

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*(3), 560-569.

Porter, G., Troscianko, T., & Gilchrist, I. D. (2007). Effort during visual search and counting: Insights from pupillometry. *The Quarterly Journal of Experimental Psychology*, *60*(2), 211-229.

Otero, S. C., Weekes, B. S., & Hutton, S. B. (2011). Pupil size changes during recognition memory. *Psychophysiology*, *48*(10), 1346–53.

Özyürek, A., Willems, R., 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.

Schluroff, M. (1986). Pupillary responses to syntactic ambiguity of sentences. *Brain and Language*, *27*, 322-344.

Searle, J. R. (1975). Indirect speech acts. *Syntax and Semantics*, *3*, 59-82.

Searle, J. R. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(6), 679-692.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd edition). Oxford: UK Blackwell Publishers.

Van Ackeren, M. J., Casasanto, D., Bekkering, H., Hagoort, P., & Rueschemeyer, S. A. (2012). Pragmatics in action: indirect requests engage theory of mind areas and the cortical motor network. *Journal of Cognitive Neuroscience*, *24*(11), 2237-2247.

Van Berkum, J. J., Van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*(4), 580-591.

Van Rijn, H., Dalenberg, J. R., Borst, J. P., & Sprenger, S. A. (2012). Pupil dilation co-varies with memory strength of individual traces in a delayed response paired-associate task. *PloS One*, *7*(12), 1-8.

Wierda, S. M., van Rijn, H., Taatgen, N. A., & Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *Proceedings of the National Academy of Sciences*, *109*(22), 8456-8460.

Zellin, M., Pannekamp, A., Toepel, U., & van der Meer, E. (2011). In the eye of the listener: Pupil dilation elucidates discourse processing. *International Journal of Psychophysiology*, *81*(3), 133-141.

Chapter 3

The relation between pragmatic abilities and cognitive effort during indirect request processing

Abstract

When we speak, we often mean something more than what we literally say. For example, in a specific context the sentence "it's dark in here" might be a request to someone to turn on the light. Most listeners can extract this 'hidden' meaning very fast and accurately, but the way in which we do this is not well understood. In other fields of pragmatics, it has been shown that the pragmatic abilities of a listener may play a role in how pragmatic meaning is understood. In this study, we investigated the role of pragmatic abilities during indirect request processing. Pupil diameter was used as a measure of cognitive effort during indirect request processing and the communication subscale of the Autism Quotient as a measure of an individual's pragmatic abilities. We predicted that there should be a relation between pragmatic abilities and cognitive effort; pragmatically more skilled participants should need less cognitive effort when they understand indirect requests. Consistent with this prediction, we found some suggestive evidence for the relation between pragmatic abilities and cognitive effort during indirect request comprehension. However, future studies with more diverse participant groups will have to confirm this finding.

The relation between pragmatic abilities and cognitive effort during indirect request processing

When we speak we often mean something more or different than what we literally say. The communicative intention of the speaker, or *speaker meaning*, is frequently not lexically encoded and thus up to the listener to infer (Grice, 1989). According to speech act theory, sentences perform actions, and for successful communication the listener has to recognize these actions, or speech acts (e.g., Austin, 1962; Levinson, 1983; Levinson, 2017). However, since sentences are often underspecified when it comes to this action level of meaning, recognizing speech acts is not always straightforward (Levinson, 2013; Searle, 1975). Especially considering the extraordinary time pressure of spoken conversation (i.e. gaps between turns are in the order of 200-300 ms; Levinson & Torreira, 2015), it is remarkable that most individuals can understand speaker meaning, or recognize speech acts, very fast and accurately (Levinson, 2015). Nonetheless, the way in which we do this is not yet well understood. Researchers have only recently begun to address these issues, even though knowledge of these mechanisms seems crucial for a full account of human language and communication (Hagoort & Levinson, 2014).

In this study, we investigated a type of utterance in which the communicative intention is not lexically encoded, namely non-conventional indirect requests. Imagine a situation in which you sit in a coffee bar and realize that your coffee is cold. You could express your request for warmer coffee by saying "my coffee is cold" to the waiter. In most instances, he or she will understand this request without much trouble and bring you a new coffee, even though you did not literally ask him or her to do so. For this type of sentences, the context, including the discourse, and the beliefs and inferences on the part of the listener, contribute greatly to the meaning of the utterance (e.g., Hagoort & Levinson, 2014; Sperber & Wilson, 1987). Upon hearing a sentence like "my coffee is cold" in the context sketched above, the waiter has to make certain assumptions about the intention with which the speaker uttered this sentence, which requires some knowledge of the communicative and social use of language. Interestingly, in other areas of pragmatics it has been shown that individuals vary significantly in whether and how they apply this 'pragmatic' knowledge (e.g., Joliffe & Baron-Cohen, 1999; Kulakova & Nieuwland, 2016; Nieuwland, Ditmar, & Kuperberg, 2010; Pijnacker, Hagoort, Buitelaar, Teunisse, & Geurts, 2009; Schindele, Ludtke, & Kaup, 2008). Thus, individual differences in pragmatic abilities might also play a role in the processing and recognition of indirect requests. In the experiment described here, we investigated the relation between an individual's pragmatic abilities and cognitive effort during comprehension of non-conventional indirect requests.

Indirect request processing

The first experimental studies on indirect requests showed that listeners could recognize *conventional* indirect requests (e.g. “Could you bring me a new cup of coffee?”) fast and without first computing the literal meaning (Holtgraves, 1994). However, the type of indirect request described above, in which a negative state (i.e. the coffee is cold) is described, is less conventional and may involve a more elaborate inference process (Holtgraves, 1994). Neuroimaging studies have shown that several brain networks outside the standard language areas are engaged when people infer the meaning of this type of indirect speech (e.g., Bašnáková, Weber, Petersson, Van Berkum, & Hagoort, 2013; Van Ackeren, Casasanto, Bekeering, Hagoort, & Ruschemeyer, 2012; for an overview see Hagoort & Levinson, 2014). These networks include the mentalizing network, the Theory of Mind network (Bašnáková et al., 2013; Van Ackeren et al., 2012) and, for indirect requests specifically, the motor system (Van Ackeren et al., 2012; Van Ackeren, Smaragdi, & Rueschemeyer, 2016). In addition, experiments using more time-sensitive measures have revealed that electrophysiological brain responses to different speech acts (i.e. naming vs. requesting) diverged as early as 120 milliseconds after word onset, suggesting that access to pragmatic knowledge can be very early and possibly occurs in parallel with other types of linguistic processing (Egorova, Shtyrov, & Pulvermüller, 2013).

Several ‘bottom-up’ and ‘top-down’ factors have also been described that are involved during speech act recognition (Levinson, 2017). Bottom-up information can be found directly in the signal; it is coded or cued by means of lexical choice, construction of the sentence, or prosody (Levinson, 2017). For example, Hellbernd and Sammler (2016) showed that different speech acts have different prosodic patterns and that prosody plays an important role in differentiating speech acts (e.g., suggestion versus warning). In contrast, top-down information includes all the contextual and sequential information. Namely, the plan of the speaker, common ground between two speakers, and the location of the speech act in the sequence of turns. For example, Gisladottir, Chwilla, and Levinson (2015) found in an ERP experiment that when speech acts are embedded in conversational turns that are highly constrained, listeners are able to recognize the speech act of the sentences already before having heard the final word of the sentence. For more complex speech acts, additional processing might be required (Gisladottir, Chwilla, & Levinson, 2015). Importantly, in addition to the top-down and bottom-up factors described here, a listener's individual traits might also play a significant role during speech act recognition. With the current study we hope to provide insight into this non-linguistic property, the pragmatic abilities of the listener, which might play a role in the process of speech act recognition.

Individual differences in pragmatic abilities

In other fields of pragmatics, the role of individual differences in pragmatic abilities has already been investigated. Research on autism spectrum disorders (ASD) has revealed that people with ASD tend to interpret metaphors and irony literally (e.g., Happé, 1993). Even within this group, differences have been found between high-functioning adults with autistic disorder and adults with Asperger Syndrome in a pragmatic reasoning task (Pijnacker, Hagoort, Buitelaar, Teunisse, & Geurts, 2009). Pragmatic ability refers to the effective application of the knowledge of the social and communicative uses of language and it is often assessed with the Autism Quotient communication-subscale (Baron-Cohen et al., 2001). Also in non-patient populations, links have been observed between an individual's pragmatic abilities and pragmatic reasoning (e.g., Kulakova & Nieuwland, 2016; Nieuwland, Ditman & Kuperberg, 2010). For example, Nieuwland, Ditman and Kuperberg (2010) investigated whether pragmatic skills are involved during the comprehension of scalar quantifiers. In this electroencephalography (EEG) study, participants were first presented with underinformative statements (e.g., some people have *lungs* that are diseased by viruses) versus pragmatically more informative (e.g. some people have *pets* that are diseased by viruses) scalar statements and were then asked to do a sentence verification task. Pragmatic skill was assessed with the Autism Quotient communication-subscale (Baron-Cohen et al., 2001). Only for pragmatically skilled participants there was a difference between the critical word in underinformative statements (*lungs*), as compared to the critical word in the informative statements (*pets*). For this group, the critical word in the underinformative sentence elicited a larger N400 component, an index of semantic processing, as compared to the informative sentence. This suggests that pragmatically skilled participants were more sensitive to the scalar quantifier "some" during the formation of online expectations for the critical word in the sentence. Thus, only participants that were pragmatically more skilled were sensitive to the pragmatic 'violation' of underinformativeness (Nieuwland, Ditman, & Kuperberg, 2010). A similar link between the N400 and pragmatic abilities was found in a study by Kulakova and Nieuwland (2016) on pragmatic abilities and counterfactual language comprehension. Together, these studies suggest that individual differences in pragmatic skills can be linked to differences in online pragmatic processing. Specifically, participants with better pragmatic skills are more sensitive to pragmatic cues in the signal which can be used to build expectations about what will come next.

The present study

The present study explores the relation between pragmatic abilities and cognitive effort during indirect request comprehension. In Chapter 2 we showed that we can reliably tap into cognitive effort during indirect request comprehension by looking at differences in pupil diameter for indirect requests versus statements. Pupil diameter has also been used to study cognitive effort during language processing (e.g. Just & Carpenter, 1993; Engelhardt, Ferreira, & Patsenko, 2010; Zellin, Pannekamp, Toepel, & van der Meer, 2011). For instance, syntactically more ambiguous sentences were associated with a larger pupil dilation than syntactically unambiguous sentences (Engelhardt, Ferreira, & Patsenko, 2010). Also, object relative sentences (e.g., *The senator that the reporter attacked admitted the error*) required more cognitive effort, as indexed by pupil diameter increase, than subject relative sentences (e.g., *The senator that attacked the reporter admitted the error*; Just & Carpenter, 1993). These fluctuations in pupil diameter as a consequence of increases in processing difficulty can been linked to the locus coeruleus-norepinephrine (LC-NE) system (see Demberg & Sayeed, 2016 for an overview). In primates, strong correlations have been found between pupil size and activity in the LC, a small region in the brainstem (Aston-Jones & Cohen, 2005). In short, neurons in the LC emit NE, a neurotransmitter. This neurotransmitter can make it easier for neurons to synchronize, which in turn facilitates the functional integration of different brain regions. This might be beneficial when difficulties in processing are encountered (Demberg & Sayeed, 2016; Sara, 2009). Thus, fluctuations in pupil diameter can be considered a by-product of activity in the LC-NE system in order to provide a ‘cognitive boost’ when processing difficulties are encountered.

In Chapter 2, participants were presented with scene-sentence combinations that could either be interpreted as an (indirect) request (e.g., a picture of a window with the spoken sentence “it’s cold in here”) or as a statement (e.g., a picture of a window with the sentence “it’s nice in here”). They were asked to decide whether or not the speaker made a request. We found that pupil diameter increase (as compared to baseline) was larger when participants processed the indirect request as compared to the statement. Thus, correctly identifying a request required more cognitive effort than identifying a literal statement. This is theoretically relevant, since different processing models diverge on their predictions concerning the effort and time necessary to understand what a speaker intends to convey. For example, direct access theories predict that, given enough context, speaker meaning is processed fast and automatically and without additional processing effort (e.g., Gibbs, 1983). On the other hand, relevance theory predicts that additional effort is necessary to comprehend speaker meaning, since additional inferential steps have to made by the listener (e.g., Sperber & Wilson, 1995). In line with the latter account, we found in Chapter 2 that additional processing effort was necessary for indirect requests, at least for the type of requests that we used and in the experimental context that we presented them. However, it is conceivable that processing these indirect requests is not equally effortful for everyone. Processing costs might vary on an individual level, as a function of pragmatic abilities. Therefore, with the present experiment, we hope to provide a more nuanced view of cognitive effort involved in indirect request processing, by taking into account the role of individual differences in pragmatic skills.

In this study, we again presented participants with scene-sentence combinations that could be interpreted as requests or as statements, while we measured pupil diameter. After each combination, participants were asked to decide whether the sentence was meant as a request or not. In addition, participants filled in the Autism Quotient questionnaire (Baron-Cohen et al., 2001) and the score of the communication subscale (AQ_comm) of this questionnaire was used as a measure of pragmatic ability. A higher score on this scale indicates poorer pragmatic skills. First, we predicted that there should be a link between pragmatic ability and accuracy for the categorization of the sentences. Pragmatically less skilled participants should be less good at recognizing the requests than more skilled participants. Second, we predicted that *if* a request was correctly identified, this would require more cognitive effort for pragmatically less skilled participants than for participants with better pragmatic skills. If participants are less effective in the application of the knowledge of the social and communicative uses of language, the request would result in more processing difficulty and would thus require a larger boost, reflected in an increase in pupil diameter, than if participants are pragmatically more skilled. Third, in line with the results reported in Chapter 2, we predicted that *if* the request was correctly identified, pupil diameter should overall be larger for requests than for statements, since requests require more cognitive effort, because the indirect meaning has to be retrieved to correctly categorize it. To clarify, mean pupil diameter in this study is the change in pupil diameter as compared to a pre-trial baseline (see the pupillometry pre-processing procedure). That is, when conditions are compared, this change can be larger for one condition than the other in a critical window after stimulus presentation. However, throughout this article we will use ‘mean pupil diameter’ instead of ‘mean pupil diameter change’ to make it easier to discuss differences between conditions and changes as a function of pragmatic abilities. Thus, increases in pupil diameter signify that processing difficulty was encountered and that performing the task required more cognitive effort from participants (i.e. they needed a ‘cognitive boost’, resulting in pupil dilation).

Method

Participants

Thirty-two right-handed native speakers of Dutch participated in the experiment (7 male, mean age = 22.13 years, range = 18-30). All had normal or corrected-to-normal vision and normal hearing. Participants gave informed consent prior to participation and were paid for taking part in the experiment. Ethical approval for the study was granted by the ethics board of the Social Sciences faculty of Radboud University (project code: ECSW2014-1003-196). One participant was excluded from the analysis since he/she did not correctly fill in the AQ questionnaire.

Materials and design

A subset of the materials constructed by Van Ackeren et al. (2012) was used, consisting of 120 sentences and 60 pictures. The design is shown in Figure 1. Each picture (e.g. a picture of a window) was combined with two sentence types: a potential indirect request (e.g. "it is very hot here") and a statement (e.g. "it is very nice here"). Utterances were recorded from a native speaker of Dutch. Since pupil size is sensitive to luminosity, the luminosity values of the pictures were adjusted so that each picture had a value between 140 and 160. Picture size was kept small (250 x 250), such that the larger part of the screen was white.

Each participant saw all 60 pictures twice, once combined with an indirect request (60 items) and once with a statement (60 items). There were always at least 45 trials between repetition of a picture. Sentences were never repeated. This resulted in 120 unique trials, divided over four blocks. Within blocks, items were pseudo-randomized and the same condition was never presented more than three times in a row. The order in which participants heard the two utterances combined with the picture was counterbalanced (i.e. for one item set (e.g., Figure 1), participant 1 first heard the indirect request and then the statement, and vice versa for participant 2). Between each block, participants were encouraged to take a break. Before the experimental blocks, participants completed 12 practice trials.

Picture	Utterance	Condition	number of items
	It is very hot in here *action*	Indirect Request (IR)	60
	It is very nice in here *no-action*	Statement	60

Figure 1. Example item set.

Apparatus and procedure

Participants sat in a medium-lit sound proof booth. The lighting was kept constant for all participants. Experiment Builder Version 1.10.1025 (SR Research, Ltd., Mississauga, Canada) was used to present the stimuli. Pupil diameter (right eye) was measured with an Eyelink 1000 Tower Mount eye-tracker (SR Research Ltd., Missauga, Canada). Sentences were presented through Sennheiser HD201 lightweight over-ear binaural headphones. A buttonbox was used to record button presses.

The time-course of a trial is presented in Figure 2. A fixation cross was presented for 1000 milliseconds (ms) at the beginning of the trial. Then, a picture was displayed and 200 ms later participants heard a sentence. The fixation cross appeared again for 2500 ms after the sentence followed by the statement "The person made a request" and participants then indicated whether or not they recognized a request by pressing a button. Response hand was counterbalanced between participants. After the participant's response, a fixation cross was presented for 2500 ms, so that the pupil size could return to baseline before the next trial.

After the experiment, participants filled out the Autism Questionnaire (Baron-Cohen et al., 2001). The Autism Questionnaire (AQ) is a self-administered questionnaire comprised of 50 questions, including 10 questions for each of five subscales (i.e., social skill, attention switching, attention to detail, communication, and imagination). Half of the items were phrased to produce a ‘disagree’ response and the other half an ‘agree’ response. Items were randomized in terms of the expected response and in terms of the subscale they assess (Baron-Cohen et al., 2001). The AQ communication subscale (henceforth “AQ_comm”) was of most significance for this experiment, since it describes a person's communicative and pragmatic abilities (e.g., item 7: “Other people frequently tell me that what I've said is impolite, even though I think it is polite”) and it has previously been found to correlate with individual responses during other types of pragmatic processing (Kulakova & Nieuwland, 2016; Nieuwland, Ditman, & Kuperberg, 2010).

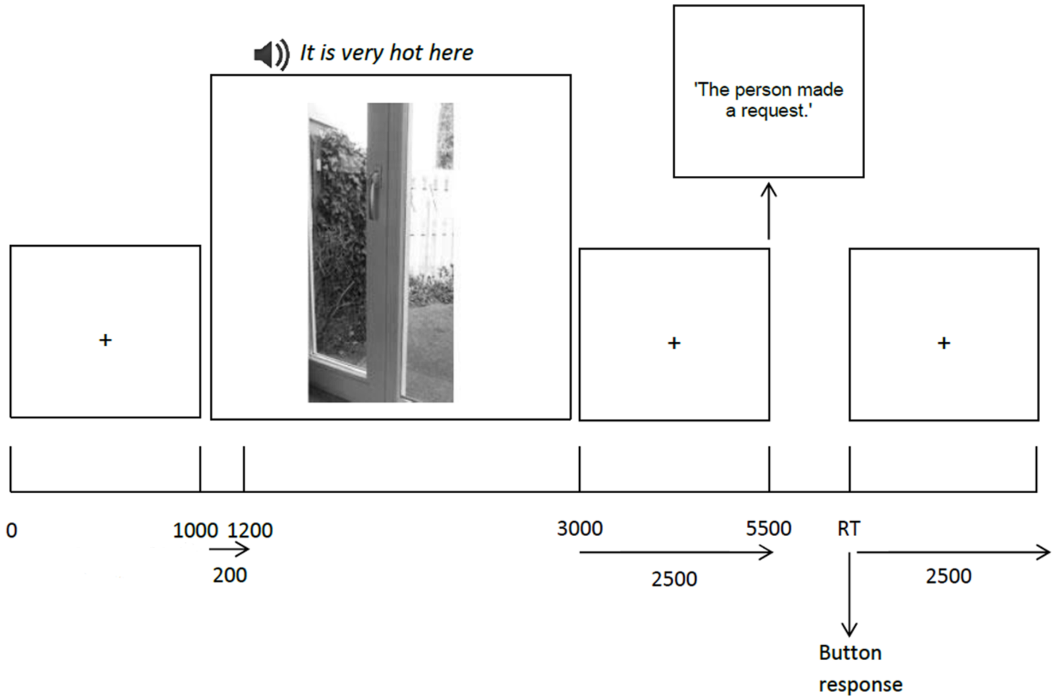

Figure 2. Time-course of a single trial (time in ms).

Pupillometry pre-processing procedure

The pupillometry data was pre-processed in R (version 3.0.3), using a procedure developed by Gerakaki, Sjerps, and Meyer (in prep.). The procedure was the same as used in Chapter 2. Pupil diameter was down-sampled to 50 Hz and the change in pupil diameter was analyzed from sample to sample to detect and remove outliers. All data points with a ratio that was one standard deviation higher or lower than the mean pupil change of the trial were considered outliers, and linear interpolation was used to replace them (for a similar procedure, see Piquado, Isaacowitz, & Wingfield, 2010). If more than 25 percent of data points were missing for a trial, the trial was removed. For each participant, absolute pupil diameter was baseline-corrected and normalized, resulting in the relative pupil diameter. Pupil size of a given trial was subtracted from each sample in the trial and then divided by the baseline to calculate the pupil diameter change. The baseline was the first window of a trial (1000 ms), in which a fixation cross was presented on the screen. Relative pupil diameter was used to compare between participants, filtering out certain tonic changes in pupil diameter (e.g. Van Rijn, Dalenberg, Borst, & Sprenger, 2012). For statistical analysis, trials were divided into four parts: baseline (0 to 1000 ms), audio (1000 ms to critical word onset (*M*=2337)), critical (1.5 ms window after critical word onset), and end of trial (from end of critical window to trial offset) (for a similar procedure, see Engelhardt, Ferreira, & Patsenko, 2010; Tromp, Hagoort, & Meyer, 2016). To determine the critical window, we took a time window of 1.5 seconds starting from critical word onset, adopting the 1.2 seconds window of Just and Carpenter (1993), plus 300 ms for word recognition (see Engelhardt et al., 2010, for a similar approach). Statistical analyses on the pupil size were only performed for the mean pupil diameter in the critical window. The relative pupil diameter multiplied by 100 was used to plot the Task-Evoked Pupillary Responses (TEPRS), which represent the average pupil diameter change (%) over time during the trial.

Analyses

Two types of analyses were performed on the data. First, correlations between pragmatic abilities and performance (i.e. categorization accuracy) were calculated for the IR condition and the statement condition separately. Furthermore, a correlation was calculated between pragmatic abilities and mean pupil diameter. For this analysis, a difference score was calculated for the pupil diameter by subtracting the score for the statement condition from that of indirect request condition. We decided to perform the correlations with a difference score rather than the absolute score for the indirect request condition directly, so that the statement condition could function as a baseline, and the correlation would not be due to a more general relation between pragmatic abilities and pupil diameter during language processing.[2]

Second, to provide a more nuanced analysis of the effects of condition and pragmatic abilities on the outcome measures, linear mixed effects regression models were used to further analyze the data (Baayen, Davidson, & Bates, 2008). R was used for all analyses (version 3.0.3; The R foundation for statistical computing; lme4 package; Bates, Mächler, Bolker, & Walker, 2014). The accuracy data were analyzed with logit mixed models (Jaeger, 2008). The full model included a fixed effect of Condition (IR, Statement), pragmatic ability

[2] To control for the possibility that pragmatic ability is correlated with pupil diameter irrespective of task demands or experimental conditions, we correlated pragmatic abilities with mean pupil diameter in the baseline window (in which a fixation cross was presented on the screen). This correlation was not significant (p = .90).

score, and the interaction of the two. Both predictors were mean-centered. The model had the maximum possible random effects structure, consisting of random intercepts and slopes for Condition by participant and by item (i.e., Picture). This was the maximum random structure leading to convergence (Barr, Levy, Scheepers, & Tily, 2013). Models were compared using a likelihood ratio test. The mean pupil diameter data was analyzed with linear mixed effects models (lme4 package; Bates, Mächler, Bolker, & Walker, 2014). The structure of the models was the same as for the accuracy data, and models were again compared using a likelihood ratio test.

Results

AQ scores

AQ scores were available for 31 participants. The overall AQ scores ranged from 7 to 25 (M = 15.81, SE = 4.87) and the AQ_comm score ranged from 1 to 6 (M =2.55, SE = 1.50).

Accuracy correlation

The average percentage of correct responses from all participants can be found in the upper row of Table 1. On average participants had high accuracy rates. There was no difference in performance between the IR condition and the statement condition (p = .09). Furthermore, there was no correlation between AQ_comm and accuracy for the IR condition (p = .86), nor for the statement condition (p = .24).

Table 1. *Percentage of correct responses per condition (IR, statement) for the entire sample (n = 31) and the pupillometry subset (n = 27).*

	n	Indirect Request (IR)	Statement
Percentage correct all (SE)	31	81.18 (0.91)	81.61 (0.09)
Percentage correct (SE)	27	86.73 (0.84)	84.07 (0.91)

Four participants were excluded from the pupillometry analyses below, since they did not have accuracy rates of > 70% for each of the two conditions. The accuracy rates for the remaining group of participants ($n = 27$) are displayed in the bottom row of Table 1. For this subgroup, there was also no correlation between AQ_comm and accuracy for the IR condition ($p = .62$), nor for the statement condition ($p = .11$). There was no difference between conditions in performance ($p = .15$).

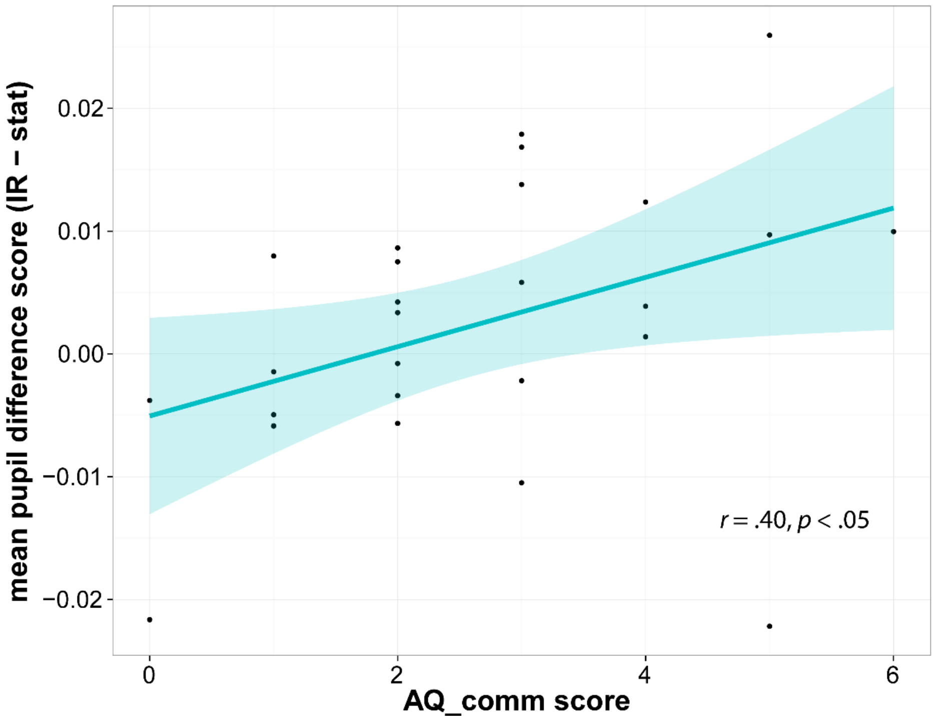

Figure 3. Scatterplot for the correlation between AQ_comm score (x-axis) and mean pupil size difference score (y-axis).

Pupillometry correlation

Incorrect trials were removed (14.60% of the data) before the pupillometry analysis, because we were interested in the relation between pragmatic abilities and cognitive effort only when indirect requests were correctly categorized. The difference between pupil diameter for the IRs versus the statements ranged from -0.022 to +0.026 ($M = +0.0023$, $SE = 0.0004$).

There was a significant positive correlation ($r(25) = .40$, $p < .05$) between the pupil difference score (IR - statement) and AQ_comm score (see Figure 3). People with higher scores on the AQ_comm (i.e., poor pragmatic skills) had a larger positive difference score (i.e., a larger increase in pupil diameter for the indirect requests as compared to statements). In other words, people with good pragmatic skills needed no or less additional cognitive effort to process indirect requests as compared to statements, whereas people with poor pragmatic skills did.

Linear mixed-effects models

The following analyses were conducted only for the subset of participants (n = 27) with accuracy scores of 70% or higher. The model for accuracy did not reveal an effect of Condition (β = -0.044, SE = 0.397, z = 0.111, p = .912). Furthermore, adding the predictor AQ_comm did not improve model fit (β = -0.053, SE = 0.111, z = 0.481, p = .630, χ^2 (1) = 0.229 p = .633). Finally, including the interaction between Condition and AQ_comm did also not improve the fit of the model (β = -0.142, SE = 0.220, z = -0.646, p = .518, χ^2 (1) = 0.413 p = .520). In sum, there was no evidence for an effect of Condition and/or pragmatic abilities score on accuracy. This result was in accordance with the absence of a correlation between the accuracy scores and pragmatic abilities.

For the pupillometry analysis, incorrect trials were removed. For the mean pupil diameter, there did not seem to be a difference between conditions, based on visual inspection of the TEPRs for the complete sample (see Figure 4). We fitted models for the mean pupil diameter in the critical window of interest. In the first model for mean pupil diameter, there was no evidence for an effect of Condition (β = 0.003, SE = 0.003, z = 0.885). We then included the predictor AQ_comm, which did not improve model fit (β = 0.003, SE = 0.002, z = 1.157, χ^2 (1) = 1.245, p = .265). Finally, including the Condition by AQ_comm interaction marginally improved model fit (β = 0.003, SE = 0.001, z = 1.740, χ^2 (1) = 3.030, p = .0812). To visualize this trend, we did a tertile split to show the effect of Condition on the pupil diameter for three levels of AQ_comm score (low, medium, high). As can be seen in Figure 5, the trend was similar to the direction of the correlation. People with high AQ_comm scores seemed to show larger pupil dilation for indirect requests than for statements, while this was not the case for pragmatically more skilled participants (i.e. those with low AQ _comm scores). We did not perform any tests on the groups created by the tertile split, since there were not enough trials to provide sufficient power for statistical analyses.

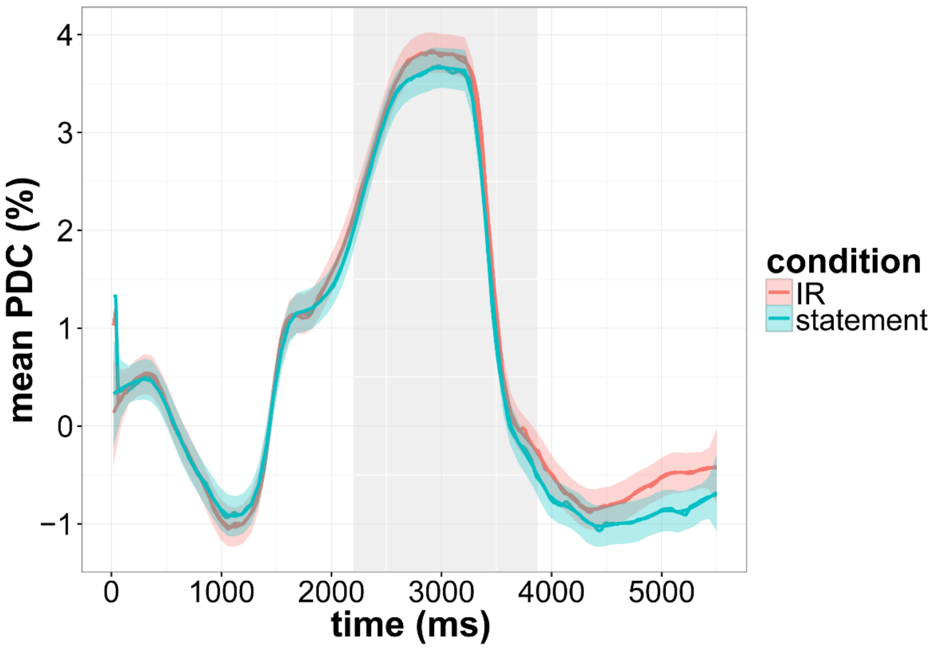

Figure 4. Mean pupil diameter change (PDC) as a function of condition (IR versus statement). Trial time (ms) is on the x-axis and pupil diameter change on the y-axis. The ribbon represents ±SE.

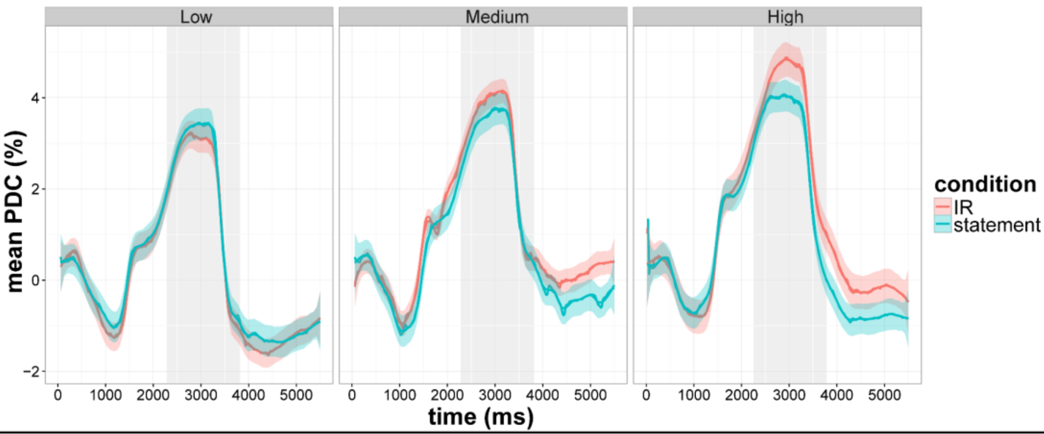

Figure 5. Tertile split based on AQ _comm score (low, medium, high) of average pupil diameter change (PDC) as a function of condition (IR versus statement). Trial time (ms) is on the x-axis and pupil diameter change on the y-axis. The ribbon represents SE. The grey section marks the critical window.

Reaction times

During the experiment, participants were asked to decide whether or not the sentences they heard were requests or not. However, they were told to withhold their response for 2500 ms after sentence offset to reduce influences of physical responses on the pupillometry data. Thus, the response was not speeded and the reaction times (RTs) obtained in this experiment should not be considered a very reliable measure of response time. However, for completeness, we describe the reaction time data and correlation with pragmatic abilities here. Again, incorrect trials (14.60%) were removed for the RT analysis. The mean RT for the indirect requests was 480.28 ms (*SE* = 5.61 ms) and for the statements 467.22 ms (*SE* = 5.66 ms). Difference scores for the correlations were calculated by subtracting the RT for the statement condition from the RT for the indirect request condition. Difference scores ranged from -38.68 to 115.35 ms (*M* = 51.89 ms, *SE* = 1.36 ms). There was no correlation between the RT difference scores and AQ_comm scores (*p* = .23).

Discussion

Our aim was to investigate whether there is a relation between pragmatic abilities and cognitive effort during indirect request comprehension. In the experiment, participants looked at pictures (e.g., a picture of a door) while they heard sentences that could either be interpreted as indirect requests (e.g., "it is cold here") or statements (e.g., "it is nice here"). They were asked to categorize the sentence as a request or a statement. First, we predicted a link between pragmatic ability and accuracy for sentence categorization. Pragmatically less skilled participants should be less good at recognizing the requests than participants with better pragmatic abilities. This prediction was not confirmed; there was no correlation between accuracy scores and pragmatic ability for the IRs or for the statements. It should be noted that participants had relatively high accuracy scores. In comparison to Chapter 2, participants were much better able to recognize requests. In Chapter 2, we removed 27 participants (55.1%) because they did not have accuracy scores higher than 70% for one or more conditions, while in this experiment we removed only 4 (12.5%). This might be due to the fact that in the current experiment 60 trials (50%) required a request response, while in Chapter 2 this was only 30 trials (25%), which might have made it easier for participants to make a decision. Hence, it is conceivable that we did not find a correlation between pragmatic abilities and accuracy due to the overall high performance and little variability for the accuracy rates.

Our second hypothesis was that *if* a request was correctly identified, this would require more cognitive effort for pragmatically less skilled participants than for participants with better pragmatic skills. This hypothesis was confirmed by the correlation analysis. There was a positive correlation between pragmatic abilities score and pupil diameter difference for IRs as compared to statements. Participants with lower pragmatic abilities scores needed a larger boost, as reflected by pupil diameter increase, than pragmatically more skilled participants. However, the results of the linear mixed-effects model revealed only a marginal increase in the fit of the model with the inclusion of an interaction between pragmatic abilities and sentence type. Consequently, we have to be careful in drawing strong conclusions from data described here. There are several reasons for why we did not detect a significant interaction in the linear mixed-effects model. First, there was not a lot of variability in the AQ_comm scores, since all participants came from a very homogenous population (i.e., students between the age of 18 and 30 from Radboud University Nijmegen). Furthermore, the sample size was small, especially for analysis of individual variability.

Finally, we expected to replicate the results of Chapter 2, namely that *if* a request was correctly identified, pupil diameter should overall still be larger for requests than for statements, since requests require more cognitive effort, because the indirect meaning has to be retrieved to correctly categorize it. However, this prediction was not confirmed; across participants, pupil diameter was not larger for indirect requests than for statements. Again, this might be due to the differences in the design between this experiment and that in Chapter 2. Here, 50% of combinations required an indirect request response, as opposed to 25% in the Chapter 2. It is conceivable that the differences observed in Chapter 2 were diminished in the current experiment, since participants encountered more indirect requests, making it easier for them to process these utterances. In fact, for scalar inferences it has been shown that tweaks to materials and the experimental context in which they are presented can cause reliable changes in processing time (Grodner, Klein, Carbary, & Tanenhaus, 2010). Also, it has previously been shown that pragmatic effects can attenuate over time as participants get more practice (Noveck, 2016). Since participants were presented with structurally similar items within each condition and the same clearly delineated response options for each trial, it is very plausible that they got better over time.

In sum, this study provided some initial evidence for a link between pragmatic abilities and cognitive effort during indirect request comprehension. We observed a significant correlation between pragmatic skills and cognitive effort for indirect requests and a marginally significant effect in the same direction in the linear mixed-effects analysis. These findings contribute to our knowledge of how we recognize speech acts even for very underspecified linguistic utterances, a process that is necessary for successful communication (e.g., Levinson, 2017). Understanding the intentions of others is a crucial component of social interaction and it is important for future studies in experimental pragmatics to recognize the role of individual variation in the machinery that makes this possible. Apart from bottom-up information in the signal and top-down contextual information (Levinson, 2017), non-linguistic factors should be considered during theorizing and experimentation in this field. Future studies with larger and more diverse participant groups or clinical populations are necessary to confirm the results presented in this study. Furthermore, complementary methods can be used to assess pragmatic abilities, such as the Assessment of Pragmatic Abilities and Cognitive Substrates (APACS), which has been successfully used to assess pragmatic abilities in clinical populations (Arcara & Bambini, 2016).

References

Arcara, G., & Bambini, V. (2016). A test for the Assessment of Pragmatic Abilities and Cognitive Substrates (APACS): Normative data and psychometric properties. *Frontiers in Psychology*, 7.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17.

Austin, J. L. (1962). *How to Do Things with Words.* Oxford: Oxford University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

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

Bašnáková, J., Weber, K., Petersson, K. M., Van Berkum, J., Hagoort, P. (2013). Beyond the language given: The neural correlates of inferring speaker meaning. *Cerebral Cortex*, *24*(10), 2572–2578.

Bates, D. Mächler, M., Bolker, B., Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7.

Demberg, V., & Sayeed, A. (2016). The frequency of rapid pupil dilations as a measure of linguistic processing difficulty. *PloS One*, *11*(1), e0146194.

Egorova, N., Shtyrov, Y., & Pulvermüller, F. (2016). Brain basis of communicative actions in language. *NeuroImage*, 125, 857–867.

Engelhardt, P. E., Ferreira, F., & Patsenko, E. G. (2010). Pupillometry reveals processing load during spoken language comprehension. *Quarterly Journal of Experimental Psychology*, *63*(4), 639–645.

Gibbs, R. W. (1983). Do people always process the literal meanings of indirect requests? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(3), 524–533.

Gisladottir, R. S., Chwilla, D. J., & Levinson, S. C. (2015). Conversation electrified: ERP correlates of speech act recognition in underspecified utterances. *PloS one*, *10*(3), e0120068.

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

Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, *116*(1), 42–55.

Hagoort, P., & Levinson, S. C. (2014). Neuropragmatics. In *The Cognitive Neurosciences* (pp. 667–674). Cambridge, MA: MIT Press.

Happé, F. G. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, *48*(2), 101–119.

Hellbernd, N., & Sammler, D. (2016). Prosody conveys speaker's intentions: Acoustic cues for speech act perception. *Journal of Memory and Language*, *88*, 70–86.

Holtgraves, T. (1994). Communication in context: Effects of speaker status on the comprehension of indirect requests. *Journal of Experimental Psychology: Learning Memory and Cognition*, *20*(5), 1215–1208.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Jolliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory: Linguistic processing in high-functioning adults with autism or Asperger syndrome: is local coherence impaired? *Cognition*, *71*(2), 149–185.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *47*, 310–339.

Kulakova, E., & Nieuwland, M. S. (2016). Pragmatic skills predict online counterfactual comprehension: Evidence from the N400. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(5), 814–824.

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology,* 6.

Levinson, S. C. (2017). Speech acts. In Y. Huang (Ed.), *Oxford handbook of pragmatics* (pp. 199–216). Oxford: Oxford University Press.

Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language*, *63*(3), 324–346.

Noveck, I. (2016). On investigating intention in experimental pragmatics. *Pre-proceedings of Trends in Experimental Pragmatics*, 103–109.

Zellin, M., Pannekamp, A., Toepel, U., & van der Meer, E. (2011). In the eye of the listener: Pupil dilation elucidates discourse processing. *International Journal of Psychophysiology*, *81*(3), 133–141.

Pijnacker, J., Hagoort, P., Buitelaar, J., Teunisse, J. P., & Geurts, B. (2009). Pragmatic inferences in high-functioning adults with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *39*(4), 607–618.

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*(3), 560–569.

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, *10*(3), 211–223.

Schindele, R., Lüdtke, J., & Kaup, B. (2008). Comprehending negation: A study with adults diagnosed with high functioning autism or Asperger's syndrome. *Intercultural Pragmatics*, *5*(4), 421–444.

Searle, J. R. (1975). Indirect speech acts. *Syntax and Semantics*, *3*, 59–82.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd edition). Oxford: UK Blackwell Publishers.

Van Ackeren, M. J., Casasanto, D., Bekkering, H., Hagoort, P., & Rueschemeyer, S. A. (2012). Pragmatics in action: Indirect requests engage theory of mind areas and the cortical motor network. *Journal of Cognitive Neuroscience*, *24*(11), 2237–2247.

van Ackeren, M. J., Smaragdi, A., & Rueschemeyer, S. A. (2016). Neuronal interactions between mentalising and action systems during indirect request processing. *Social Cognitive and Affective Neuroscience*, *11*(9), 1402–1410.

van Rijn, H., Dalenberg, J. R., Borst, J. P., & Sprenger, S. A. (2012). Pupil dilation co-varies with memory strength of individual traces in a delayed response paired-associate task. *PLoS One*, *7*(12), e51134.

Zellin, M., Pannekamp, A., Toepel, U., & van der Meer, E. (2011). In the eye of the listener: Pupil dilation elucidates discourse processing. *International Journal of Psychophysiology*, *81*(3), 133–141.

Chapter 4

The combined use of virtual reality and EEG to study language processing in naturalistic environments

Abstract

When we comprehend language, we often do this in rich settings in which we can use many cues to understand what someone is saying. However, it has traditionally been difficult to design experiments with rich three-dimensional contexts that resemble our everyday environments, while maintaining control over the linguistic and non-linguistic information that is available. Here we test the validity of combining electroencephalography (EEG) and Virtual Reality (VR) to overcome this problem. We recorded electrophysiological brain activity during language processing in a well-controlled three-dimensional virtual audiovisual environment. Participants were immersed in a virtual restaurant, while wearing EEG equipment. In the restaurant participants encountered virtual restaurant guests. Each guest was seated at a separate table with an object on it (e.g. a plate with salmon). The restaurant guest would then produce a sentence (e.g. "I just ordered this salmon."). The noun in the spoken sentence could either match ("salmon") or mismatch ("pasta") with the object on the table, creating a situation in which the auditory information was either appropriate or inappropriate in the visual context. We observed a reliable N400 effect as a consequence of the mismatch. This finding validates the combined use of VR and EEG as a tool to study the neurophysiological mechanisms of everyday language comprehension in rich, ecologically valid settings.

Tromp, J., Peeters, D., Meyer, A. S., & Hagoort, P. (2018). The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behavior Research Methods, 50*(2), 862-869.

The combined use of virtual reality and EEG to study language processing in naturalistic environments

In everyday life, we often communicate about the things in our immediate environment. The information we can use to understand what someone is saying therefore often extends beyond words. For example, when visiting a restaurant we may listen to a friend talking about the food on her plate and the drinks on the table. We use visual information to understand what is being said. Consequently, realistic models of language comprehension should be able to explain language processing in this and many other types of contextually rich environments. Unfortunately, this is not always the case. In a recent overview, Knoeferle (2015) argues that psycholinguistic theorizing has been mostly “language- centric”. Most models (e.g., Friederici, 2002; Bornkessel & Schlesewsky, 2006) can explain a range of semantic and syntactic processes very well, but it is more difficult to derive hypotheses from them about how people comprehend language when they can use all sorts of information from the non-linguistic environment (Knoeferle, 2015). One reason for the limited number of models with predictions on language processing in rich ‘real-life’ contexts is that it is experimentally challenging to test them. It is difficult to design experiments with rich three-dimensional contexts that resemble our everyday environments, while maintaining control over the linguistic and non-linguistic information that is provided. It becomes even more difficult if neurophysiological methods like electroencephalography (EEG) are used, which require strict control over the linguistic and non-linguistic input and are sensitive to many non-relevant signals from the environment. Here we test the validity of combining Virtual Reality (VR) and EEG to overcome this problem.

A virtual environment is a digital space in which sensory experiences are recreated and a user’s movements can be tracked (Fox, Arena, & Bailenson, 2009). VR can be used to create a three-dimensional world in which people can move and interact, which makes it a very suitable method to study psychological and social phenomena (Fox et al., 2009). By offering the possibility to recreate very complex, rich, everyday environments, VR allows researchers to increase the ecological validity of a study, while maintaining full experimental control. It makes it possible to study behavior in different environments, without interference from uncontrollable cues and it allows manipulations of variables that have traditionally been hard to replicate or control in the lab (Blascovich et al., 2002; Blascovich & Bailenson, 2011; Fox et al., 2009). Also, since virtual environments are often very engaging, they can be considered a motivational tool (Bayliss & Ballard, 2000). Finally, the use of virtual agents provides a good alternative to the use of human confederates, which is often problematic (Kuhlen & Brennan, 2013).

VR and EEG have been successfully combined, for instance, to study driving behavior (Bayliss & Ballard, 2000), spatial navigation (e.g. Bischof & Boulanger, 2003) and spatial presence (Baumgartner, Valko, Esslen, & Jänke, 2006). However, we are not aware of any studies that combine VR and EEG to study language behavior. The reason for this might be the assumption that human-computer interactions are necessarily different from human-human interactions. This could be problematic if one wants to study everyday language behavior. However, recent evidence suggests that this is not the case. In a study by Heyselaar, Hagoort, and Segaert (2015), participants performed the same syntactic priming task with a human confederate and with a human-like virtual agent and showed comparable priming effects in both situations. In addition, it has been shown that people adapt their speech rate and pitch to a virtual interlocutor in the same way as they do with a human interlocutor (Casasanto, Jasmin, & Casasanto, 2010; Gijssels, Casasanto, Jasmin, Hagoort, & Casasanto, 2016). Thus, VR has proven to be a useful tool to study language processes on a behavioral level. With the experiment proposed here we hope to extend this to the neurophysiological level. As a proof of concept, we used VR and EEG to study language comprehension in an engaging visually rich three-dimensional environment. In particular, we investigated electrophysiological brain responses to mismatches between visual and auditory information.

In our experiment people were immersed in a rich virtual environment (VE), a restaurant, while wearing EEG equipment. Several virtual agents (henceforth "restaurant guests") were seated at different tables in the restaurant and participants were moved through the restaurant from table to table. Upon arrival, the participant looked at the object on the table in front of the guest (e.g. a plate with salmon), after which the guest produced a sentence (e.g. "I just ordered this *salmon*"). The noun in the sentence could either match ("*salmon*") or mismatch ("*pasta*") with the object on the table, creating situations in which the auditory information was either appropriate or inappropriate with respect to the visual context. Thus, if successful, this setup will allow us to investigate electrophysiological brain activity during the simultaneous processing of auditory and visual information in a well-controlled three-dimensional virtual environment.

Although not in VR, previous studies have used designs comparable to the one used here to investigate the neural correlates of language processing in an audiovisual context. For example, in a study by Peeters, Hagoort and Özyürek (2015), participants viewed static pictures while they heard sentences that could either match or mismatch with the information in the picture. For instance, participants saw a picture of a woman pointing at a mango, while they heard a sentence including either a matching noun (e.g. "I have just found this *mango* in the cupboard") or a noun that did not match the visual information (e.g. "I have just found this *spoon* in the cupboard"). Incongruency between the spoken word and the physical object in the visual scene was reflected in an enhanced N400. The N400 is an event-related potential (ERP) component that peaks around 400 milliseconds (ms) after onset of a critical stimulus. The N400 has been linked to meaning processing and is sensitive to a wide variety of stimuli, including spoken and written words, objects and sounds. Several theories exist on the functional significance of the N400 component (see Kutas & Federmeier, 2011 for an overview). One view is that the N400 reflects semantic integration (Brown & Hagoort, 1993; Hagoort, Baggio, & Willems, 2009). Semantic integration is the process by which listeners use the global semantic representation from the context to immediately integrate the meaning of upcoming words into the overall message representation (Hagoort, 2003). In everyday language comprehension the brain combines meaningful information from incoming speech with information about objects in the visual environment that are in the current focus of attention. Willems, Özyürek and Hagoort (2008), for instance, investigated the neural integration of words and pictures into a preceding sentence context. In their ERP experiment, participants heard a word (e.g. flower) and saw a picture (e.g. a picture of a flower) that had to be integrated with the previous sentence context (e.g. "The man gave his wife a nice..."). Pictures and words could either fit well (e.g. flower) or less well (e.g. cherry) with the previous sentence context. If the item presented did not match the previous sentence context well, an N400 effect was observed. This effect was very similar for pictures and words in terms of latency and amplitude, suggesting that no differentiation between verbal and visual semantic information was made at this level of processing. In addition, there was an effect in an earlier window (225-325 ms), which was also not specific to the picture or word conditions (Willems, Özyürek, & Hagoort, 2008).

In addition to pictures, researchers have used videos to provide visual context to investigate semantic processing in more real-world environments (e.g. Sitnikova, Kuperberg, & Holcomb, 2003; Sitnikova, Kuperberg, West, & Holcomb, 2006). Sitnikova et al. (2008) presented participants with movie clips of everyday events (e.g. cutting bread). The clips consisted of a context (e.g. "a man placing a cutting board on a kitchen counter and then placing a loaf of bread on the cutting board") and a final scene. The final scene could either match the previous scene (e.g. "the man cuts off a piece of bread with a knife"), or violate goal-related action requirements (e.g. "the man slides an electric iron across the loaf of bread"), or be completely unexpected (e.g. "the man uses an electric iron to press wrinkles from his pants"). Importantly, both mismatch conditions resulted in a larger N400 than the match condition. Furthermore, an early semantic congruency effect was observed in the N300 window (250-350 ms). The authors suggested that this N300 effect reflected the fast access that visual images have to semantic memory networks (see also Sitnikova et al., 2006; McPherson & Holcomb, 1999). Finally, when the goal-related action requirement was violated (i.e. the ironing scene), a posterior late positivity was observed (Sitnikova, Holcomb, Kiyonaga, & Kuperberg, 2008). Although this experiment did not investigate the integration of visual and auditory information, as the violations occurred within the visual domain, the results offer predictions for the latency and distribution of ERP effects when participants that are looking at a non-static environment.

In the field of gesture and sign language research the use of videos is common, since semantic processing here critically hinges on the visual information provided (Andric & Small, 2012; Dick, Mok, Beharelle, Goldin-Meadow, & Small, 2014; Özyürek, 2014). For example, Özyürek, Willems, Kita, and Hagoort (2007) investigated the online integration of semantic information from speech and gesture. Participants listened to sentences with a critical verb (e.g. "He slips on the roof and *rolls* down"), combined with a video of an iconic gesture (e.g. a rolling gesture). The verbal and/or gestural semantic content could either match ("rolls" and a rolling gesture) or mismatch ("walks" and a walking gesture) with the part of the sentence before the critical verb (e.g. "He slips of the roof and..."). The results revealed an effect in the N400 window for both gestural and spoken mismatches, suggesting that information from both modalities is integrated at the same time.

Although the use of videos to study language comprehension in context is already a step away from using static pictures on a computer screen, it still has certain limitations that could be overcome by exploiting recent advances in VR technology. First, videos provide only a two-dimensional scene on a very small computer screen, while in VR participants experience a very large realistic three-dimensional environment. Furthermore, in VR it is possible for participants to look at a dynamic speaker and even interact with him or her, rather than just observing a person on a screen. Recently it has been argued that to study the brain basis of interaction, we should move away from passive spectator science to studies with engaged participants (Hari, Henriksson, Malinen, & Parkkonen, 2015). VR is a useful method to do so, provided that reliable effects can be observed in an environment that is much more complex and dynamic but also more distracting than a simple computer screen. With the experiment described here, we aimed to test the feasibility of combining VR and EEG to study language comprehension in a rich setting. Based on the studies mentioned above, we predicted an N400 effect for our study as well. The amplitude in the N400 window should be more negative for the noun in the mismatch condition (e.g. "I just ordered this *pasta*" when there is a piece of *salmon* on the table) than in the match condition (e.g. "I just ordered this *salmon*" when there is a piece of *salmon* on the table). Finding an N400 effect would validate the combined use of VR and EEG as a tool to study everyday language comprehension in rich, ecologically valid settings, thereby paving the way for future experimental studies of the neurophysiological mechanisms involved in everyday language use.

Method

Participants

Twenty-three participants (two males) with an average age of 21 years (range 18-26 years) participated in the experiment. All were right-handed native speakers of Dutch, had normal or corrected-to-normal vision and normal hearing and had no history of speech problems or neurological diseases. Participants provided written informed consent and were paid to participate in the experiment. Ethical approval for the study was granted by the ethics board of the Social Sciences Faculty of Radboud University. Two participants were excluded from the analysis due to technical failure during the experiment. Data from one additional participant was excluded because too many trials (> 30 percent per condition) had to be discarded due to EEG artifacts.

Materials and design

The experiment took place in a Virtual Environment (VE) that was custom-made using Vizard (version 4.08, WorldViz, Santa Barbara). It consisted of a restaurant with eight tables in one row and a virtual restaurant guest sitting at each table (see Fig. 1).

Figure 1. Screenshot of the Virtual Environment (VE).

Participants passively moved from table to table through the restaurant via a pre-programmed procedure (i.e. they did not physically walk themselves). This procedure was chosen to reduce the amount of movement artifacts in the EEG and to control the amount of time a participant was able to look at the object on each table. All restaurant guests resembled Caucasian males or females between the ages of 25 and 35, in line with the age, gender, and background of the speakers who recorded the sentences. The restaurant guests kept a neutral facial expression throughout the experiment. The voices of the virtual guests were rendered with a stereo speaker set.

Materials consisted of 80 objects and 96 sentences. There were 80 experimental sentences and 16 fillers. On each trial, participants saw an object on the table in the VE (see below) and then heard a sentence from a restaurant guest seated at the table. All sentences and objects were relevant to a restaurant setting. The sentences (e.g. the Dutch equivalent of "I just ordered this salmon.") were paired with objects (e.g. a plate with salmon) so that the critical noun in the sentence could either match (e.g. "salmon") or mismatch (e.g. "pasta") with the object on the table. The demonstrative preceding the noun always matched with the gender of the noun corresponding to the object on the table and with the noun spoken by the virtual agent (which differed in the mismatch condition). The filler sentences were general statements that could be uttered in a restaurant setting, but did not refer specifically to an object in the virtual environment (e.g. "I always come here for lunch"). During presentation of the filler sentences, a generic cup, plate, or bowl was presented on the table. The sentences were recorded by eight native speakers of Dutch (four male), had an average duration of 1973 ms (*SD* = 354) and were equalized in maximum amplitude in the speech analysis package Praat (version 5.1; www.praat.org). The onset of the critical noun was determined in Praat. The experimental sentences had ten different sentence frames (e.g. "Ik heb deze … net besteld.", "*I have just ordered this ...* "*).* Each speaker used each sentence frame only once and each frame was present in each round (or block) only once. Half of the sentences were presented in the match condition and half in the mismatch condition. This was counterbalanced across participants, which resulted in two lists. Sentences were never repeated for one participant. Objects were repeated once, with a minimum of 32 trials (four blocks) between two presentations of the same object.

Procedure

Participants were seated in a chair while they wore an EEG cap and an NVIS nVisor SX60 head-mounted display (HMD). The HMD presented the Virtual Environment (VE) at a 1280x1024 resolution, with a 60 degrees monocular field of view. Eight reflective markers were mounted onto the HMD, which were linked to a passive infrared DTrack 2 motion tracking system (ART Tracking, Munich). The data from this system was used to update the participant's viewpoint when the participant moved. Prior to entering the VE, participants were told that they would move through a restaurant and that the guests in the restaurant would say something to them. They were instructed to pay close attention to the objects on the tables and to what the restaurant guests said. To familiarize participants with the food and drinks served in the virtual restaurant, they were asked to look at the menu of the restaurant prior to the start of the experiment, which contained all the objects, and their labels, that could be presented in the VE.

The trial sequence was as follows: from the beginning of the trial participants “arrived” at the table in two seconds (i.e. the movement took 2 sec.). Upon arrival, the participant had four seconds to look at the object on the table. Then, the restaurant guest looked up and two seconds later he or she began to speak. At the end of the sentence, the participant was moved backwards again automatically. Before the start of the experiment, participants were instructed to keep eye contact with the restaurant guest from the moment the guest looked up to the end of the sentence. They were also encouraged not to blink their eyes during this period.

Participants made twelve rounds through the restaurant. During each round, each restaurant guest said one sentence, resulting in eight sentences per round. After each round participants were encouraged to take a short break. Before the first experimental round, the participant completed a practice round, in which they were moved past each table. During this round participants could get used to the movement and they were encouraged to practice looking up at the restaurant guest and not blinking while they had eye-contact with him/her. There were no objects on the table during the practice round and the restaurant guests only looked up and did not speak.

After the experiment participants were asked to complete two questionnaires. The first evaluated whether they paid attention during the experiment. It contained eight statements; four about the sentences (e.g. "An avatar said that he/she always comes here for breakfast.") and four about the objects (e.g. "One of the objects in the restaurant was a mandarin."). Participants were asked to choose "true", "false" or "I don't know". The percentage of correct responses was calculated on the basis of the "true" and "false" responses. If participants filled in "I don't know" (5% for the object questions, 3.75% for the sentence questions), this was not counted as a response. The aim of the second questionnaire was to assess the participant's perception of the virtual agents. It consisted of eight questions about the appearance and behavior of the restaurant guests (e.g. "How human-like did you find the avatars?"). Participants were asked to response on a scale from 1 (not human-like) to 7 (very human like).

EEG recording and analysis

The electroencephalogram (EEG) was continuously recorded from 59 active electrodes held in place by an elastic cap (see Fig. 2 for the equidistant electrode montage). In addition to the electrodes in the cap, three external electrodes were attached; one below the left eye to monitor for blinks and one on the lateral canthus to the side of each eye to monitor for horizontal eye movements. Finally, two electrodes were placed over the left and right mastoid respectively. Electrodes were referenced online to the electrode placed over the left mastoid and offline to the average of the left and right mastoids. Electrode impedance was kept below 20 KΩ. The EEG was recorded with a low cut-off filter of 0.01 Hz and a high cut-off filter of 200 Hz at a sampling rate of 500 Hz. A high-pass filter at 0.01 Hz and a low-pass filter at 40 Hz were applied offline. Brain Vision Analyser software (Version 2.0.2, Brain Products, Munich) was used to process the EEG. Epochs from 100 ms preceding the onset of the critical noun to 1200 ms after the critical noun were selected. Trials containing ocular artifacts were excluded (8.88% in the match condition, 9.63% in the mismatch condition; not statistically different). The 100 ms period preceding the critical noun was used as a baseline. Average event-related potentials (ERPs) were calculated per participant and condition in three time-windows. In addition to the N400 window (350-600 ms), an earlier window (250-350 ms) was included based on previous studies that observed early effects as a result of visual or audiovisual mismatches (e.g. Willems et al., 2008; Sitnikova et al., 2008; Peeters et al., 2015). Finally, a 200-milliseconds window after the N400 window was analyzed (600-800) to test for the presence of a sustained N400 effect. Repeated measures analyses of variance (ANOVAs) were performed in the different time-windows with the factors Condition (match, mismatch), Region (vertical midline, left anterior, right anterior, left posterior, left anterior) and Electrode. The Greenhouse-Geisser correction (Greenhouse & Geisser, 1959) was applied to all analyses with more than one degree of freedom in the numerator, the adjusted values are reported.

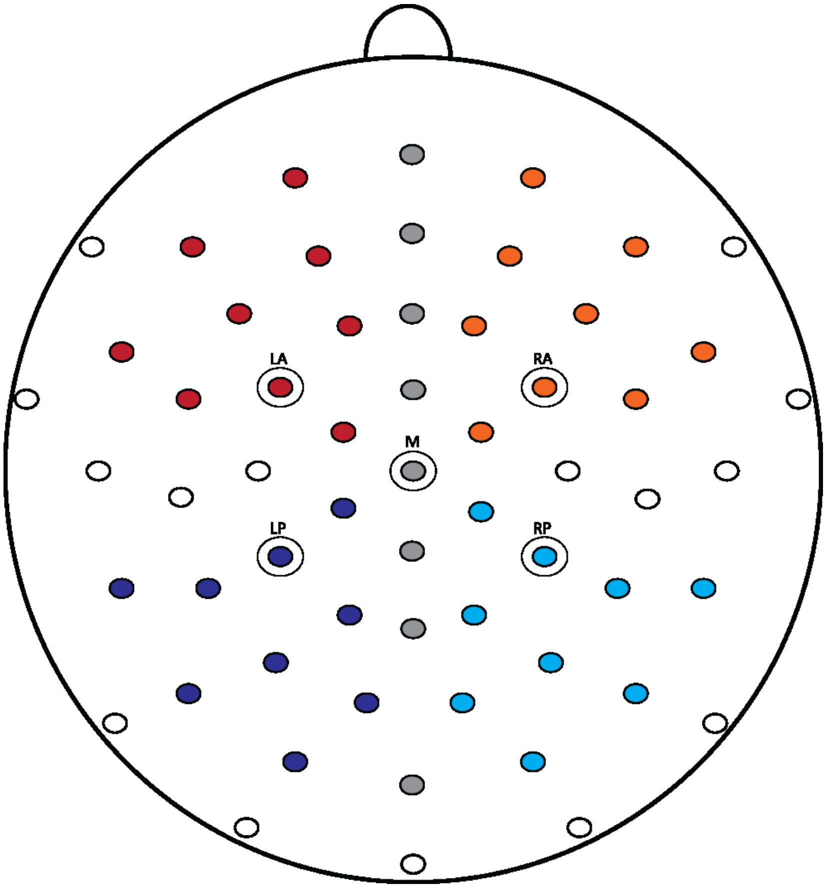

Figure 2. Equidistant electrode montage. The electrode sites displayed in Figure 3 are encircled (left anterior, LA; right anterior, RA; midline, M; left posterior, LP; right posterior, RP). The five regions used in the analysis are highlighted in different colors (LA: red; RA: orange; M: grey; LP: dark blue; RP: light blue).

Results

On average participants answered 86.46% (*SE* = 0.90%) of the questions correctly in the attention questionnaire. They scored 77.92% (*SE* = 4.54%) on the object questions (e.g. "One of the objects in the restaurant was a mandarin.") and 95.00% (*SE* = 2.29%) on the questions about the sentences (e.g. "An avatar said that he/she always comes here for breakfast"). The results from the second questionnaire indicated that the restaurant guests were rated as relatively human like (*M* = 4.6, *SE* = 0.06).

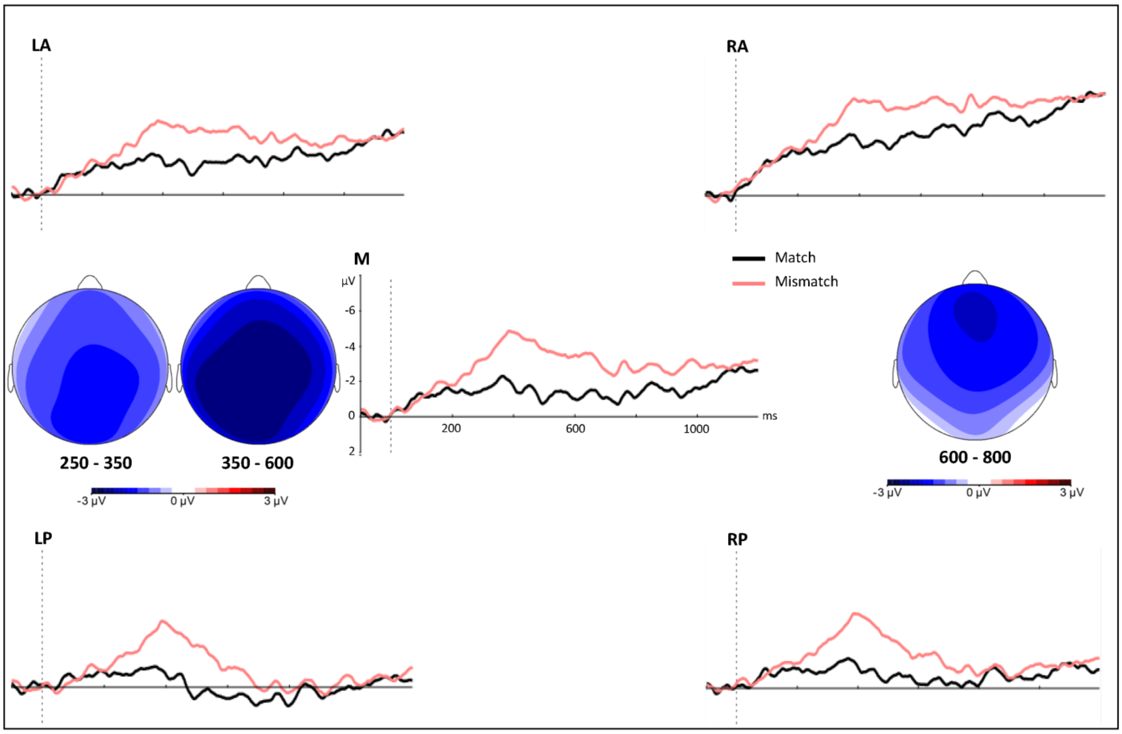

Figure 3. Grand-average waveforms time-locked to the onset of the critical noun in the match and the mismatch condition. The topographic plots display the voltage differences between the two conditions (mismatch - match) in the three different time-windows.

Figure 3 displays the grand average waveforms time-locked to the onset of the critical noun. The ANOVA for the early time-window (250-350 ms) revealed a significant main effect of Condition ($F(1,19) = 6.22$, $p = .03$, partial $\eta^2 = .25$). ERPs were more negative for the mismatch condition ($M = -2.42$ μV, $SE = .32$) than for the match condition ($M = -1.34$ μV, $SE = .46$). This effect was not modulated by Region ($F < 2$).

In the N400 window (350-600 ms) there was also a significant main effect of Condition ($F(1,19) = 18.03$, $p = .001$, partial $\eta^2 = .49$), with a more negative ERP for the mismatch condition ($M = -2.99$ μV, $SE = .42$) than the match condition ($M = -1.30$ μV, $SE = .58$). The effect was widespread, confirmed by the absence of a Condition by Region interaction ($F < 2$).

Visual inspection of the waveforms in Fig. 3 indicated a continuation of the N400 effect in the 200 ms epoch right after the standard N400 window (600-800 ms). This observation was confirmed by the ANOVA for this window, which revealed a main effect of Condition ($F(1,19) = 10.10$, $p < .01$, partial $\eta^2 = .35$). Amplitude was still more negative for the mismatch condition ($M = -2.33$ µV, $SE = .67$) than for the match condition ($M = -1.23$ µV, $SE = .61$). This effect was not modulated by Region ($F < 3$).

Discussion

The aim of this study was to test the validity of the combined use of VR and EEG to study language comprehension in a visually rich context. Participants were immersed in a virtual environment, a restaurant, in which virtual restaurant guests were seated at tables with food or drinks in front of them. The guests produced sentences that could match (e.g. "I just ordered this salmon") or mismatch (e.g. "I just ordered this pasta") with the object on the table (e.g. a piece of salmon). As a result of this manipulation, we observed a reliable N400 effect, in line with our predictions. This shows that VR and EEG combined can be used to study language comprehension in realistic three-dimensional environments. Neither the VR helmet (head-mounted display) that was placed over the EEG cap, nor the noise caused by the VR equipment, limited us in acquiring a reliable EEG signal. There were not more artifacts, due to movement or blinking, than in an average EEG study. The rich virtual environment was not too distracting for the participants, as they paid attention to the restaurant guests and objects and they judged the restaurant guests as human-like. It should be noted that the percentage of correct answers in the attention questionnaire was lower for the objects than for the sentences, which might suggest that participants did not pay enough attention to the objects. We believe, however, that this difference was due to the fact that participants were presented with the menu of the restaurant, which contained all objects, prior to seeing a subset of the objects in the actual experiment. Thus, they might have remembered objects from the menu rather than from the experiment itself, which resulted in the higher percentage of errors.

In all time-windows, ERPs were more negative for the mismatch condition as compared to the match condition. Importantly, in the N400 window there was a widely distributed ongoing negativity similar in onset latency and distribution to the effects observed in previous studies that investigated the integration of visual and auditory information (e.g. Peeters et al., 2015; Willems et al., 2008). This negativity extended into the 600 to 800 ms time-window. The extended nature of the N400 effect in our study could simply be a carryover effect from the strong N400 effect (e.g. Willems et al., 2008) or it could reflect the extended presentation time of the incongruous information. In our study participants were able to see the object even after the restaurant guest had already stopped speaking, which resulted in a more prolonged negativity than the typical N400 effect evoked by short presentation of written or spoken words (Sitnikova et al., 2008). Finally, ERPs were more negative for the mismatch than the match condition in an early time-window (250-350 ms). In Sitnikova et al., (2008) the negativity in this window was interpreted as a separate N300 effect, reflecting the rapid access of visual information to semantic memory networks. However, since in our study the mismatching information came from the speech signal (in the context of visual information), it is unlikely that this account holds for the present data. Rather, the effect resembles early effects observed in other studies investigating mismatches in auditory speech processing (e.g. Connoly & Philips, 1994; Hagoort & Brown, 2000). In these studies it has been suggested that a negativity in this window is an indication of a mismatch between the expected word forms, based on the context, and the activated lexical candidates generated on the basis of the speech signal (Phonological Mismatch Negativity or PMN, Connoly & Philips, 1994). In our experiment, participants could build up a strong expectation or prediction for the word form of the upcoming noun on the basis of the visual context (i.e. they saw the object on the table well in advance). In addition, for most of the stimuli a mismatch could already be detected at the first segment of the noun (in 98.96 % of our item sets the onset of the mismatching noun was different from the word form expected on the basis of the visual context). Thus, it is very probable that the negativity observed in the early window (250-350 ms) was due to a mismatch between the expected and the encountered word form.

Although the present study was successful in providing evidence for the reliability of the combined use of VR and EEG, it has certain limitations. First, in a few cases there was some difficulty in setting up the EEG cap and VR helmet. The head-mounted display used in this study was meant to fit relatively tightly around the head, which in some instances made it more challenging to use it in combination with an EEG. More recently developed head-mounted displays (e.g. the Oculus Rift) are lighter and more flexible than the one used in the current study, which will allow for longer experiments and reduced EEG preparation time preceding the start of an experiment. Moreover, limitations of head-mounted displays can easily be overcome by using virtual reality equipment (such as a CAVE system) that does not necessarily makes use of a head-mounted display but has participants wear 3D shutter glasses to experience immersion into a virtual environment. Finally, because of the combination with EEG, the virtual environment could not be used to its full potential. In real life, people move their head, look around and interact with the environment, which is all possible in VR as well. However, in our experiment such behavior was restricted because of the sensitivity of the EEG to movement artifacts.

The combination of VR and EEG has the potential to address several under-researched questions in the field of psycholinguistics and neurobiology of language. It can be used to study how we comprehend language when we use multiple sources of information in our environment, which is necessary for the development of more complete models of language processing (Knoeferle, 2015). Also, it can shed light on how we listen and speak in interactive real-world situations. The need for a shift away from spectator science and toward more interactive and realistic paradigms to study the human brain and human behavior is also echoed in other fields of neuroscience. Social interaction plays a central role in human brain function and it has been argued that studies in social neuroscience should shift their focus toward studies with engaged participants and dynamic stimuli (Hari et al., 2015; see also Willems, 2015). Along similar lines, Schilbach and colleagues (2013) highlight the necessity of studying real-time social encounters in an interactive manner. VR is very well suited to help us understand how we interact with others (virtual agents, avatars, or humans) during real-time communication. Research into the electrophysiology of language comprehension has been virtually "speakerless", which has left the social, pragmatic and dynamic functions of communication severely under-researched (Hoeks & Brouwer, 2014). VR provides a way to include a well-controlled speaker in our experiments to study aspects of language and communication in a more natural, dynamic way, even in combination with electrophysiological recordings.

Acknowledgements:
We are very grateful to Jeroen Derks for programming the virtual restaurant and to Albert Russel for technical support. We thank Annelies, Annika, Bart, Ellen, Erik, Linda, Rick and Stef for giving the restaurant guests a voice. We thank two anonymous reviewers for their valuable comments.

References

Andric, M., & Small, S. L. (2012). Gesture's neural language. *Frontiers in Psychology, 3*.

Baumgartner, T., Valko, L., Esslen, M., & Jäncke, L. (2006). Neural correlate of spatial presence in an arousing and noninteractive virtual reality: an EEG and psychophysiology study. *CyberPsychology & Behavior, 9*(1), 30-45.

Bayliss, J. D., & Ballard, D. H. (2000). Single trial P3 epoch recognition in a virtual environment. *Neurocomputing*, *32*, 637-642.

Bischof, W. F., & Boulanger, P. (2003). Spatial navigation in virtual reality environments: an EEG analysis. *CyberPsychology & Behavior*, *6*(5), 487-495.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, *13*(2), 103-124.

Blascovich, J., & Bailenson, J. (2011). *Infinite reality: Avatars, eternal life, new worlds, and the dawn of the virtual revolution*. New York, NY: William Morrow & Co.

Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages. *Psychological Review*, *113*(4), 787.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*(1), 34-44.

Casasanto, L., Jasmin, K., & Casasanto, D. (2010). Virtually accommodating: Speech rate accommodation to a virtual interlocutor. In *32nd Annual Meeting of the Cognitive Science Society (CogSci 2010)* (pp. 127-132). Cognitive Science Society.

Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, *6*(3), 256-266.

Dick, A. S., Mok, E. H., Beharelle, A. R., Goldin-Meadow, S., & Small, S. L. (2014). Frontal and temporal contributions to understanding the iconic co-speech gestures that accompany speech. *Human Brain Mapping*, *35*(3), 900-917.

Gijssels, T., Casasanto, L. S., Jasmin, K., Hagoort, P., & Casasanto, D. (2016). Speech accommodation without priming: The case of pitch. *Discourse Processes*, *53*(4), 233-251.

Fox, J., Arena, D., & Bailenson, J. N. (2009). Virtual reality: A survival guide for the social scientist. *Journal of Media Psychology*, *21*(3), 95-113.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78-84.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95-112.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia*, *38*(11), 1518-1530.

Hagoort, P. (2003). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *Neuroimage*, *20*, S18-S29.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences, 4th ed.* (pp. 819-836). Cambridge, MA: MIT Press.

Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of social interaction in human brain function. *Neuron*, *88*(1), 181-193.

Heyselaar, E., Hagoort, P., & Segaert, K. (2015). In dialogue with an avatar, language behavior is identical to dialogue with a human partner. *Behavior Research Methods*, 1-15.

Hoeks, J. C., & Brouwer, H. (2014). Electrophysiological Research on Conversation and Discourse. Holtgraves, T. (Ed.), *Oxford Handbook of Language and Social Psychology*, pp 365-386. Oxford: Oxford University press.

Knoeferle, P. (2015). Language comprehension in rich non-linguistic contexts: combining eye tracking and event-related brain potentials. In R.M. Willems (Ed), *Cognitive Neuroscience of Natural Language Use* (pp. 77-100). Cambridge: Cambridge University Press.

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: when confederates might be hazardous to your data. *Psychonomic Bulletin & Review*,*20*(1), 54-72.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621.

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology*, *36*(1), 53-65.

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.

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: insights from brain and behaviour. *Phil. Trans. R. Soc. B*, *369*(1651), 20130296.

Peeters, D., Hagoort, P., & Özyürek, A. (2015). Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives. *Cognition*, *136*, 64-84.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, *36*(04), 393-414.

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

Sitnikova, T., West, W. C., Kuperberg, G. R., & Holcomb, P. J. (2006). The neural organization of semantic memory: Electrophysiological activity suggests feature-based segregation. *Biological Psychology*, *71*(3), 326-340.

Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience, 20*(11), 2037-2057.

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science, 5*(6), 679-692.

Willems, R. M., Özyürek, A., & Hagoort, P. (2008). Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. *Journal of Cognitive Neuroscience, 20*(7), 1235-1249.

Willems, R. M. (Ed.). (2015). *Cognitive neuroscience of natural language use*. Cambridge: Cambridge University Press.

Chapter 5

Behavioral and electrophysiological signatures of indirect request processing in a naturalistic environment

Abstract

Linguistic utterances are frequently underdetermined when it comes to their meaning. For example, in a restaurant one could say “My soup is cold” to a waiter to request him or her to warm up one’s soup. Although in recent years a growing body of research has been dedicated to the question of how we process the meaning of underdetermined utterances like indirect requests, one of the main challenges that remains is how to address this question in naturalistic ecologically valid settings. The present study takes on this challenge. In two experiments, participants were immersed in a virtual restaurant in which they were asked to listen to and respond to utterances from virtual guests in the restaurant. They were presented with possible requests (e.g., “My soup is cold”) and statements (e.g., “My soup is nice”). By looking into a virtual mirror, participants were assigned a role before the start of the experiment, namely to be a waiter or a restaurant critic. Experiment 1 showed that waiters interpreted the possible request as requests and the statements as statements, while restaurant critics interpreted both sentence types as statements. In Experiment 2 we repeated the same experiment while we measured the electroencephalogram (EEG). We predicted to find a difference in the event-related potentials (ERPs) between possible requests and statements for the waiters, but not for the critics. However, we did not find any evidence in line with this prediction. Rather, we observed only an effect of the participant’s role on the ERPs. These findings indicate that we were successful in biasing participants to interpret the possible requests and statements differently as a function of their role. However, in the neurophysiological brain responses, we did not find any reliable evidence for a difference in the processing of requests versus statements. At a methodological level, this study represents an important step forward in studying the behavioral and neural correlates of pragmatic language processing in rich interactive environments.

Behavioral and electrophysiological signatures of indirect request processing in a naturalistic environment

Linguistic utterances are often underdetermined when it comes to the meaning that they convey (Hagoort, 2017). That is, the coded meaning of an utterance can be very different from the speaker meaning (i.e. what the speaker intended to communicate). Furthermore, speakers often intend to achieve something with their use of an utterance in a certain context (Grice, 1957; Levinson, 2017). For example, in a restaurant one could say to the waiter "My soup is cold", to request the waiter to warm up one's soup. It is the communicative intention of the speaker that drives the listener's behavior, rather than his or her actual words (Levinson, 2017). Often the intention of the speaker can be conceptualized in terms of a speech act, such as a request in the example above (e.g., Searle, 1969). Understanding speaker meaning usually involves making a pragmatic inference in which listeners take into account information from other sources than the linguistic code. One can think of contextual factors like discourse and the implicit background knowledge shared between speaker and listener (Hagoort & Levinson, 2014). One of the main goals of experimental pragmatics and neuropragmatics is to account for processes involved in constructing speaker meaning. With the current study we aim to contribute to this by investigating the behavioral and neurophysiological correlates of how people construct meaning when presented with non-conventional indirect requests (henceforth 'indirect requests'). This type of request most often comes in the form of a negative state remark, where the speaker asserts a negative state (e.g., "My soup is cold") that can be resolved by the hearer (e.g., by warming up the soup; Holtgraves, 1994). Negative state remarks are commonly interpreted as requests when the speaker is of a higher status (e.g., a customer to a waiter) and the interaction occurs in a particular context or setting (e.g., a restaurant; Holtgraves, 1994; 2008).

In recent years the question of how we bridge the gap between coded meaning and speaker meaning for indirect requests has received more attention. There is a considerable amount of evidence to suggest that brain areas outside the core language system are involved in understanding speaker meaning in general (Bašnáková, Weber, Petersson, Van Berkum, & Hagoort, 2013; for an overview see Hagoort & Levinson, 2014; Hagoort, 2017). For indirect requests in particular, these networks include the Theory of Mind network (Van

Ackeren, Casasanto, Bekkering, Hagoort, & Rueschemeyer, 2012) and the motor system (Van Ackeren et al., 2012; Van Ackeren, Smaragdi, & Rueschemeyer, 2016). Furthermore, earlier research has shown that patients with right hemisphere damage have difficulty comprehending non-conventional indirect requests (e.g., Stemmer, Giroux, & Joanette,1994; see also Holtgraves, 2012). Unfortunately, less is known about the neurophysiological time course of meaning construction during indirect request processing. Moreover, theoretical accounts differ considerably in their predictions regarding the time course of understanding speaker meaning. Most of the accounts relevant for the current study were formulated in terms of the processing of literal as compared to non-literal language, for example to provide an account of how we process and understand irony (e.g., Filik, Lethold, Wallington, & Page, 2014; Gibbs & Colston, 2012). The most important differences between them are concerned with the necessary steps and the influence of context in accessing the indirect meaning. For example, the indirect access model, derived from the work by Grice (1975) and Searle (1975) postulates that listeners first compute the literal meaning of an utterance and then, if this meaning is inappropriate in the context, infer the speaker's communicative intention. Thus, because the literal meaning is first computed and only thereafter, if necessary, the non-literal meaning, this will entail additional processing costs in the case of non-literal language. On the other hand, the direct access view claims that, in an appropriate context, the listener can completely by-pass the literal meaning of an utterance and compute the indirect meaning immediately (Gibbs, 1994; 2002). This is based on the idea that contextual information and lexical processes can interact very early on. Somewhat in between, the graded salience hypothesis argues that the most salient reading of the sentence will be the initial reading (Giora, 2003). For example, conventional indirect request (e.g., "Can you pass the salt?"), are highly frequent and thus the initial reading will be the request reading (i.e., pass me the salt) rather than a question about the physical ability of the listener to pass the salt (Coulson & Lovett, 2010).

A useful technique to investigate the temporal dynamics of language processing is the Event-related potential (ERP) technique. The ERP is a time-locked measure of electrophysiological brain activity with a high temporal resolution that has been used often to study sentence processing over time (for an overview see Kutas & Federmeier, 2011). This technique has been used to investigate process of meaning construction for other types of non-literal language, like idioms (e.g., Canal, Pesciarelli, Vespignani, Molinaro, & Cacciari, 2015), irony (e.g., Regel, Gunter, & Friederici, 2011; Spotorno, Cheylus, Van Der Henst, & Noveck, 2013), metonymy (e.g., Schumacher, 2014), and metaphor (e.g., Bambini, 2016; Lai, Curran, & Menn, 2009). In the ERP literature on non-literal language comprehension, the predictions regarding the processing steps involved in comprehending indirect utterances are often operationalized in terms of predictions about the influence of certain manipulations on specific ERP components, in particular on the N400 (350 -600 ms) and the Late Positivity Component (LPC) or P600 (~600 - 900 ms). Even though the functional significance of these components is still under debate, they can be a useful tool in disentangling different processes during indirect speech comprehension. Effects in one and/or both (i.e. a biphasic pattern) of these windows have been observed in studies on non-literal language comprehension. Overall, the studies that report only an effect on the N400 component argue against the indirect view, while studies that report a biphasic pattern or only a modulation in the P600 window have been interpreted in favor of indirect access theories (e.g., Coulson & van Petten, 2002; De Grauwe, Swain, Holcomb, Ditman, & Kuperberg, 2010; Yang, Bradley, Huq, Wu, & Krawczyk, 2013; for a recent overview see Bambini, 2016).

Coulson and Lovett (2010) were among the first to use ERPs to investigate the processing of non-conventional indirect requests. In their study, participants read a short story, for example about a couple in a restaurant, followed by a target sentence (e.g., “my soup is too cold to eat”). The story could either bias the interpretation of the target sentence toward a request, for example when the woman uttered the sentence to the waiter, or toward a statement, for example when the woman uttered the target sentence to her husband. Because the target sentences were identical in both contexts, ERPs were analyzed for each word in the sentence. Several differences in the ERPs were observed between indirect requests and the statements. First, the second and third word of the indirect requests as compared to statements elicited larger Late Positivity Complex (LPC). The authors argued that this is likely to reflect the reader drawing on information from the preceding story context. Readers retrieved information from semantic memory to correctly interpret the indirect requests. Second, the ERP to the fifth word (i.e. “cold” in “my soup is too cold to eat”) was more negative for the literal statements than the indirect request in a time-window of 400 to 700 ms after word onset on frontal electrodes. The authors suggested that this effect was due to the concreteness of the adjective (i.e. cold). In the literal statement, the reader may have imagined the coldness of the soup, while for the indirect requests, the word cold served as an important clue to the speaker's intentions (Coulson & Lovett, 2010). Finally, over-sentence averages were analyzed showing a larger positivity for indirect requests as compared to statements, beginning at the second word. Low frequency over-sentence effects have been suggested to index easy of processing (King & Kutas, 1995) and the authors suggested that this result might indicate that processing indirect requests was actually less taxing than statements. The scenario's in their study were indeed specifically designed to promote a target reading of the utterances as indirect requests (Coulson & Lovett, 2010). In sum, the authors argued that their results are more compatible with the direct access view (Gibbs, 1994) than with a more indirect view, considering that there was an early effect of discourse context on the construction of non-literal meaning. However, it was also stated that the study needs replication since there were only a relatively small number of stimuli and many trials were excluded from the analysis due to eye blinks, which rendered the signal-to-noise ratio less than optimal (Coulson & Lovett, 2010).

The study by Coulson and Lovett (2010) provides an important first insight into the electrophysiological processing of indirect requests processing in context, however it has one significant limitation, aside from the limited number of trials included in the analysis. The interpretation of indirect requests, and most other forms of non-literal language, relies heavily on the context, including the situational context, the discourse, and the beliefs and inferences of the participants (Hagoort & Levinson, 2014). Yet, especially when electrophysiology (EEG) is used, in studies on non-literal language processing this context has been reasonably limited. It is hard for researchers to design experiments that offer the contextual richness of everyday situations, while maintaining control over linguistic, non-linguistic and other signals that the EEG is sensitive to. Thus, the challenge is to develop a controlled experimental paradigm that can tap into indirect request processing without impoverishing the situational context in which they most often occur, namely during sequentially structured interactions in everyday environments (Levinson, 2013). With the experiments described in this chapter we took on this challenge. On the basis of the reliable findings obtained in Chapter 4, we created a new paradigm, the virtual restaurant paradigm (described below), to investigate how people construct meaning when encountering indirect requests. In Experiment 1 participants listened to utterances (possible requests vs. statements) and they were asked to briefly respond to what was said. In Experiment 2 we used the same paradigm while we recorded the participants' electroencephalogram (EEG).

The virtual restaurant paradigm

We used virtual reality (VR) to create a rich three-dimensional setting, namely a restaurant, to test indirect requests comprehension (see Figure 1). By offering the possibility to recreate very complex, yet well controlled, everyday environments, VR allows researchers to increase the ecological validity of a study, while maintaining full experimental control. In recent years, VR has been used more often to study language behavior (Casasanto, Jasmin, & Casasanto, 2010; Eichert, Peeters, & Hagoort; 2017; Gijssels, Casasanto, Jasmin, Hagoort, & Casasanto, 2016; Heyselaar, Hagoort, & Segaert, 2015) and it has produced reliable results even in combination with continuous EEG recordings (Peeters & Dijkstra, 2017; Tromp, Peeters, Meyer, & Hagoort, 2017). The restaurant that was used in the experiments described here was also used successfully in Chapter 4 of this thesis to study audio-visual semantic integration in a naturalistic environment.

Figure 1. Screenshot of the Virtual Environment (VE).

In the current experiments, participants were moved from table to table through the virtual restaurant in an automatic fashion. At each table, a restaurant guest was seated who would produce and utterance, either a possible request (e.g., “My soup is cold”) or a statement (e.g., “My soup is nice”). However, under many conditions, “My soup is cold” can also be understood as a statement. In other words, it can easily be used to perform the speech act of informing, rather than requesting. There is not always a one-to-one mapping of speech acts onto utterances (Levinson, 2017). Thus, to encourage or discourage the request interpretation of the possible request we gave participants a specific role during the experiment, namely a waiter or a restaurant critic. This role assignment was done by means of brief instruction at the beginning of the experiment and by having participants look in a mirror in the virtual restaurant. In this mirror they saw a virtual representation of themselves (male or female) who was wearing the appropriate outfit for their role (see Yee & Bailenson, 2007 for other manipulations of self-representation in VR). The participant’s head was tracked by motion tracking software, so the head in their mirror image would move with them. Both groups heard exactly the same utterances, yet as a function of their role we expected differences in how the sentences would be processed. We know that guests often speak to waiters about their food or drinks and that waiters are there to perform or suggest an action that can solve a problem if there is one (i.e. relieve the negative state "cold soup"). In contrast, restaurant critics do not do this. Rather, they listen to opinions about food and drinks and observe environments to form judgements about them later. Thus, they will most likely not interpret "My soup is cold" as a request. The only task of the participants in the experiment was to briefly respond to what the restaurant guest had said while taking into account their own role in the experiment. The virtual restaurant paradigm differs in several aspects from more traditional paradigms; 1) there is no reference to the experimental manipulation in the task; 2) participants are in a rich meaningful environment and 3) participants are direct addressees in a small but somewhat interactive context.

In Chapters 2 and 3 of this thesis, participants were asked to judge whether the sentence that they heard was meant as a request or not. This type of categorization task is a common practice in research on speech acts and indirect requests and it is often necessary to ensure that participants actually understood the correct intention before underlying cognitive processes or neural activity of this process can be described (Gisladottir, Chwilla, and Levinson, 2015; Tromp, Hagoort, & Meyer, 2016; Van Ackeren et al., 2012; Van Ackeren et al., 2016). However, an explicit judgement like this is almost never required during real interaction. Rather, interaction is characterized by the reciprocity of roles (e.g., speaker and addressee), and a speech act produced in an interactive context often sets up an expectation for a response (Levinson, 2006). Thus, in our experiment we asked participants to respond to the speaker, yielding a simple but natural sequence of turns in the absence of any meta-judgement on the speech act performed by the restaurant guest.

The second novel attribute of our paradigm is the virtual environment. In psycholinguistics, language comprehension in frequently studied in isolation, while very often when we listen to someone in real life we make use of many non-linguistic cues from the environment to understand what they are saying (Knoeferle, 2015). In the virtual restaurant paradigm, participants can use both auditory and visual information, since each utterance is related to a restaurant setting and the guests will often refer to the food or drinks on the table in front of them. On a more pragmatic note, a restaurant setting is a meaningful place in which the indirect speech act of requesting is very often performed (Holtgraves, 1994).

A third property of the virtual restaurant paradigm is that participants are active addressees rather than passive listeners and that they are in a simple, but somewhat interactive context. In our paradigm, participants were directly addressed and asked to produce a short sentence in response to the utterance, without any further constraints. Very often, in pursuit of strict experimental control, psycholinguists make the participants in their experiments detached overhearers of single sentences or short fragments of conversation between people they cannot see and to which they have no relation (Hoeks & Brouwer, 2014). This is not always problematic, since certain psycholinguistic processes might be best studied under these conditions (e.g., word recognition, syntactic processing). However, it *is* often problematic in the case of pragmatic phenomena. Studies in pragmatics are concerned with how we understand utterances on the basis of the communication at hand and what is known about the conversational context (Hoeks & Brouwer, 2014). Thus, to study pragmatic phenomena it is often necessary to provide a conversational context and the situational information on which to base their beliefs and inferences. Although it is by no means optimal our paradigm does provide some more context than other paradigms do. The guests in the restaurant are virtual, but visible speakers and the relation between the participant and the guest is evident from the role of the participants. Furthermore, participants have an active role in the experiment, which might make them more engaged and motivated.

The present study

In Experiment 1, we used the virtual restaurant paradigm to investigate how people interpret communicative actions in spoken utterances as a function of the interaction between their role in the experiment (waiter vs. critic) and the type of utterance (possible request vs. statement). The responses of participants were recorded and coded and we measured response duration and gap duration (i.e. the length of the gap between the utterance of the guest and the response of the participant). We predicted that there should be an interaction between the sentence and role for the responses; waiters should more often interpret the possible requests as requests as compared to the statements, while restaurant critics should interpret both sentence types as statements.

In Experiment 2, we used the same paradigm while we recorded the participants' EEG and analyzed event-related potentials (ERPs). Since ERPs are very time-sensitive they can be useful in elucidating the relative contributions over time of the different manipulations in our paradigm. We predicted that if there is a difference in interpretation between the possible requests and statements, as evident from the verbal responses of the participants, this should be visible in the online processing of these utterances as well. Thus, since we predict a behavioral difference between the sentence types for the waiters, but not for the critics, we also predict a difference in the ERPs to the critical words for the waiter, but not for the critics. As discussed above, studies on non-literal language comprehension often take the time window in which ERPs are affected by the manipulation as support for either more direct or indirect processing models (e.g., Bambini, 2016; Spotorno et al., 2013). In short, effects in the N400 window are often taken as support for early more direct access to the non-literal meaning while biphasic effects of effects in the LPC window are considered as evidence for more indirect views on non-literal meaning retrieval. Thus, for the current study, effects in the N400 window, in the absence of later effects, would support a direct access view for indirect request processing. Conversely, a biphasic pattern or effects only in the LPC window would support more indirect processing models. However, there are two issues to consider regarding the ERP window predictions for the current study in light of previous studies. First, effects in neuropragmatics are usually very dependent on the saliency and amount of contextual clues provided to the reader/listener. For example, on the same target sentences, Bambini (2016) observed an N400 effect for metaphors compared to literal sentences only when there were very little contextual clues to guide the interpretation of the metaphors. An effect in the P600 window was observed both in the absence and presence of contextual cues (Bambini, 2016). Second, the paradigm used in the current study varies substantially from previous studies, since participants are embedded in a very rich virtual environment and they are not asked to perform any explicit comprehension task. Previous research has shown that ERP effects, especially the P600 effect, are task-dependent (e.g., Kolk, Chwilla, Van Herten, & Oor, 2003, for a discussion see Brouwer, Fitz, & Hoeks, 2012; Kuperberg, 2007). Thus, for the present study one has to consider the possibility that the absence of a comprehension task, although more ecologically valid, might attenuate the ERP effects predicted as compared to studies with more explicit task demands (see also Brouwer & Crocker, 2017).

Experiment 1

Participants

Thirty-three native speakers of Dutch (5 males) participated in the experiment. All were right-handed, had normal or corrected-to-normal vision and hearing and had no history of speech problems. Average age was 24 years (range: 21 - 28 years). The study was approved by the ethics board of the Social Science Faculty of Radboud University. Participants provided informed consent and were paid to participate in the experiment. One participant was excluded because of a technical failure during the experiment.

Materials

Eighty experimental sentences and 40 filler sentences were constructed. 40 experimental sentences were negative state remarks (e.g., “My soup is cold”), which could be interpreted as a request (i.e. “possible request” or PR). To construct the other 40 experimental sentences (i.e. “statements” or S), the final word of the possible request was replaced by a different word so that it could no longer be interpreted as a request (e.g., “My soup is nice”). Thus, sentences in the PR condition and the S condition were identical apart from the final word. The filler sentences were generic statements (e.g., “I always come here for lunch”). The complete stimulus list can be found in Appendix A. All sentences were related to a restaurant setting and most sentences referred to a food or drink item. The item that was referred to in the sentence (e.g., a bowl of soup) was presented in the virtual environment (VE) on the table in front of the restaurant guests that uttered the sentence. If the utterance did not refer to a food or drink item, a generic cup, bowl or plate was present on the table. All items that were displayed had high name agreement (M = 90.63%). They were taken from the database described in Peeters (2017). The sentences were recorded by ten native speakers of Dutch (5 males and 5 females). These speakers were not aware of the experimental manipulations or of the virtual environment that the experiment would take place in. They were asked to produce the sentences as neutrally as possible. The speech analysis package Praat (version 5.1) was used to equalize the sentences in maximum amplitude. The experimental sentences had an average duration of 1550 ms (not statistically different across conditions).

Materials pre-test

Sentences in the experimental conditions were matched on length (p = .463), frequency (p = .631) and lemma frequency (p = .278). The properties of the stimuli can be found in Table 1. In addition, the sentences were rated on paper in a pre-test by 32 participants. Sixteen participants were instructed to imagine being a waiter in a restaurant and 16 participants were instructed to image to be a restaurant critic. There were two different versions of the rating document to avoid order effects. Participants were asked to assess, on a scale from 1 to 9, whether the sentence was a request (1) or literal statement (9). In addition, they rated whether the sentence was very positive (1) or very negative (9) and whether they felt a lot of emotion with the sentence (1) or no emotion (9). These ratings of request, valence and arousal respectively are presented in the lower part of Table 1.

Table 1. *Stimuli properties*

Measure	Role	Condition			
		Possible Request (PR)		Statement (S)	
		M	*SE*	*M*	*SE*
Length		6.15	0.44	6.55	0.37
Frequencymil		27.17	8.37	32.94	9.57
Frequencylemma		1889.85	561.81	3371.95	1144.29
Request	Waiter	4.85	0.14	7.94	0.07
	Critic	6.74	0.15	8.41	0.05
Valence	Waiter	7.70	0.06	3.29	0.09
	Critic	7.97	0.06	3.19	0.10
Arousal	Waiter	4.61	0.08	5.25	0.11
	Critic	4.59	0.11	5.15	0.12

ANOVAs were conducted to statistically test the difference between the experimental sentences on the three criteria. For the request ratings, possible requests were overall rated lower (i.e., ‘more a request’) than statements as evident by a main effect of Sentence ($F(1,30) = 54.84$, $p < .001$). Importantly, there was also a significant interaction between Role and Sentence ($F(1,30) = 4.87$, $p < .05$). The difference between the PR and S was larger for the waiters ($M = 3.10$) than the critics ($M = 1.67$). For valence there was also a main effect of Sentence; PRs were rated as more negative than statements ($F(1,30) = 778.78$, $p = < .001$). However, there was no interaction with Role ($F(1,30) = 1.21$, $p = .280$). Thus, valence differences between sentence conditions did not differ between the waiters and critics. Finally, there was a main effect of Sentence ($F(1,30) = 16.13$, $p = < .001$) for arousal, but there was no interaction between Sentence and Role ($F(1,30) = .76$, $p = .759$). In sum, only for the request ratings there was a difference between the waiters and critics, while there was no difference between the roles for valence, or for arousal.

Design and Procedure

As described in the Introduction, we created a new method to study indirect request comprehension in virtual reality. Participants were randomly assigned to imagine themselves as either a waiter or a restaurant critic during the experiment. They were encouraged to behave according to their role by means of the instruction before the experiment and during the experiment itself by looking in the virtual mirror before going through the restaurant. There were 10 tables in the restaurant and at each table a virtual restaurant guest was seated who produced a sentence. Participants made 12 rounds through the restaurant. All virtual agents that uttered the sentences in the restaurant (henceforth "restaurant guests") were Caucasian females or males that looked similar in age and background to the ten speakers who recorded the sentences.

Each experimental sentence was produced by a woman (list 1) and a man (list 2) and each item appeared either in the first half (list A) or in the second half (list B). This resulted in four lists. If version one of an item-set (e.g., "My soup is cold") appeared in the first half of the experiment, the other version of the sentence (e.g., "My soup is nice") appeared in the second half. Items were randomized within a block (i.e. a round through the restaurant), but the same condition was not repeated more than three times in a row. Before the onset of the experiment, participants were shown a menu so that they could familiarize themselves with the objects that could be displayed in the restaurant. They then received the instructions, which were the same for the waiter and critic apart from the description of the role. The description for the waiter condition was: "You will see that you are dressed as a waiter/waitress, this is because you work in the restaurant" and for the critic it said: "You will see that you are dressed formally, this is because you are a restaurant critic for a magazine". Participants were instructed to briefly respond to the utterance of the restaurant guests they would encounter while keeping in mind their role in the experiment.

Before each round participants saw the mirror in the restaurant and they were given a few seconds to look into it to reinforce their role. Then, participants were automatically moved from table to table through the restaurant (i.e. they did not walk themselves). The trials sequence was as follows: when the participant arrived at a table the restaurant guest would look up and turn his or her head to directly gaze at the participant. He or she would wait for 1000 ms and then utter the sentence, after which the participant had 4000 ms to respond before the restaurant guest would turn his/her head back and the participant was moved to the next table.

Apparatus

Participants were seated in a CAVE system consisting of three screens (VISCON GmbH, Neukirchen- Vluyn, Germany) which covered the entire horizontal visual field of the participant. In the CAVE a custom-made virtual environment (VE), the restaurant, was projected (see Figure 1). The VE was made using Vizard (version 4.08, WorldViz, Santa Barabara, CA). Participants wore 3D shutter glasses equipped with reflectors. The position of the head of the participants was tracked by means of infrared motion capture cameras that tracked the reflectors on the glasses (Vicon Motion Systems Ltd, UK). This system was used to update the participant's viewpoint when he or she moved. The utterances produced by the restaurant guests were presented through two speakers. The spoken utterances produced by the restaurant guests, the gaps, and the answers given by the participants were recorded for all trials by means of a wireless microphone.

Analysis

Praat (version 5.1, www.praat.org) was used to measure the response durations and gap durations. Gap duration was measured from the offset of the sentence uttered by the restaurant guest to the onset of the response from the participant. Responses from all participants were transcribed and then coded by the first author. They were coded as Request Response (RR) if the response indicated that the participant understood the utterance of the restaurant guest as request or Neutral Response (NR) if the participant understood it as a statement. A typical example of an RR to "My soup is cold" was "I will go warm it up" and a typical NR was "Okay, I am sorry to hear that" For more examples of responses and how they were coded, see Appendix B. Missing data and responses that indicated that participants did not hear/understand the sentence (e.g., "Sorry, I did not understand you") were coded as missing values (waiter; PR: 0.20%, S: 0.14%; critic; PR: 0.14%, S: 0.20%). To ensure reliability of the coding, the responses from 10 participants (5 in the waiter condition and 5 in the critic condition) were coded also by an independent rater. The inter-rater reliability was 97.41% (not statistically different per condition).

For all analyses trials with missing values were removed (0.96% of the data). The response type analysis was conducted with logit mixed model in R (Jaeger, 2008). The outcome measure was Response Type (RR vs. NR). The full model included the fixed effects of Sentence (PR vs. S), Role (waiter versus critic) and the interaction. Predictors were mean-centered. The random structure contained random intercepts and slopes for Sentence by participant and item. Response Duration and Gap Duration were analyzed separately with linear mixed-effects regression models in R (version 3.0.3; The R foundation for statistical computing; lme4 package, Bates, Mächler, Bolker, & Walker, 2014). The predictors (Role, Sentence) were mean-centered and the random structure was the same as for the model for Response Type. To evaluate the effect of the different predictors, models were compared using a likelihood ratio test.

Results

The average percentage of Request Replies per condition can be found in Table 2 and the results of the final model for Response Type can be found in Table 3.

Table 2. *Percentage of Request Replies (RRs) per sentence condition (PR, S) and per role (waiter, critic).*

Role	Condition			
	Possible Request (PR)		Statement (S)	
	M	*SE*	*M*	*SE*
Waiter	84.99	1.42	4.10	0.79
Critic	3.46	0.73	0.16	0.16

Table 3. *Results of the logit mixed model analyses of response type. The estimated coefficient (β), standard error (SE) and z-value (z) and p-value (p) are presented.*

Measure	Fixed Effects	*β*	SE	z	p
Response type	Intercept	-3.68	0.43	-8.53	<.001
	Sentence	5.14	0.87	5.88	<.001
	Role	4.97	0.59	8.37	<.001
	Sentence x Role	3.03	1.22	2.49	<.05

As can be seen in Table 3 there was a main effect of Role. Participants in the waiter condition made significantly more request replies than participants in the critic condition (see Table 2). In addition, there was a main effect of Sentence, reflecting that there were more request replies to possible request than to statements. Importantly, the significant interaction indicated that, as can be seen from Table 2, this effect was significantly larger for the waiters than for the critics.

For the response duration model, including the interaction between Sentence and Role ($\beta = 0.42$, SE = 0.19, $z = 2.12$, χ^2 (11) = 7.106, $p = .008$) significantly improved model fit. Response duration in the waiter condition was longer for the possible requests ($M = 2.40$, $SE = 0.04$) than for the statements ($M = 1.80$, $SE = 0.03$), while this difference was much smaller in the critic condition (PR: $M = 1.79$, $SE = 0.03$; S: $M = 1.61$, $SE = 0.04$).

For the outcome variable gap duration only the inclusion of the predictor Role improved model fit ($\beta = 0.39$, SE = 0.21, z = 1.90, χ^2 (10) = 7.290, $p = .007$), gaps were on average longer for the critics ($M = 1.72$, $SE = 0.03$) than for the waiters ($M = 1.32$, $SE = 0.02$). Including the predictor Sentence did not improve model fit (χ^2 (9) = 1.402, $p = .246$), neither did including the interaction term (χ^2 (11) = 0.600, $p = .439$).

Interim Discussion

Participants were presented with spoken utterances (possible request vs. statement) and they were asked to briefly respond taking into account their role (waiter vs. critic) in the experiment. Importantly, the results revealed that waiters more often interpreted the possible requests as requests as compared to the statements. In contrast, the restaurant critics interpreted both sentence types as statements. Thus, it can be concluded that in the absence of a meta-linguistic task, we successfully biased participants to interpret identical negative state remarks differently depending on their role in the restaurant. Second, waiters gave slightly longer responses when they were presented with a request. This might be due to the fact that often in their answers to these items, they outline a specific course of action to comply with the request, while for the statements they gave shorter, more general answers. Finally, the gap between the end of the critical utterance and the speech onset of the participant was longer for the critics. We return to this point in the General Discussion.

Having established that our experimental manipulation successfully induces participants to respond as either waiters or restaurant critics in our virtual restaurant, we now turn to testing the neurophysiological time course of indirect request comprehension in this naturalistic environment.

Experiment 2

Participants

Eighty-four individuals participated in the experiment (16 males, mean age 22 years, range 18-29 years). One participant was removed from the analysis since he/she revealed after the experiment that he/she was diagnosed with Asperger's syndrome. The remaining participants had no history of neurological problems or diseases, or speech problems. All were right-handed native speakers of Dutch and had normal hearing and normal or corrected-to-normal vision. Ethical approval was granted by the ethics board of the Social Sciences faculty of Radboud University. Participants provided written informed consent and they were paid to take part in the experiment. The data of five participants was unusable due to technical failure during the EEG recording. In addition, the data of 15 participants was removed since for a large percentage of the experimental sentences (more than 35% in one or both conditions) their reply did not indicate that they had interpreted the sentences as would be expected according to their role. Finally, data from 17 participants had to be discarded due to excessive EEG artifacts. The remaining data, of 46 participants (23 waiters; 23 critics), were used for the analyses.

Materials, design and procedure

Based on a by-item response analysis of the stimuli of Experiment 1, three item-sets were removed (to reduce 'incorrect' responses). These item-sets (i.e., 'pizza', 'pannenkoek', 'donut') yielded less than 40 percent correct responses (i.e., the expected response (RR vs. NR) given the condition and role) in one or more of the four conditions. In addition, four filler items were removed that had less than 40 percent correct responses in one or more condition. Thus, the stimulus set of this experiment consisted of 37 possible requests, 37 control sentences and 36 fillers. In Experiment 1, two versions of each sentence were presented, but for Experiment 2 we chose the version with the highest accuracy score based on the responses in Experiment 1. The reason for this was to choose the best version of the item so that chances were higher that it would be interpreted in line with the participants' role and to reduce variability in the prosodic properties of the critical words for the Event-Related Potential (ERP) analysis. Thus, critics and waiters heard exactly the same sentences. The onset, offset and length of the critical words (e.g., 'cold' in "My soup is cold") were measured using Praat.

There were 110 items in total. Participants made 11 rounds through the restaurant. There were two lists so that in list one the first item (e.g., "My soup is cold") appeared in the first half and its counterpart (e.g., "My soup is nice") in the second half of the Experiment, while in list two this was the other way around. Items were randomized within each block, but one condition never occurred more than three times in a row. The trial sequence differed slightly from Experiment 1, as we wanted to reduce movement artifacts in EEG signal caused by articulation. After the participant arrived at a table, the restaurant guest would turn his/her head to the participant, then wait 1000 ms and produce the sentence. Then, the restaurant guest would move his/her head back (and look forward, not at the participant) for 2000 ms after which he/she turned her/his head back to the participant for 4000 ms so that the participant could respond to the restaurant guest. Thus, participants were asked to withhold their respond for 2000 ms rather than to answer directly. Before the experiment participants were instructed to keep eye contact with the restaurant guest while he or she spoke and to not blink during this period. Otherwise, the design and procedure were identical to Experiment 1.

Behavioral analyses

The responses for Experiment 2 were measured and coded in the same way as for Experiment 1 except that in Experiment 1 the audio (of restaurant guest and participant) was continuously recorded, while for this Experiment, the audio was recorded and saved per trial by an in-house Presentation script (Neurobehavioral Sytems). The responses were transcribed during the experiment and rated by the main author. Additionally, the responses of 10 randomly chosen participants were rated by an independent rater; inter-rater reliability was 97.18% (no difference between conditions). Trials with missing values were removed (0.76% of the data).

EEG recording and analysis

The electroencephalogram was initially continuously recorded from 59 active electrodes (Brain Products, Munich, Germany). The electrode montage was the same as in Chapter 4 (Figure 2). However, for many participants four electrodes from the original set (2 on each side of the head above the ears) had to be removed, since it was impossible to obtain a reliable signal due to the presence of the shutter glasses. These electrodes were excluded from the analysis. Furthermore, two electrodes were attached on the canthus to the side of each eye to monitor for horizontal eye-movements and one electrode was attached under the left eye to monitor for eye blinks. Two electrodes were placed over the left and right mastoids respectively and the electrodes were referenced online to the electrode on the left mastoid and offline to the average of the electrodes on the left and right mastoid. Electrode impedance was kept below 20 KΩ. The EEG was recorded with a high cutoff filter of 200 Hz and a low cutoff filter of 0.01 Hz at a sampling rate of 500 Hz. A low-pass filter of 40 Hz was applied offline. Epochs from 100 ms before the onset of the critical word to 1000 ms after critical word onset were selected. ERPs were time-locked to the onset of the critical word and the 100 ms before the onset of the critical word was used as a baseline. Trials with missing values for the behavioral responses were removed. Subsequently, independent component analysis (ICA) was used to correct for ocular artifacts. After the ICA, trials that still contained artifacts were removed (24.03% of the data including the trials with missing values for responses), ERPs were calculated by averaging across trials for each participant and condition separately.

The ERPs were analyzed by means of a nonparametric cluster-based permutation tests (Maris & Oostenveld, 2007). To determine which time points and electrodes show a significant effect, this test used a clustering algorithm based on the physiologically plausible assumption that ERP effects are clustered in time and space over adjacent electrodes and samples. It has the advantage of controlling for the family-wise error rate that arises when an effect is evaluated at multiple time point and electrodes (Maris & Oosterveld, 2007). In short it worked as follows; for every data point a dependent samples *t*-test quantified the difference between two conditions. All neighboring data points with a significance level of $p < .05$ were then grouped into clusters. For each cluster, the sum of the *t*-statistic was used for the cluster-level *t*-statistic. Then, a null distribution was calculated that assumed no difference between conditions (2000 randomizations, calculating the largest cluster-level statistic) after which the observed cluster-level statistic was compared against the null-distribution. Clusters falling in the highest or lowest 2.5th percentile were considered significant (Bonferroni corrected, a p-value < .025 reflects a significant effect). To test for interactions between factors, the difference between the two levels of each factor was used as input to the initial *t*-test. Since Role (waiter, critic) is a between-subjects manipulation, we used an independent samples *t*-test to obtain the *t*-statistics. For the first analysis, we submitted the complete epoch from 0 to 1000 ms after critical word onset to cluster-based permutation testing. Additionally, we tested for the interaction in the N400 window (350 to 600 ms) and, based on Coulson & Lovett (2010), a late (LPC) window (700 – 900 ms). Since we hypothesized to find an effect of Sentence (PR vs. S) for the waiters but not for the critics, we tested for the effect of Sentence separately for the waiters and the critics. Finally, we compared the waveforms of the critics and the waiters collapsed over sentence type (i.e. the main effect of role) and we compared the sentence types collapsed over role (i.e. the main effect of Sentence type).

Results

Behavioral results

Table 4 shows the percentage of request replies per condition and Table 5 the results of the logit mixed-effects model analysis of Response Type. As in Experiment 1, the predictors Sentence and Role significantly improved model fit. There were more RRs in the possible request condition than the statement condition and the waiters gave more RRs than the critics. Importantly, there was a significant interaction between Role and Sentence;

waiters gave significantly more RRs in the possible request condition than the statement condition, while this difference was not observed for the critics.

Table 4. *Experiment 2: Percentage of Request Replies (RRs) per sentence condition (PR, S) and per role (waiter, critic).*

Role	Condition			
	Possible Request (PR)		Statement (S)	
	M	*SE*	*M*	*SE*
Waiter	81.63	1.34	4.17	0.70
Critic	3.18	0.60	0.24	0.17

Table 5. *Experiment 2: Results of the logit mixed model analyses of response type. The estimated coefficient (β), standard error (SE) and z-value (z) and p-value (p) are presented.*

Measure	Fixed Effects	*β*	SE	z	p
Response type	Intercept	-3.25	0.28	-11.93	<.001
	Sentence	3.93	0.55	7.08	<.001
	Role	4.73	0.46	10.20	<.001
	Sentence x Role	3.49	0.95	3.67	<.001

ERP results: interaction

The waveforms are presented Figure 3. The cluster-based permutation test over the full epoch (0 - 1000 ms) did not reveal any reliable clusters when the interaction between Sentence and Role was tested (all p-values > 0.311). Also, the analysis in the N400 window (350 – 600 ms) did not yield any significant clusters (all p-values > 0.600), nor did the analysis of the LPC window (all p-values > 0.325).

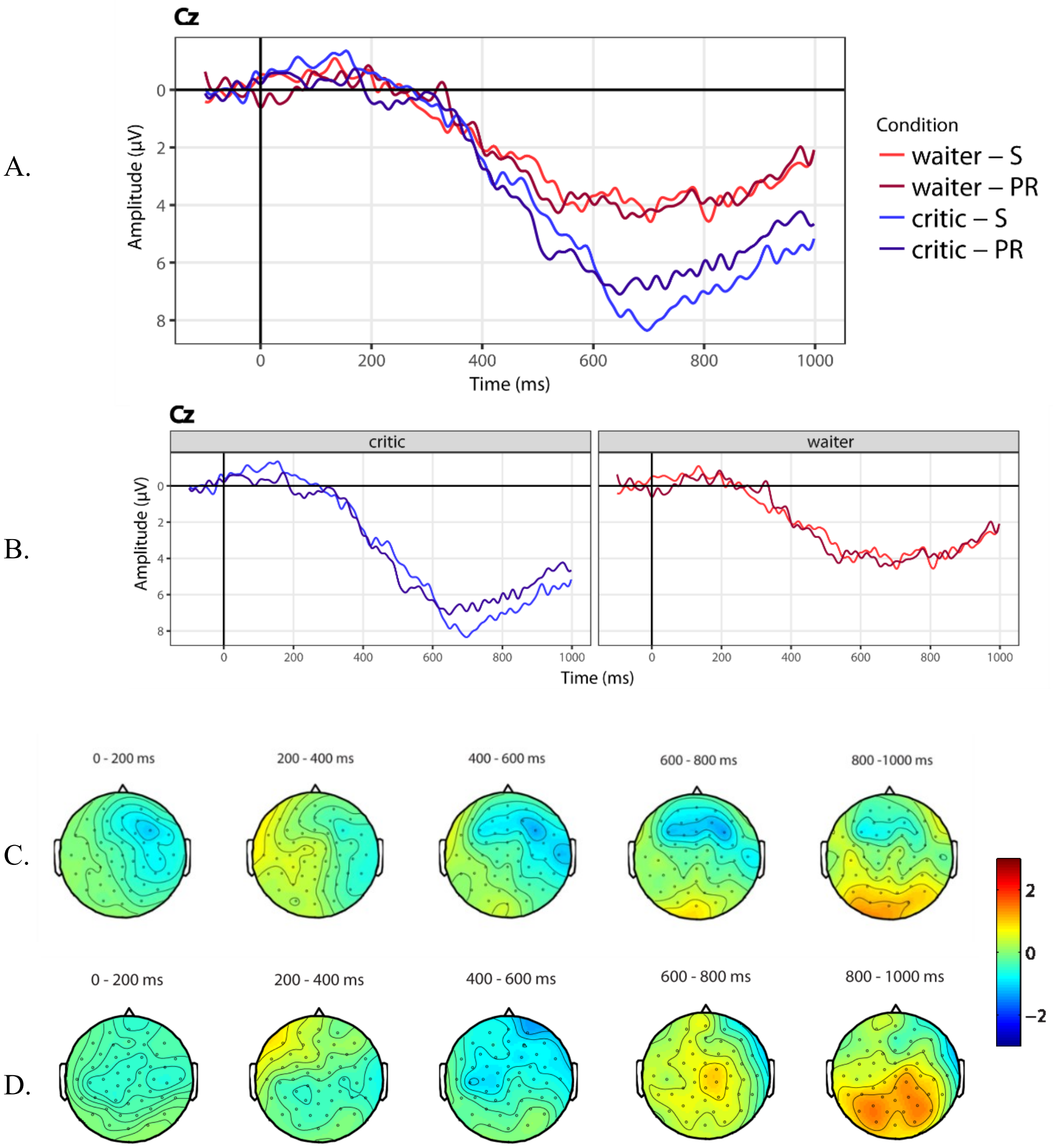

Figure 3. Panel A: event-related potentials for all four condition combinations for electrode Cz, time-locked to the critical word. Panel B: ERPs time-locked to the critical word for critics and waiters separately for electrode Cz. Panel C: Topoplots (statement – PR) for the waiter condition. Panel D: Topoplots (statement – PR) for the critic condition.

ERP results: waiter and critic separately

For the waiters, no reliable clusters were detected in the comparison between possible requests and statements in the entire 0 to 1000 ms epoch (all *p*-values > 0.101), the N400 window (all *p*-values > 0.063) or the LPC window (all *p*-values > 0.199).[3] Also for the critics this comparison did not reveal any significant cluster in the 0 to 1000 ms epoch (all *p*-values > 0.055). The same holds true for the N400 (all *p*-values > 0.396) and the LPC window (all *p*-values > 0.068). Thus, for either group (waiters or critics), the cluster-based permutation test did not provide any evidence for a difference between sentence types (PR versus S).

ERP results: role

Based on visual inspection of the waveforms in Figure 3, the ERPs for the waiters were compared to those of the critics (collapsed over sentence type; see Figure 4). The cluster-based permutation test on the entire epoch (0 – 1000 ms) revealed a significant difference between the waiter and critic condition; a reliable cluster was detected from 360 until 998 ms (p = 0.014). The effect of role can be characterized as a widespread sustained positivity starting from around 0.36 s for the critics as compared to the waiters (see Figure 4).

[3] Trials in which the waiters did not provide the expected reply (i.e. a neutral reply to a possible request) were included in the analyses reported above. Thus, even though participants had on average more than 65% correct responses in both conditions, it is a possibility that these items had an effect on the average ERPs. To exclude this possibility, we ran all analyses again after removing these trials from the data (on average 3 trials per participant, 10.34% of the data). It should be noted that this resulted in a more unequal distribution of trials over the two conditions, hence the results should be interpreted with caution. However, the analysis revealed that also when excluding these trials, there were no significant clusters (all *p*-values > 0.109).

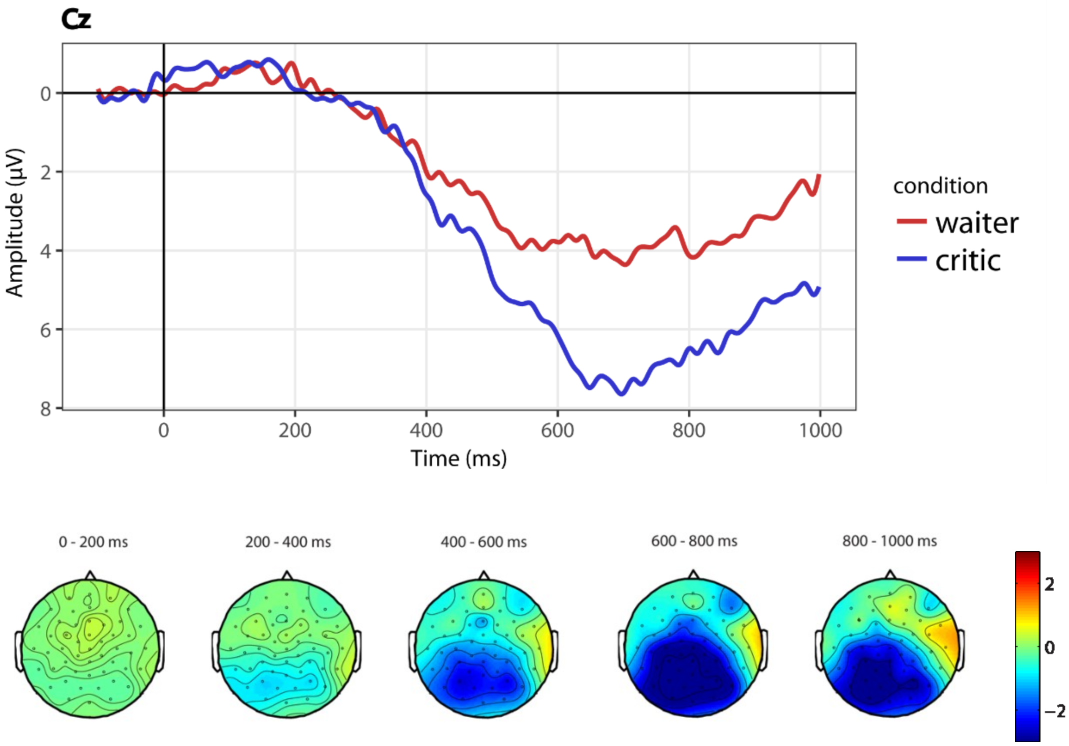

Figure 4. Event-related potentials for the two roles (collapsed of sentence condition) for electrode Cz, time-locked to the critical word. The topoplots represent the difference between the two roles (waiter - critic).

Exploratory analysis: gap window

In the analyses reported above, a window of 1000 ms after the onset of the critical word was tested. However, it is conceivable that differences in the ERPs for the PR versus S condition only manifested themselves after this time window (i.e. in the gap between stimulus offset and the onset of the participant's reply). To explore this possibility an additional time window of 1500 ms after critical word offset was analyzed. This analysis did not yield any significant results (more details can be found in Appendix C).

Interim Discussion

The behavioral results of Experiment 1 were replicated; waiters indicated that they understood the possible requests as requests and the statements as statements, while the critics understood both types of sentences as statements. For the ERPs, we predicted to find an interaction between a participant's role in the experiment and the type of sentence they heard, specifically a difference between possible requests and statements for the waiters, but not for the critics. However, the interaction analyses reported above did not reveal any evidence to support this prediction. In addition, the separate analysis for the waiters also did not reveal a difference between the two sentence conditions. However, there was a significant difference between the critics and the waiters collapsed over sentence condition, namely a widespread sustained positivity starting from around 0.36 s for the critics as compared to the waiters.

General discussion

In two experiments, we used the 'virtual restaurant paradigm' to investigate how people interpret possible request (e.g., "my soup is cold") and statements (e.g., "my soup is nice") as a function of their role in the experiment (waiter vs. critic). In both experiments the responses of the waiters indicated that they understood the possible requests as requests and the statements as statements. In contrast, the critics understood both types of sentences as statements. These findings indicate that we were successful in biasing participants to a certain interpretation of the sentences spoken by the virtual restaurant guests in a visually rich virtual environment and in the in the absence of a meta-linguistic task. In light of the recent debates in cognitive neuroscience, these results provide a first step toward more naturalistic experiments (e.g., Willems, 2015). Participants listened to and saw a 'real' speaker (Hoeks & Brouwer, 2014), they had an active role in the experiment, and a social relation was established between speaker and hearer (Holtgraves, 1994).

In Experiment 2 ERPs were analyzed and we predicted to find a difference for the waiters, as a function of how they understood the sentence (i.e., request or statement) but not for the critics. However, we failed to obtain any evidence in line with this prediction. Rather, only the role that participants were assigned to affected the ERPs. Before we discuss the implications of these results, three potential concerns are addressed.

First, although the number of participants and trials included in the analysis is comparable to other ERP studies on language processing in VR (Peeters & Dijkstra, 2017; Tromp et al., 2017) it might be the case that there was not enough power in the current study to observe reliable differences in the ERPs. This concern pertains not only to the amount of trials included in the analysis, but also to the design itself. Our predictions were based on studies in which a meta-linguistic judgement task was present which could have directed the attention of the participant to the experimental manipulation. For example, in the study by Coulson & Lovett (2010) participants were presented with a probe to assess whether they interpreted the sentence correctly. As mentioned in the Introduction, ERP effects, especially the P600 effect, can vary significantly as a function of the task demands of the experiment. For example, in a study by Kolk and colleagues (2003) a difference in the size of the P600 effect was observed depending on whether an explicit acceptability judgement was required from participants or not. Although this study was not concerned with pragmatic processing or indirectness, the results do attest to the fact that removal of an explicit task might have an influence on the P600 effect. Thus, when comparing our paradigm to that of previous ERP studies on non-literal language comprehension, it is a possibility that the absence of a task in our study attenuated any potential ERP effects, at least in the P600/LPC window.

The second issue to consider before interpreting the results of this study is that it is conceivable that there were differences in the ERPs, but that they manifested earlier in the sentence rather than at the final word or in the gap after that (i.e. the two windows that were tested). The sentences in the experiment were recorded in full to keep them as natural as possible, however this could have introduced prosodic patterns that could bias interpretation to requests or statements earlier than the critical word. Previous research has shown that spoken single (non-) words can reliably be recognized by participants as having different intentions (or speech acts) depending on their prosodic features (Hellbernd & Sammler, 2016). Although this is a possibility, it is unlikely for the following reasons; the sentences were recorded as neutral as possible and the speakers were not aware of the experimental manipulation when they recorded the sentences. Afterwards they were asked what they thought the sentences might be for and they did not allude to the manipulation that we intended. Thus, if anything, the possible requests were produced in a more neutral way than they would be in real life. Finally, the final word was the unique point at which the meaning of the sentence became clear and thus most likely the point at which the participant could have understood that the sentence was a request or not.

Third, as evident from the results in Chapter 3 of this thesis, individual differences in pragmatic abilities can play a role in the processing of indirect requests. One could argue that participants might have used different strategies to understand the speaker meaning in the possible requests, which might have influenced the group averaged ERPs. However, visual inspection of the individual participant's waveforms does not support this explanation.

While taking into account these concerns, one could speculate that in our study there was little or no difference in the neurophysiological signatures for the processing of requests versus statements. Although the absence of evidence for an effect is not evidence of the absence of an effect, we did create a very contextualized environment in which the requests were very natural and maybe even very predictable. Holtgraves (1994) presented participants with short introductory scenes followed by a conventional request (e.g., "Would you fill the water jug?") or negative state remark (e.g., "The water jug is almost empty"). The results revealed that when speaker and hearer in the scene were of equal status, processing the negative state remark was more time-consuming and involved activation of the literal meaning of the remark. However, if the interpersonal context indicated that the speaker was higher in status than the listener (like the guest-waiter relation in our experiment) this difference disappeared. It was concluded that processing of indirect requests is sensitive to the social context (Holtgraves, 1994). Regarding the mechanism by which this happens, the author stated that "it appears that people are simply primed to recognize the utterances of a higher status speaker as performing directives" (Holtgraves, 2008, p. 628). This idea is also alluded to in Coulson & Lovett (2010) when the authors argued that the over sentence positivity for indirect request as compared to statements, might indicate that processing indirect requests was actually less taxing than statements, because the scenario's in their study were indeed specifically designed to promote a target reading of the utterances as indirect requests (Coulson & Lovett, 2010). For our study, one could argue that the expected utterance that a waiter might hear in a restaurant is some form of a request, since he or she is there to help make the stay for the guest as pleasant as possible. Specifically, there might be little difference in expectation for the final word, since both a request and a statement are equally likely under the contextual restrictions of our paradigm. This might have reduced or eliminated the effects that were observed for non-literal language processing in previous studies with less supporting context (e.g., Bambini, 2016; Spotorno et al., 2013). However, again, one should be careful in making claims based on the absence of an effect in our study.

In addition, our study, unlike the Holtgraves (1994) study, was not designed to test the effect of interpersonal context on indirect request comprehension directly, rather we aimed to use the context for the purpose of inducing the indirect request reading of the possible requests.

We did observe one difference in the ERPs, in the form of a sustained positivity starting from around 400 ms for the critics as compared to the waiters. Again, we should be careful with the interpretation, since the effect started earlier in our study as compared to other studies. However, late positivities have been observed for more context-dependent phenomena, such as and the processing of ambiguous idioms (Canal et al., 2015), irony (e.g., Spotorno et al., 2013), and jokes (Coulson & Kutas, 2001). Coulson & Lovett (2010) also observed a positivity for indirect requests as compared to statements. Though, rather than for the different sentence types, we only observed this difference when we compared the roles directly. It could be hypothesized that this positivity reflects a feature of our design that we did not anticipate. We created the critic condition as a baseline against which to compare the differences we predicted to observe for the different sentence types in the waiter condition. If there would be a difference for the waiters in the absence of a difference for the critics, this would serve as good evidence that the difference we would observe for the waiter were really due to the request versus statement processing rather than other properties of the stimuli (i.e., valence or arousal). However, the overall positivity for the critics might indicate that when participants were in the restaurant critic role, sentences were more difficult to integrate with the overall context of the experiment. Coulson and Lovett (2010) interpret the LPC for indirect requests in their study as reflecting the reader drawing on information from the preceding story context, they retrieved information from semantic memory to correctly interpret the indirect requests. Although we do not want to make any claims about the functional nature of the effect, it is possible that in our study the critics in general had more difficulty with integrating the sentences they heard in the experiment with the context. One piece of evidence that could support this interpretation is the larger gap durations for the critics as compared to waiters observed in Experiment 1. Critics took longer to form and plan their response, even though their responses were not significantly longer than those of the waiters. In sum, participants in the critic role might have had more difficulty with interpreting the sentences in the experiment, which is reflected by a sustained positivity for the critics as compared to the waiters. In conclusion, we did not observe a reliable difference between the requests and statements in our study, thus we cannot make claims about the implications of these results for theories in neuropragmatics. One could only speculate that if there was really no difference between processing indirect requests and statements, this would be evidence in favor of more direct access models (e.g., Gibbs, 2002) and it would stress the importance of (interpersonal) context during indirect request comprehension (Holtgraves, 1994; 2008). However, since the results are not conclusive in this respect, more research is necessary to investigate the processing of indirect requests in a naturalistic environment. A fruitful approach that could address some of the problems of the current design might be to take a step back in ecological validity to make the design more comparable to other designs in the field. One could create a paradigm in virtual reality in which the participant sees and hears the possible requests and statements being uttered to a waiter or, for example, to someone at the table. This would make the design more similar to the one of Holtgraves (1994) and Coulson & Lovett (2010), while still increasing ecological validity. However, this would require a meta-linguistic task to assess whether the participant interpreted the sentences correctly. For example, in Holtgraves (1994) participants were asked to perform a sentence verification judgement that were either direct requests (i.e., warm up the soup) or unrelated to the setting.

To conclude, we created a new paradigm in virtual reality to investigate indirect request processing. On a behavioral level, we succeeded in biasing participants to a certain interpretation of the sentences spoken by the virtual restaurant guests in a visually rich virtual environment without a meta-linguistic task. Regarding the neurophysiological signatures of indirect request comprehension our results remain inconclusive and more research is necessary. We hope this study still inspires future research in that it represents an important attempt toward studying engaged participants in more realistic dynamic and interactive environments (Hari, Henriksson, Malinen, & Parkkonen, 2015; Willems, 2015; Schilbach et al., 2013). For study of pragmatic phenomena in particular, it provides a way to introduce multiple sources of information that we use in everyday life into the experimental setting (Knoeferle, 2015), including a virtual, yet visible speaker with an established social relation to the listener (Hoeks & Brouwer, 2014).

References

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016). Disentangling metaphor from context: an ERP study. *Frontiers in Psychology*, *7*.

Bašnáková, J., Weber, K., Petersson, K. M., Van Berkum, J., Hagoort, P. (2013). Beyond the language given: The neural correlates of inferring speaker meaning. *Cerebral Cortex, 24*(10), 2572–2578.

Bates, D. Mächler, M., Bolker, B., Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7.

Brouwer, H., & Crocker, M. W. (2017). On the Proper Treatment of the N400 and P600 in Language Comprehension. *Frontiers in Psychology*, *8*.

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.

Canal, P., Pesciarelli, F., Vespignani, F., Molinaro, N., & Cacciari, C. (2015). Electrophysiological correlates idioms comprehension: semantic composition does not follow lexical retrieval. In V. Pirrelli, C. Marzi, & M. Ferro (Eds.), *Proceedings of the NetWordS Final Conference on Word Knowledge and Word Usage: Representations and Processes in the Mental Lexicon,* Pisa, Italy, March 30–April 1, pp. 98–101.

Casasanto, L., Jasmin, K., & Casasanto, D. (2010). Virtually accommodating: Speech rate accommodation to a virtual interlocutor. In S. Ohlsson & R. Catrambone (Eds.) Cognition in flux: Proceedings of the *32nd Annual Meeting of the Cognitive Science Society* (pp. 127-132). Austin, TX: Cognitive Science Society.

Coulson, S., & Kutas, M. (2001). Getting it: human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters*, *316*(2), 71-74.

Coulson, S., & Lovett, C. (2010). Comprehension of non-conventional indirect requests: An event-related brain potential study. *Italian Journal of Linguistics*, *22*(1), 107-124.

Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition*, *30*(6), 958-968.

De Grauwe, S., Swain, A., Holcomb, P. J., Ditman, T., & Kuperberg, G. R. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia*, *48*(7), 1965-1984.

Eichert, N., Peeters, D., & Hagoort, P. (2017). Language-driven anticipatory eye movements in virtual reality. *Behavior Research Methods*, 1-14.

Filik, R., Leuthold, H., Wallington, K., & Page, J. (2014). Testing theories of irony processing using eye-tracking and ERPs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 811.

Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge: Cambridge University Press.

Gibbs, R. W., & Colston, H. L. (2012). *Interpreting figurative meaning*. Cambridge: Cambridge University Press.

Giora, R. (2003). *On our mind: Salience, context, and figurative language*. Oxford: Oxford University Press.

Gijssels, T., Casasanto, L. S., Jasmin, K., Hagoort, P., & Casasanto, D. (2016). Speech accommodation without priming: The case of pitch. *Discourse Processes*, *53*(4), 233-251.

Gisladottir, R. S., Chwilla, D. J., & Levinson, S. C. (2015). Conversation electrified: ERP correlates of speech act recognition in underspecified utterances. *PloS One*, *10*(3), e0120068.

Grice, H. P. (1957). Meaning. *The Philosophical Review*, *66*(3), 377-388.

Hagoort, P., & Levinson, S. C. (2014). Neuropragmatics. In M.S. Gazzaniga, & G.R. Mangun (Eds.), *The Cognitive Neurosciences* (pp. 667–674). Cambridge, MA: MIT Press.

Hagoort, P. (2017). The core and beyond in the language-ready brain. *Neuroscience & Biobehavioral Reviews*, 81, 194-204.

Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of social interaction in human brain function. *Neuron*, *88*(1), 181-193.

Hellbernd, N., & Sammler, D. (2016). Prosody conveys speaker's intentions: Acoustic cues for speech act perception. *Journal of Memory and Language*, *88*, 70–86.

Heyselaar, E., Hagoort, P., & Segaert, K. (2015). In dialogue with an avatar, language behavior is identical to dialogue with a human partner. *Behavior Research Methods*, 1-15.

Hoeks, J. C., & Brouwer, H. (2014). Electrophysiological research on conversation and discourse. Holtgraves, T. (Ed.), *Oxford Handbook of Language and Social Psychology*, pp 365-386. Oxford: Oxford University press.

Holtgraves, T. (1994). Communication in context: Effects of speaker status on the comprehension of indirect requests. *Journal of Experimental Psychology: Learning Memory and Cognition*, *20*(5), 1215–1208.

Holtgraves, T. (2008). Automatic intention recognition in conversation processing. *Journal of Memory and Language*, *58*(3), 627-645.

Holtgraves, T. (2012). The role of the right hemisphere in speech act comprehension. *Brain and Language*, *121*(1), 58-64.

King, J. W., & Kutas, M. (1995). Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, *7*(3), 376-395.

Knoeferle, P. (2015). Language comprehension in rich non-linguistic contexts: combining eye tracking and event-related brain potentials. In R.M. Willems (Ed), *Cognitive Neuroscience of Natural Language Use* (pp. 77-100). Cambridge: Cambridge University Press.

Kolk, H. H., Chwilla, D. J., Van Herten, M., & Oor, P. J. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, *85*(1), 1-36.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23-49.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621-647.

Lai, V. T., Curran, T., & Menn, L. (2009). Comprehending conventional and novel metaphors: An ERP study. *Brain Research*, *1284*, 145-155.

Levinson, S. C. (2006). On the human" interaction engine". In N. J. Enfield, & S. C. Levinson (Eds.), Roots of human sociality: Culture, cognitition and interaction (pp. 39-69). Oxford: Berg.

Levinson, S. C. (2013). Action formation and ascription. In T. Stivers, & J. Sidnell (Eds.), *The handbook of conversation analysis* (pp. 103-130). Malden, MA: Wiley-Blackwell.

Levinson, S. C. (2017). Speech acts. In Y. Huang (Ed.), Oxford handbook of pragmatics (pp. 199–216). Oxford: Oxford University Press.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177-190.

Peeters, D. (2017). A standardized set of 3-D objects for virtual reality research and applications. *Behavior Research Methods*, 1-8.

Peeters, D., & Dijkstra, T. (2017). Sustained inhibition of the native language in bilingual language production: A virtual reality approach. *Bilingualism: Language and Cognition*, 1-27.

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.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, *36*(04), 393-414.

Schumacher, P. B. (2014). Content and context in incremental processing: “the ham sandwich” revisited. *Philosophical Studies*, *168*(1), 151-165.

Searle, J. R. (1975). Indirect speech acts. *Syntax and Semantics*, *3*, 59–82.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language* (Vol. 626). Cambridge: Cambridge University Press.

Spotorno, N., Cheylus, A., Van Der Henst, J. B., & Noveck, I. A. (2013). What’s behind a P600? Integration operations during irony processing. *Plos One, 8*(6), e66839.

Stemmer, B., Giroux, F., & Joanette, Y. (1994). Production and evaluation of requests by right hemisphere brain-damaged individuals. *Brain and Language, 47*(1), 1-31.

Tromp, J., Peeters, D., Meyer, A. S., & Hagoort, P. (2017). The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behavior Research Methods,* 1-8.

Van Ackeren, M. J., Casasanto, D., Bekkering, H., Hagoort, P., & Rueschemeyer, S. A. (2012). Pragmatics in action: Indirect requests engage theory of mind areas and the cortical motor network. *Journal of Cognitive Neuroscience, 24*(11), 2237–2247.

Van Ackeren, M. J., Smaragdi, A., & Rueschemeyer, S. A. (2016). Neuronal interactions between mentalising and action systems during indirect request processing. *Social Cognitive and Affective Neuroscience, 11*(9), 1402–1410.

Willems, R. M. (Ed.). (2015). *Cognitive neuroscience of natural language use.* Cambridge: Cambridge University Press.

Yang, F. P. G., Bradley, K., Huq, M., Wu, D. L., & Krawczyk, D. C. (2013). Contextual effects on conceptual blending in metaphors: An event-related potential study. *Journal of Neurolinguistics, 26*(2), 312-326.

Yee, N., & Bailenson, J. (2007). The Proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, *33*(3), 271-290.

Appendix A: Sentence stimuli

Table A1: Experimental items

Note: The same context sentence was used to make the possible request and the statement, only the critical word was changed. Critical words are underlined.

Possible indirect request (PR)	Statement (S)
Deze appel is beurs	Deze appel is groen
Mijn banaan is bedorven	Mijn banaan is vers
Mijn biefstuk is taai	Mijn biefstuk is mals
Dit bord is vet	Dit bord is keurig
Dit brood is zwart	Dit brood is bruin
Deze cocktail is smerig	Deze cocktail is passend
Mijn cola is warm	Mijn cola is koel
Mijn croissant is uitgedroogd	Mijn croissant is frans
Mijn donut is kleverig	Mijn donut is smeuig
Dit ei is rot	Dit ei is gebakken
Mijn friet is klef	Mijn friet is knapperig
Deze frikandel is vreselijk	Deze frikandel is top
Dit glas is bevlekt	Dit glas is groot
Deze ham is slecht	Deze ham is gekruid
Mijn hamburger is flauw	Mijn hamburger is heerlijk
Dit ijs is gesmolten	Dit ijs is gevarieerd
Dit ijsje is smakeloos	Dit ijsje is fris
Deze kiwi is overrijp	Deze kiwi is smaakvol
Deze koek is goor	Deze koek is zoet
Deze koffie is slap	Deze koffie is smakelijk
Deze lepel is vuil	Deze lepel is schoon
Het licht is fel	Het licht is sfeervol
Deze maaltijd is zout	Deze maaltijd is bijzonder
Dit mes is bot	Dit mes is scherp
Deze muffin is droog	Deze muffin is nieuw
De muziek is luid	De muziek is apart
Mijn pannekoek is smoezelig	Mijn pannekoek is geslaagd
Mijn paprika is ongewassen	Mijn paprika is rood
Deze peper is aangevreten	Deze peper is sappig
Deze pizza is plakkerig	Deze pizza is krokant

Deze pudding is mismaakt	Deze pudding is origineel
Deze salade is vies	Deze salade is aangenaam
Deze sinasappel is verschrompeld	Deze sinasappel is oranje
Mijn soep is koud	Mijn soep is lekker
Mijn spaghetti is hard	Mijn spaghetti is perfect
Mijn sushi is muf	Mijn sushi is verfijnd
Mijn taart is grauw	Mijn taart is authentiek
Deze tompouce is onsmakelijk	Deze tompouce is roze
Mijn wafel is zompig	Mijn wafel is voortreffelijk
Mijn wijn is zuur	Mijn wijn is geurig

Table A2: Filler items

Filler items
Ik heb deze yoghurt net gekregen
Ik heb deze cake net gekregen
Ik heb deze kaas net gekregen
Ik heb deze rijst net gekregen
Ik heb deze worst net gekregen
Ik heb deze bonbons net gekregen
Ik heb deze aardbeien net besteld
Ik heb deze asperges net besteld
Ik heb deze spiesjes net besteld
Ik heb deze spinazie net besteld
Ik heb deze sandwich net besteld
Ik heb deze tomaten net besteld
Ik heb deze smoothie net besteld
Ik heb deze poffertjes net besteld
Ik heb deze olijven net besteld
Ik heb deze omelet net besteld
Ik heb deze champignons net besteld
Ik heb deze kroket net besteld
Ik heb deze kersen net besteld
Ik heb deze druiven net besteld
Ik heb deze bonen net besteld
Ik kom hier altijd voor lunch
Ik kom hier altijd voor ontbijt
Ik kom hier altijd voor diner
Ik kom hier altijd voor brunch

Ik heb net eten besteld
Ik heb net drinken besteld
Ik heb net een voorgerecht besteld
Ik heb net een dessert besteld
Dit restaurant heeft een grote kaart
Dit restaurant heeft een uitgebreide kaart
Ik heb deze kip net gekregen
Ik heb deze aardappeltjes net gekregen
Ik heb deze wortels net gekregen
Ik heb deze garnalen net gekregen
Ik heb deze mandarijn net gekregen
Ik heb deze peer net gekregen
Ik heb deze wrap net gekregen
Ik heb deze chocomel net gekregen
Ik heb deze melk net gekregen

Appendix B: Response coding

Note: For both experiments, all responses from 10 participants (5 in the waiter condition and 5 in the critic condition) were coded also by an independent rater. While coding the responses the raters read the experimental sentence (i.e. "My soup is cold") and the participant's reply. If the reply indicated that the participant wanted to perform an action to relieve the negative state intrinsic to the possible request (i.e. do something about the cold soup) it was coded as a request reply. If not, it was coded as a neutral reply. Replies in which the participant indicated that he or she did not hear or understand the utterance were coded as missing values. Examples of request replies and neutral replies taken from the actual data can be found in table B1 below.

Table B1: Examples from the data of request replies and neutral replies in response to two possible requests.

Possible request	Reply type	Example
Mijn soep is koud *My soup is cold*	Request Reply (RR)	Dat is rot ik zal even een nieuwe voor u halen *That is a pity I will get a new one for you*
		Ik zal u een nieuwe warme soep komen brengen *I will come bring you a new warm soup* Ik zal een nieuwe voor u gaan halen *I will get a new one for you* Ik zal hem even voor u opwarmen *I will warm it up for you*
		Excuses ik ga een nieuwe soep voor u halen *My apologies, I will get you a new soup*
	Neutral Reply (NR)	Wat vervelend *How annoing*
		Dat klinkt niet goed *That does not sound good* Dat is niet de bedoeling *That is not how it's supposed to be* Dat is jammer *That is a shame*
		Bedankt voor uw commentaar *Thank you for your commentary*
Deze lepel is vuil *This spoon is dirty*	Request Reply (RR)	Ik kan even een nieuwe voor u halen *I can get a new one for you*

	Ik pak meteen een nieuwe voor u *I am getting a new one for you right away*
	Excuses u krijgt een nieuwe lepel *My apologies, you will get a new spoon*
	Ik zal het doorgeven aan de bediening *I will tell the service staff*
	Ik zal hem meenemen *I will take it away*
Neutral Reply (NR)	Dat is niet goed om te horen *That is not good to hear*
	Ik hoop dat ze u een nieuwe geven *I hope they will give you a new one*
	Dat hoort niet *That is not right*
	Weer een minpunt *Another bad* point
	Dat is niet hygienisch *That is not hygenic*

Appendix C: Response gap ERP analysis

The average gap duration was around 2300 ms (see Figure C1). To minimize the amount of motor artefacts we analyzed a gap duration window of 1500 ms (indicated by the red line if Figure C1). The pre-processing procedure was the same as for the main analyses of Experiment 2 with the exception that a segment of 2500 ms after critical word onset was used (1000 ms for the critical word and 1500 ms for the gap). The baseline was the same (i.e. 100 ms before word onset). Time-locked grand averages were computed for the four conditions and a cluster-based permutation test was conducted for the complete epoch from (0 to 1000 ms after critical word onset), the N400 window (350 to 600 ms) and the LPC window (700 – 900 ms). Both the interaction and the effect of sentence within each role were tested. The analysis revealed that also in the gap there were no significant clusters (all *p*-values > 0.158).

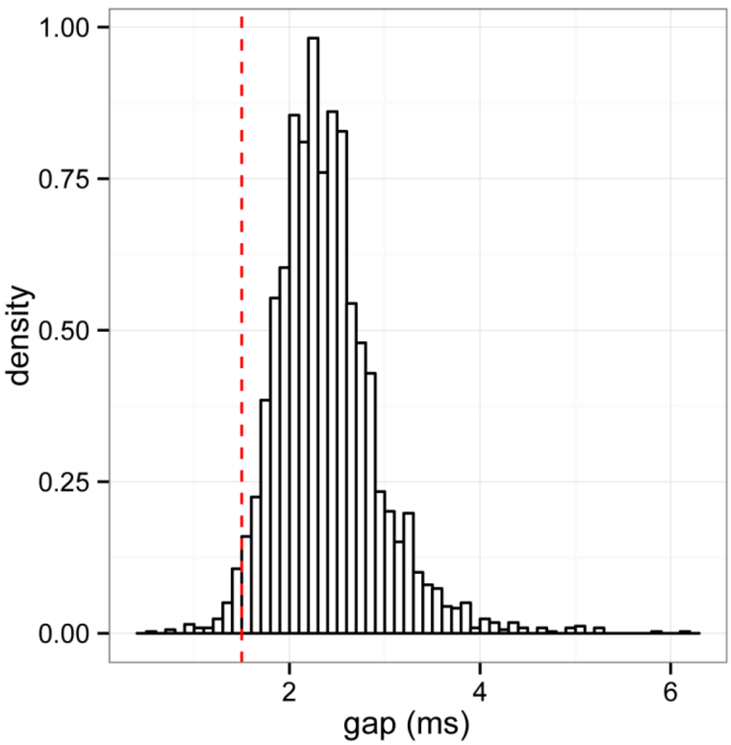

Figure C1: Histogram of the gap duration (ms). The red line indicates the end of the gap analysis window (i.e. 1500 ms)

Chapter 6

Summary and discussion

Summary and discussion

The first aim of this thesis was to shed light on the processes involved in bridging the gap between coded meaning (i.e. the linguistic code) and speaker meaning (i.e. what the speaker intended to communicate) for indirect request. In addition, virtual reality (VR) was explored as a new method to investigate language processing in a naturalistic but controlled experimental setting. In this chapter the main findings are summarized and discussed, and suggestions for future research are provided.

Summary of the main findings

Previous research has shown that fluctuations in pupil diameter can be used as an online monitor for cognitive effort during language comprehension (Sirois & Brisson, 2014). In Chapter 2 pupillometry was used to shed light on the cognitive effort involved in understanding indirect requests. Participants were presented with combinations of scenes and spoken sentences that taken together could be interpreted as either a request (e.g., a picture of a window with the sentence “It’s very hot here”) or a statement (e.g., a picture of a window with the sentence “It’s very nice here”). Participants were asked to indicate whether the speaker made a request or not. An increase in pupil diameter was observed in a time window of 1500 milliseconds following the critical word (“hot” or “nice” in “It is very hot/nice here”) for indirect request compared to statements. This difference was observed regardless of whether pupil size was compared in the experimental conditions or on the basis of the participant’s categorizations. These results suggest that processing the indirect requests required cognitive effort beyond the effort necessary to process a (direct) statement. These findings are in line with the claim derived from relevance theory that pragmatic inferencing requires processing effort (Sperber & Wilson, 1995). On a more methodological level, these results show that pupillometry can be used to study pragmatic processing, even in the context of visual information.

In Chapter 3, the relation between pragmatic abilities and cognitive effort during indirect request processing was explored. In other fields of pragmatics, it has been shown that an individuals’ pragmatic abilities play a role in how pragmatic meaning is understood (e.g., Nieuwland, Ditmar, & Kuperberg, 2010). Therefore, it might be the case that pragmatic abilities play a role in indirect request processing as well. As in Chapter 2, participants were presented with scene-sentence combinations interpreted as either a request (e.g., a picture of a window with the sentence "It's very hot here") or a statement (e.g., a picture of a window with the sentence "It's very nice here"). The communication subscale of the Autism Quotient (Baron-Cohen et al., 2001) was used to assess pragmatic abilities. It was hypothesized that if a request was correctly identified, this would require more cognitive effort for pragmatically less skilled participants than for participants with better pragmatic skills. This hypothesis was confirmed by a correlation analysis. However, the results of the linear mixed-effects analysis revealed only a marginal increase in the fit of the model with the inclusion of an interaction between pragmatic abilities and sentence type. These results provide some initial evidence for a link between pragmatic abilities and cognitive effort during indirect request comprehension.

The second aim of Chapter 3 was to replicate the results of Chapter 2, namely that pupil diameter should overall be larger for requests than for statements. However, this prediction was not confirmed. The possible reasons for this are discussed in more detail in the section below.

Chapter 4 tested the feasibility of the combined use of EEG and VR to study language processing in a contextually rich virtual environment. This is important, since it is hard to design experiments that resemble the everyday environments in which we often comprehend language, while maintaining control over the linguistic and non-linguistic information that is provided, especially if EEG is used. Participants were immersed in a virtual restaurant in which they were presented with spoken utterances from virtual restaurant guests. Each guest was seated at a table with a food or drink item on in front of them (e.g., a plate of salmon). The utterance of the guest could either match (e.g., "I just ordered this salmon") or mismatch (e.g., "I just ordered this pasta") with the object on the table. The results revealed a reliable N400 effect; ERPs were more negative for the mismatch than the match condition. This suggests that VR and EEG can be combined to study language processing in a more naturalistic environment.

In Chapter 5, indirect request processing was investigated in a more naturalistic virtual environment. As in Chapter 4, participants were immersed in a virtual restaurant where they encountered restaurant guests seated at different tables. The guest produced a sentence related to the object on the table in front of them (e.g., a bowl of soup). The sentences could either be a possible indirect request (e.g., “My soup is cold”) or a statement (e.g., “My soup is nice”). In addition, participants were assigned a role before the start of the experiment, namely to be a waiter or a restaurant critic. This was accomplished by means of a brief instruction and a virtual mirror in the restaurant in which they could see themselves in the outfit appropriate for the role. This new *virtual restaurant paradigm* has several advantages over more classical paradigms, namely that (1) there is no reference to the experimental manipulation in the task; (2) participants are in a rich meaningful environment, and (3) participants are direct addressees in a small but somewhat interactive context.

The results of Experiment 1 revealed that waiters interpreted the possible request as requests and the statements as statements, while restaurant critics interpreted both sentence types as statements. In Experiment 2 the same paradigm was used while the electroencephalogram (EEG) was measured. The prediction was a difference in the event-related potentials (ERPs) between possible requests and statements for the waiters, but not for the critics. However, this prediction was not confirmed. There were no differences in the ERPs between possible requests and statements for the waiters, there was only a main effect of the participant’s role on the ERPs (see Chapter 5 for a discussion). On a methodological level, these experiments can be considered a first attempt toward a more engaging interactive paradigm to study pragmatic processing in a more ecologically valid environment.

Cognitive effort and indirect request comprehension

In this thesis, pupillometry was used to investigate the cognitive effort involved in indirect request comprehension (Chapter 2) and the relation between pragmatic abilities and cognitive effort during this process (Chapter 3). Before the implications of the results from these chapters are discussed in the broader context of indirect language processing, there are some important differences between these chapters that need to be addressed. In Chapter 2, an increase in pupil diameter (and presumably cognitive effort) was observed for indirect requests compared to statements. However, this effect was not replicated in Chapter 3. Rather, in this chapter, there was evidence for a link between pragmatic abilities and cognitive effort; better pragmatic skills were associated with less cognitive effort during indirect request processing.

The experiments in these chapters were conducted with different participant groups and different materials. However, since both groups came from the Radboud University and they were all between 18 and 30 years old it seems unlikely that this caused the differences in the observed results. A more likely explanation has to do with the second difference; the experimental materials. In Chapter 2, only 25% percent of the scene-sentence combinations could be interpreted as requests and thus required a request reply. In contrast, in Chapter 3, 50% of scene-sentence combinations could be identified as a request. Since all other experimental factors (i.e. instruction, procedure) were the same in the two chapters, the list context (i.e. the mixture of items included in the stimulus list) is the most likely candidate to explain the difference in results. Influences of list context have also been observed on other effects in language processing. For example, in the priming literature it has been found that list context, for example the proportion of repetition versus unrelated primes in list of experimental items, can influence the size of priming effect (e.g., Bodner & Masson, 2001).

A list context effect on the pupillometry results in the current thesis is not necessarily problematic. However, it does have consequences for the generalizability of the findings from these chapters. Specifically, it does not seem to be the case that processing indirect requests always requires more cognitive effort than processing statements. Rather, the cognitive effort necessary for indirect request comprehension depends on the experimental list context and on the pragmatic skills of the listener. This is interesting, since it suggests that, contrary to the claims made in relevance theory (Sperber & Wilson, 1995), pragmatic inferencing might require different amounts of processing effort for different people, and in different context. In addition, even though theories on more direct access (Gibbs, 1994) versus indirect access (Searle, 1975) to non-literal meaning were not directly tested in these experiments, one could argue that if indirect language processing always requires activation of the literal meaning first, additional processing costs should always be observed whenever participants interpreted a negative state remark as a request. However, the results from Chapter 3 suggest that this might not be the case every individual and in all experimental contexts.

An important challenge for future research is to systematically characterize the factors that may play a role in the cognitive effort necessary to understand indirect requests. Rather than investigating whether or not indirect requests require more effort to process than statements, it might be more fruitful to investigate the conditions under which potential difference in cognitive effort can be observed. Most likely these factors are not limited to the stimulus list and individual differences in pragmatic abilities. Other factors might include task demands (i.e. what response is required), timing (i.e. within what time frame is the response required) and context (i.e. what information is available to aid the interpretation of the utterance). In addition, future research on individual differences in indirect request processing could focus on more diverse participant groups, such as older adults or patient groups, and on the underlying causes for the observed individual differences. For example, do pragmatically skilled individuals use qualitatively different strategies for indirect request processing than those that are less skilled?

Direct or indirect access to the meaning of indirect requests

In the Introduction of this thesis, several theories were discussed regarding the processing steps involved in indirect language comprehension (Gibbs, 1994; Giora, 2003; Grice, 1975). These theories differ in whether access to the literal meaning is necessary before one can understand the speaker meaning (i.e. more indirect access views; Grice, 1975, Searle, 1969) or whether it is possible to access the indirect meaning of an utterance more directly and without activation of the literal meaning (direct access view; Gibbs, 1994). As discussed in Chapter 5, predictions from these theories are often operationalized in terms of the (timing of) ERP components that they would affect (e.g., Bambini, 2016). The aim of Chapter 5 was to contribute to this debate by investigating the ERPs to requests versus statements in a new more naturalistic paradigm (i.e. the virtual restaurant paradigm). The behavioral results of Chapter 5 are encouraging in that participants were biased toward a specific interpretation of the experimental sentences without reference to the experimental manipulation in the task, they were directly addressed, and they understood the sentences in a richer more meaningful and somewhat more interactive environment. However, no reliable differences were observed between requests and statements in the ERPs. Thus, the results of Chapter 5 cannot shed light on the indirect versus direct access debate and more research on this issue is necessary.

Indirect request processing

The first aim of this thesis was to provide insight into the nature of indirect request processing. Although, as discussed above, it is not possible to discuss this process in terms of the indirect versus direct access debate, there are some general ideas on indirect request processing that follow from the results in this thesis.

First, as discussed above, the results from Chapters 2 and 3 suggest that the cognitive effort involved in indirect request processing is sensitive to experimental list context and individual differences in pragmatic abilities. Thus, more generally, for the investigations of how we bridge the gap between coded meaning and speaker meaning, researchers should take into consideration that that individual differences in pragmatic skills and differences in experimental list composition might have a considerable influence on the possible outcomes.

Second, the behavioral results of both experiments in Chapter 5 suggest that the participant's role in the experiment played a large part in how the negative state remarks were processed. Overall, only waiters interpreted the possible indirect request as requests, while the critics did not. This suggests that indirect request processing is sensitive to contextual information. Furthermore, this contextual information (i.e. the role of the participant) invoked the use of world knowledge and knowledge of social conventions. World knowledge played a role in the sense that we know that the role of a waiter in a restaurant is (usually) to keep people satisfied and happy (i.e., relieve negative states if there are any). Social conventions play a role in that they dictate that a restaurant guest can make a request to a waiter, but not to a restaurant critic. This idea is in line with the results from a study by Holtgraves (1994). This study revealed that speaker status has a significant effect on the speed of indirect request comprehension. Specifically, comprehending negative state remarks was time-consuming and involved activation of the literal meaning of the remark if the interlocutors were of equal status. However, when the speaker was of higher status, this difference disappeared. Furthermore, as discussed in Chapter 5, Coulson & Lovett (2010) take the LPC observed for indirect requests in their study to reflect the reader drawing on information from the preceding story context. Thus, the results of Chapter 5 provide converging evidence for the hypothesis that processing indirect request is sensitive to several (contextual) factors, in particular, the relation between the speaker and the listener, world knowledge and social conventions.

An important avenue for future research is to investigate the mechanisms that underlie the use of different types of information during pragmatic processing in different (experimental) contexts in real time. A possible candidate for this mechanism is prediction (Van Berkum, 2010). Recently it has been suggested that prediction plays a significant role in language comprehension; listeners regularly use their knowledge of the wider discourse context to predict specific upcoming words (e.g., Delong, Urbach, & Kutas, 2005; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort., 2015; Wicha, Moreno, & Kutas, 2004). However, this story becomes more complicated for pragmatic processing, especially in a rich context like the virtual restaurant paradigm, since many cues are provided that can be used to anticipate what comes next. Furthermore, one would have to consider what the contents of these 'pragmatic' predictions could be, since predicting the upcoming word might be of little help for understanding the speaker meaning of, for example, a negative state remark. Rather, one might have to predict on a different level, such as a communicative 'move' (Van Berkum, 2010) or a speech act (Gisladottir, Chwilla, & Levinson, 2014).

One final remark regarding the investigation of indirect request processing and pragmatic processing in general is concerned with the nature and ecological validity of experimental paradigms. In Chapters 2 and 3 participants were asked to make a meta-linguistic judgment regarding the utterance that they heard (i.e. Did the speaker make a request or not?), while in the virtual restaurant paradigm (Chapter 5), participants were asked to listen to (indirect) sentences and then come up with a response that was relevant to the (implicit) meaning of the sentence they heard. Although no reliable differences in the ERPs were observed in Chapter 5, there were significant differences in the behavioral responses. This suggests that one can study indirect request processing in a more naturalistic setting without an explicit task or instruction. Although the virtual restaurant paradigm is by no means real conversation, it could be considered more 'conversation-like', since the participant has his/her own turn in response to the turn of the restaurant guest. On a conceptual level, it can be argued that the demands on the participants in these two different paradigms (metalinguistic judgement versus responding) might have had considerable consequences for the way in which indirect requests were processed. Previously, it has been argued that conversation is cognitively demanding, because interlocutors have to produce and comprehend sentences very quickly (Clark, 1996; see also Holtgraves, 2008). Furthermore, it has been suggested that because of these real-time demands of conversation, participants in conversation might not fully process each turn in the conversation. Listeners might resort to a form of good-enough-processing (Ferreira, Bailey, & Ferraro, 2002) instead of engaging in a full-blown syntactic and semantic analysis (Holtgraves, 2008). This idea can be considered one of the ways in which the results obtained in more conversation-like paradigms might differ from more passive listening paradigms. The demands on the participant are quite different. It might be interesting for future research to develop and test more specific processing models for understanding indirect language in different contexts and under different task demands. Especially for indirect requests, this would include a more systematic description of what is meant by "understanding", "processing", and "recognizing" an indirect request. What does it really mean to *understand* an indirect request in conversation? One could argue that understanding a request goes beyond the mere recognition of the intended message as a request. It is our action in response to a request that makes communication successful. If this is the case, these actions should play a role in future experimental paradigms as well.

Indirectness versus speech acts

Throughout this thesis, indirect requests (negative state remarks) were compared to direct statements. Thus, the manipulations of indirectness (direct versus indirect) and speech acts (request versus statement) were confounded. In other words, indirect requests were compared not only to a direct utterance (a direct statement), but also to a different speech act (a statement). These dimensions are more often confounded in studies on indirect request processing (e.g., Coulson & Lovett, 2010; Holtgraves, 1994), especially if studies are theoretically grounded in the field of indirect language processing. One reason that these dimensions are not often manipulated orthogonally is that it is already quite difficult to construct carefully matched linguistic materials in experimental pragmatics, especially when negative state remarks are used. For example, the inclusion of a direct request would require using sentences like "Warm up my soup", which have a very different surface structure from negative state remarks and statements. Thus, it would be very difficult to use this type of sentence as a control for a negative state remark in an ERP experiment. However, for future research it would be valuable to disentangle these different manipulations to assess the relative effects of differences in speech acts and differences in directness on sentence processing. Furthermore, this type of experiment might be able to take a step forward in uniting theoretical and experimental research on speech acts (e.g., Egorova, Shtyrov, & Pulvermüller, 2013; Gisladottir, Chwilla, Levinson, 2014) and indirect language processing (e.g., Coulson & Lovett, 2010; Spotorno, Cheylus, Van Der Henst, & Noveck, 2013).

Virtual reality

In this thesis, VR was explored as a new method to investigate language processing in a naturalistic but controlled experimental setting. Researchers in social science, and in particular social neurocognition, have stressed the need for new more ecologically valid interactive paradigms in which to study social phenomena (Hari, Henriksson, Malinen, & Parkkonen, 2015; Schilbach et al., 2013). Although the use of VR as a method in the social sciences is relatively new (Fox, 2009), it provides some interesting opportunities to create more realistic and engaging, well controlled experiments. As discussed in Chapter 4, for language research in particular, VR allows researcher to include a standardized interaction partner (Bombari, Mast, Canadas, & Bachmann, 2015), thus allowing for the inclusion of a completely controlled yet visible speaker (Hoeks & Brouwer, 2014). In addition, it allows for the manipulation of contextual factors that are hard to replicate or control in the lab, for example the environment setting (e.g., a restaurant), the identity of the listener (e.g., a waiter) and referents (e.g., objects on the table).

In the present work, VR was used in combination with EEG to study audio-visual integration and indirect request processing. The replication of the N400 effect to audio-visual mismatches in VR (Chapter 4) provides an encouraging first step toward the combined use of VR and EEG to study language processing. In addition, the behavioral results of Chapter 5, namely that utterances were interpreted differently depending on the role of the participant, suggest that VR might be a useful tool for experimental pragmatics as well. The virtual restaurant paradigm is an example of a paradigm that would be much harder to realize outside VR, since it required a restaurant setting, a mirror for role assignment and several different speakers that were kept identical for each participant. Another topic in pragmatics that might especially benefit from the use of VR as a method is the study of non-verbal behavior, such as gestures or eye-gaze. For example, for indirect requests it has been shown that pointing gestures can aid the comprehension of indirect requests (Kelly, Barr, Breckinridge Church, Lynch, 1999). It could be of interest to investigate when and how information from gestures is integrated during indirect request processing.

Even though VR is a promising tool for increasing ecological validity in language research, it still has many challenges. As discussed in Chapter 5, the absence of reliable ERP effects might in part have been due to some aspects of the VR paradigm (see Chapter 5 for a discussion). Furthermore, a substantial amount of programming work is necessary to create a virtual environment and virtual interlocutors. Thus, before more standardized virtual worlds are available, VR experiments will be time-consuming. Also, specifically in combination with EEG, VR can increase the amount of movement artifacts in the data relative to EEG studies in which participants are required to sit still in front of a computer screen. Researchers should take this into account an adjust the number of participants and trials accordingly. Finally, because VR is such a new technique it is difficult to formulate hypotheses for VR experiments solely based on theories and results from experiments with more classical paradigms. It is conceivable that results from more classical experimental settings do not translate directly to richer virtual environments. To facilitate the comparison between results from more classical paradigms and VR paradigms, researchers should carefully consider which steps they take in 'going ecologically valid'. As discussed in Chapter 5, rather than changing the environment, the task and the design all at once, it might be more valuable to test smaller changes step by step. Then, even if different results are obtained in VR than in other experimental settings, these differences are more meaningful and they can help in the development of theoretical models of language processing in more realistic settings (e.g., Knoeferle, 2015). In sum, VR might proof to be a great tool to shed new light on many linguistic phenomena, as long as researchers will bound themselves by a bit more than just the limits of their imagination.

Conclusion

Understanding language goes beyond coding and decoding linguistic utterances. It is the intention of the speaker that drives the listener's behavior rather than his or her actual words. A waiter in a restaurant can easily interpret the sentence "Our bottle is empty" as a request to get a new bottle rather than a literal comment on the state of the bottle. He can understand what the speaker intended to communicate, even when the linguistic code itself did not specify a request at all. The aim of this thesis was to contribute to our understanding of how we bridge the gap between what is said and what is intended by investigating the processes involved in understanding non-conventional indirect requests. Furthermore, virtual reality was explored as a method to investigate these processes in a naturalistic but controlled experimental setting.

This thesis shows that comprehending indirect requests can be more effortful than comprehending statements. However, this increase in cognitive effort is modulated by the experimental (list) context in which the requests appear as well as by the pragmatic abilities of the listener. Furthermore, it shows that listeners use several pieces of information to guide their interpretation of possible indirect requests, specifically their role and relation to the speaker, world knowledge and knowledge of social conventions.

On a methodological level, this thesis provides evidence for the feasibility of the combined use of EEG and VR to study language processing. In addition, it might inspire future research in pragmatics by providing a first step towards a more realistic and interactive paradigm that could do justice to the rich and dynamic environments in which we usually engage in pragmatic processing.

References

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016). Disentangling metaphor from context: an ERP study. *Frontiers in psychology*, *7*.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17.

Bodner, G. E., & Masson, M. E. (2001). Prime validity affects masked repetition priming: Evidence for an episodic resource account of priming. *Journal of Memory and Language, 45*(4), 616-647.

Bombari, D., Mast, M. S., Canadas, E., & Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: Virtues, pitfalls, and future challenges. *Frontiers in Psychology*, 6.

Coulson, S., & Lovett, C. (2010). Comprehension of non-conventional indirect requests: An event-related brain potential study. *Italian Journal of Linguistics, 22*(1), 107-124.

Clark, H. H. (1996). *Using language.* Cambridge: Cambridge University press.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117-1121.

Egorova, N., Shtyrov, Y., & Pulvermüller, F. (2013). Early and parallel processing of pragmatic and semantic information in speech acts: neurophysiological evidence. *Frontiers in Human Neuroscience,* 7.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*(1), 11-15.

Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding.* Cambridge: Cambridge University Press.

Giora, R. (2003). *On our mind: Salience, context, and figurative language.* Oxford: Oxford University Press.

Gisladottir, R. S., Chwilla, D. J., & Levinson, S. C. (2015). Conversation electrified: ERP correlates of speech act recognition in underspecified utterances. *PloS One, 10*(3), e0120068.

Hagoort, P., & Levinson, S. C. (2014). Neuropragmatics. In M.S. Gazzaniga, & G.R. Mangun (Eds.), *The Cognitive Neurosciences* (pp. 667–674). Cambridge, MA: MIT Press.

Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of social interaction in human brain function. *Neuron, 88*(1), 181-193.

Hoeks, J. C., & Brouwer, H. (2014). Electrophysiological Research on Conversation and Discourse. Holtgraves, T. (Ed.), *Oxford Handbook of Language and Social Psychology*, pp 365-386. New York: Oxford University press.

Holtgraves, T. (1994). Communication in context: Effects of speaker status on the comprehension of indirect requests. *Journal of Experimental Psychology: Learning Memory and Cognition, 20*(5), 1215–1208.

Holtgraves, T. (2008). Automatic intention recognition in conversation processing. *Journal of Memory and Language, 58*(3), 627-645.

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.

Knoeferle, P. (2015). Language comprehension in rich non-linguistic contexts: combining eye tracking and event-related brain potentials. In R.M. Willems (Ed), *Cognitive Neuroscience of Natural Language Use* (pp. 77-100). Cambridge: Cambridge University Press.

Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language, 63*(3), 324-346.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences, 36*(04), 393-414.

Searle, J. R. (1975). Indirect speech acts. *Syntax and Semantics, 3*, 59–82.

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science, 5*(6), 679-692.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Oxford: UK Blackwell Publishers.

Spotorno, N., Cheylus, A., Van Der Henst, J. B., & Noveck, I. A. (2013). What's behind a P600? Integration operations during irony processing. *Plos One, 8*(6), e66839.

Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 443.

Van Berkum, J. J. (2010). The brain is a prediction machine that cares about good and bad-any implications for neuropragmatics? *Italian Journal of Linguistics, 22*, 181-208.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*(7), 1272-1288.

Nederlandse samenvatting

Stel je voor dat je in een restaurant bent met je vrienden of familie. Op een gegeven moment komt de ober langs om te vragen of je nog iets nodig hebt en je zegt tegen de ober "Onze fles wijn is leeg". Waarschijnlijk snapt de ober meteen dat dit een verzoek is om een nieuwe fles wijn, ook al heb je hem of haar niet letterlijk om een nieuwe fles gevraagd. De ober heeft de *speaker meaning* (wat de spreker wilde communiceren) begrepen, ook al was de *coded meaning* (de linguïstische code) helemaal geen verzoek. Binnen het gebied van de taalkunde dat dit soort processen bestudeert, pragmatiek, wordt dit een pragmatische analyse van de zin genoemd. In alledaagse communicatie zijn zinnen vaker indirect en ondergespecificeerd wat betreft betekenis. Het gedrag van de luisteraar wordt regelmatig bepaald door wat de spreker bedoelt te communiceren en niet per se door wat er daadwerkelijk wordt gezegd. Het doel van dit proefschrift was om meer te weten te komen over de processen die nodig zijn om indirecte verzoeken te begrijpen. Is het bijvoorbeeld moeilijker om een indirect verzoek te begrijpen dan een letterlijke zin zonder verborgen betekenis of hangt dit af van de omgeving waarin de zinnen worden uitgesproken? In dit proefschrift werden verschillende methoden gebruikt, namelijk pupillometrie (het meten van de pupilgrootte) en electro-encephalografie (EEG; het meten van de elektrische activiteit van de hersenen). Verder werd in dit proefschrift een relatief nieuwe methode, namelijk virtual reality (VR), gebruikt om deze processen te onderzoeken in een meer realistische omgeving.

Eerder onderzoek heeft laten zien dat fluctuaties in de grootte van de pupil gebruikt kunnen worden als een maatstaf voor cognitieve inspanning tijdens het begrijpen van taal. Hoe groter de pupil, hoe groter de inspanning. In Hoofdstuk 2 werd deze techniek gebruikt om te onderzoeken hoeveel cognitieve inspanning er nodig is om indirecte verzoeken te begrijpen. Proefpersonen zagen foto's en hoorden daarbij gesproken zinnen, die samen geïnterpreteerd konden worden als een verzoek (bijvoorbeeld een plaatje van een raam met de zin "Het is warm hier") of een letterlijke uitspraak (hetzelfde plaatje maar nu met de zin "Het is mooi hier"). De resultaten lieten zien dat proefpersonen meer cognitieve inspanning nodig hadden om indirecte verzoeken te verwerken dan directe uitspraken met dezelfde grammaticale structuur. Dit suggereert dat het begrijpen van een verzoek dat niet letterlijk wordt uitgesproken meer moeite kost dan het begrijpen van de letterlijke boodschap van een vergelijkbare zin. Daarnaast geven deze resultaten aan dat pupillometrie een goede methode is om pragmatische taalverwerking te bestuderen, zelfs als naast spraak ook visuele informatie wordt aangeboden.

In Hoofdstuk 3 werd de relatie tussen individuele verschillen in pragmatische capaciteiten en cognitieve inspanning tijdens het verwerken van indirecte verzoeken onderzocht. Pragmatische capaciteiten zijn een maatstaf voor hoe goed iemand zijn of haar kennis van het gebruik van taal in sociale situaties kan toepassen. Uit onderzoek is gebleken dat verschillen in pragmatische capaciteiten een rol spelen tijdens het maken van een pragmatische analyse van een zin. Het is dus mogelijk dat zulke individuelen verschillen ook een rol spelen bij het begrijpen van indirecte verzoeken. Net als in Hoofdstuk 2 werden proefpersonen gepresenteerd met combinaties van scènes en gesproken zinnen die samen geïnterpreteerd konden worden als een verzoek of een letterlijke uitspraak. De pupilgrootte werd ook hier gebruikt als maat voor cognitieve inspanning. Daarnaast werden de pragmatische capaciteiten van de proefpersoon gemeten door middel van een vragenlijst. Uit dit experiment bleek dat het correct identificeren van een indirect verzoek minder moeite kostte voor personen die beter waren in pragmatiek dan voor personen met minder goede pragmatische capaciteiten.

In Hoofdstuk 4 en 5 werd VR gebruikt als methode. VR is een virtuele werkelijkheid waarin een omgeving wordt gesimuleerd via een computer. De gebruiker draagt vaak een headset of een bril zodat hij of zij een driedimensionaal beeld heeft van de visuele wereld. Het gebruik van VR kan nuttig zijn voor onderzoek naar taalverwerking omdat het vaak moeilijk is om experimenten te ontwerpen die lijken op alledaagse omgevingen (waarin we normaal taal gebruiken). Tegelijkertijd kan er met VR strenge controle worden gehouden over de informatie die wordt gepresenteerd aan de proefpersonen.

In Hoofdstuk 4 werd onderzocht of het mogelijk is om op een betrouwbare manier hersenactiviteit, door middel van EEG, te onderzoeken tijdens taalverwerking in een meer realistische en natuurlijke omgeving dan in de meeste traditionele experimenten. In traditionele experimenten worden zinnen vaak gepresenteerd zonder verdere informatie over de omgeving waarin ze normaal gesproken worden uitgesproken. Vaak is dit geen probleem voor het onderzoek van taalverwerking, maar voor pragmatiek kan dit problematisch zijn omdat personen vaak informatie uit de directe omgeving gebruiken om een pragmatische analyse te maken. Voordat VR echter gebruikt kan worden in taalonderzoek is het belangrijk om te weten of we in VR dezelfde effecten in het hersensignaal kunnen observeren als in andere experimenten. In dit experiment ‘zaten’ proefpersonen in een virtueel restaurant waarin ze zinnen hoorden die werden uitgesproken door de virtuele restaurantgasten. Elke gast zat aan een tafel met iets te eten of te drinken voor zich (bijvoorbeeld een bord met zalm). De gesproken zin kon kloppen met het item op de tafel (bijvoorbeeld “Ik heb deze zalm net besteld”) of niet kloppen met het item op tafel (bijvoorbeeld “Ik heb deze pasta net besteld”). De resultaten lieten een betrouwbaar verschil zien in het hersensignaal tussen de juiste zinnen en de onjuiste zinnen, zoals al eerder werd aangetoond in traditionele experimenten. Dit geeft aan dat VR en EEG gecombineerd kunnen worden om taalverwerking te onderzoeken in een meer naturalistische omgeving.

In Hoofdstuk 5 werd de verwerking van indirecte verzoeken onderzocht door middel van VR en EEG. Tijdens het verwerken van indirecte verzoeken in een alledaagse omgeving gebruiken mensen zowel hun bestaande wereld kennis als informatie vanuit de directe omgeving om de zin goed te interpreteren. Net als in Hoofdstuk 4 waren proefpersonen in een virtueel restaurant waarin ze gesproken zinnen hoorden van restaurant gasten. De proefpersoon was in dit experiment echter zelf geen gast, maar een ober of een restaurantcriticus. Deze rol werd toegekend aan het begin van het experiment aan de hand van een korte instructie en een spiegel in de virtuele wereld waarin de proefpersoon zichzelf zag in de passende outfit (bijvoorbeeld een gilet voor de ober). In dit hoofdstuk zei de gast een zin gerelateerd aan het object op de tafel voor zich (bijvoorbeeld een kom soep). De zin kon een mogelijk indirect verzoek zijn (bijvoorbeeld “Mijn soep is koud”) of een letterlijke uitspraak (bijvoorbeeld “Mijn soep is lekker”). De taak van de proefpersoon was om kort antwoord te geven op wat de restaurantgast zei. De resultaten lieten zien dat de obers mogelijke verzoeken ook daadwerkelijk interpreteerden als een verzoek, terwijl de restaurantcritici dit bijna nooit deden. Dit geeft aan dat de proefpersoon zijn rol in het experiment gebruikte om de zin te interpreteren. Oftewel, de context speelt een belangrijke rol bij het begrijpen van indirecte zinnen. De resultaten van de hersenmetingen waren minder duidelijk, er werd geen verschil gevonden tussen de hersenactiviteit tijdens het verwerken van indirect verzoeken en letterlijke zinnen. Dit kan een aantal redenen hebben. Het kan bijvoorbeeld zo zijn dat er met veel informatie uit de directe omgeving een minder groot verschil is tussen het verwerken van indirecte verzoeken en letterlijke uitspraken. In dit geval zou de informatie uit de omgeving het dus makkelijker maker om de indirecte verzoeken te verwerken. Op basis van de resultaten in Hoofdstuk 5 is het echter niet mogelijk om hier een betrouwbare uitspraak over te doen, dus meer onderzoek in deze richting is nodig in de toekomst.

Samengevat laten de resultaten in dit proefschrift zien dat het begrijpen van indirecte verzoeken meer cognitieve inspanning vereist dan het begrijpen van letterlijke zinnen. Echter, de pragmatische capaciteiten van de persoon spelen hier ook een rol in: voor iemand met meer pragmatische capaciteiten kost het begrijpen van indirecte verzoeken minder cognitieve inspanning dan voor iemand met minder pragmatische capaciteiten. Ook geven de resultaten aan dat luisteraars verschillende soorten informatie gebruiken om indirecte verzoeken te interpreteren, zoals hun wereld kennis en de directe context waarin ze de zin horen. Vanuit een methodologisch oogpunt kunnen de VR experimenten in dit proefschrift worden gezien als een eerste stap voor het onderzoeken van taalverwerking in een meer realistische maar toch gecontroleerde omgeving.

Acknowledgements

As during any adventure there were good times and bad times. I consider myself so lucky that there were always people around for both.

First of all, I would like to thank my supervisors, Antje Meyer and Peter Hagoort. Antje, thank you for letting me pursue my own interests, for your support, and for providing me with such valuable feedback so fast. You taught me how to structure my (often quite messy) ideas and you gave me more confidence in myself as a researcher. Peter, thank you for your great ideas and valuable feedback. I really appreciated your focus and open-mindedness during our meetings, even when half way during my PhD I suggested that we should now do three VR experiments.

David, thank you for going on the VR adventure with me. My thesis would not have been the same without your ideas, feedback, and advice. I still have the first drawing you made of the restaurant… I will sell it to the highest bidder when you are famous for leading the VR revolution in the language sciences (or for making the most inappropriate joke in the language sciences).

I would also really like to thank the members of my reading committee, Prof. Ivan Toni, Dr. Roel Willems, and Prof. Max Louwerse, for taking the time to read and evaluate my thesis. I really appreciate it.

Suzanne and Martine, thank you so much for agreeing to be my paranymphs. Suus, you really were my 'steun en toeverlaat' at the MPI. Thank you for taking the time to listen to me, to help me, and to distract me with your jokes and stories outside of work. Martine, thank you for always being there for me. You helped me to relativize things, relax and believe in myself. Let's keep on taking trips, talking for hours and having a glass of wine forever... otherwise life will get so boring!

A big thanks also to my office mates. Will, thank you for all the good conversation and for the fun times. I wish I had hugged you more when you were still in Nijmegen. Merel, thank you for always finding the time to take a walk with me and for all the help with Praat. I think our friendship and our (very real) podcast are two of the best things that came out of my time at the MPI!

I am very grateful to all the members of the Psychology of Language department that I got to meet over the years. Agnieszka, Alastair, Alma, Amie, Andrea, Annelies, Caitlin, Chu, Conny, Dan, Eirini, Evelyn, Falk, Florian, Greta, Hans Rutger, Jana, Jeroen, Joe, Joost, Limor, Markus, Maartje, Matthias, Merel, Nina, Philip, Renske, Sara, Saoradh, Shiri, Suzanne, Svetlana, Will, Xuefei and Zeshu. Thank you for your scientific input and for the fun times at and outside of work! A special thanks to Annelies, for the great lunch chats.

I am also grateful to the members of Neurobiology of Language department for all the inspiration and advice! This section will get way to long if I mention all of you, but I do want to thank Birgit in particular for helping me with all the testing in the last phase of my PhD.

For conversations and good times at conferences, pizza-evenings, and in and around the MPI, I want to thank Elliot, Elliott, Evelien, Franziska, Gwilym, Linda, and Lotte. Also a big thanks to my fellow 'Nijmeegse Taalmiddag' organizers, it was a lot of work but super fun!

I would like to thank Dirkje, Els and Kevin for all the great work for the IMPRS. You were always so enthusiastic and you were really there for us. I would also like to thank the administration and the library for creating such an organized and pleasant working environment.

I am grateful to the technical group for all their support. Especially Reiner Dirskmeyer, Ronal Fisher, Maarten van den Heuvel, and Alex Dukers. I want to thank Jeroen Derks in particular for creating the first virtual restaurant in Nijmegen for my experiments and for all your patience while listening to my often messy explanations and wishes. Also a big thanks to Albert Russel for all the work for the VR experiments.

Annika, Ashley, Bart, Emma, David (Italy), Giulio, Lisa, Maarten, Martine, Suzanne, Rick, and Tobias, life in Nijmegen would not have been the same without you! Thank you for your company and for the endless amount of conversations, barbeques, and cultuurcafedrinks.

A very big thank you also to my family. Mam, tijdens mijn PhD was de Sterappel mijn veilige thuishaven, alle momenten dat ik even weg was uit Nijmegen hebben me zo goed gedaan. Bedankt voor je support, je optimisme en je eeuwige geloof in mij als onderzoeker en persoon. Pap, toen ik klein was heb je me geleerd hoe auto's en computers werken en nog heel veel meer. Het meeste hiervan ben ik alweer vergeten, maar zonder je

support en de nieuwsgierigheid die je me toen al hebt bijgebracht was dit PhD avontuur nooit gelukt. Jouke, wat is het fijn om zo’n slim, lief en supportive broertje te hebben. Bedankt voor alle gesprekken en voor al je advies! Bernie, bedankt voor de fijne momenten en al het lekkere eten aan de Sprengenbergweide!

Nicco, thank you for being the wiser half of Tromberti. Thank you for listening, arguing, supporting and most of all believing. From the first cheesecake to the last bottle of wine, we made this thesis happen together. Piacere.

Curriculum Vitae

Johanne Tromp was born in 1989 in Ermelo, The Netherlands. She obtained her Bachelor's degree from University College Roosevelt (Utrecht University) in 2010, after which she completed the Research Master Cognitive Neuroscience at Radboud University in 2012. In 2013, she started her PhD project in the Psychology of Language department at the Max Planck Institute for Psycholinguistics in Nijmegen, funded by an International Max Planck Research School for Language Sciences Fellowship from the Max Planck Society. In 2016 she worked for six months as a policy officer at the Netherlands Organisation for Scientific Research in The Hague, after which she returned to Nijmegen to finish her PhD project. Currently, she works in Berlin as a postdoctoral researcher at the Berlin School of Mind and Brain and the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig.

Publications

Tromp, J., Peeters, D., Meyer, A. S., & Hagoort, P. (2018). The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behavior Research Methods, 50*(2), 862-869.

Tromp, J., Hagoort, P., & Meyer, A. S. (2016). Pupillometry reveals increased pupil size during indirect request comprehension. *The Quarterly Journal of Experimental Psychology, 69*(6), 1093-1108.

Verhees, M. W. F. T., Chwilla, D. J., Tromp, J., Vissers, C. T. W. M. (2015). Contributions of emotional state and attention to the processing of syntactic agreement errors: evidence from P600. *Frontiers in Psychology, 6*.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*

18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*

34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*

53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. *Split intransitivity in Rotokas, a Papuan language of Bougainville. Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*

69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*
83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*

84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*
85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*
86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*
87. The acquisition of morphophonological alternations across languages. *Helen Buckler*
88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*
89. Rediscovering a forgotten language. *Jiyoun Choi*
90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. lncrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*

100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*

101. Assessing birth language memory in young adoptees. Wencui Zhou

102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*

103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*

104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition. *Rósa Signý Gisladottir*

105. Modelling Multimodal Language Processing. *Alastair Smith*

106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*

107. Situational variation in non-native communication. *Huib Kouwenhoven*

108. Sustained attention in language production. *Suzanne Jongman*

109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*

110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*

111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*

112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*

113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*

114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*

115. Music and language comprehension in the brain. *Richard Kunert*

116. Comprehending Comprehension: Insights from neuronal oscillations on the neuronal basis of language. *Nietzsche H.L. Lam*

117. The biology of variation in anatomical brain asymmetries. *Tulio Guadalupe*

118. Language processing in a conversation context. *Lotte Schoot*

119. Achieving mutual understanding in Argentine Sign Language. *Elizabeth Manrique*

120. Talking Sense: the behavioural and neural correlates of sound symbolism. *Gwilym Lockwood*

121. Getting under your skin: The role of perspective and simulation of experience in narrative comprehension. *Franziska Hartung*

122. Sensorimotor experience in speech perception. *Will Schuerman*

123. Explorations of beta-band neural oscillations during language comprehension: Sentence processing and beyond. *Ashley Lewis*

124. Influences on the magnitude of syntactic priming. *Evelien Heyselaar*

125. Lapse organization in interaction. *Elliott Hoey*

126. The processing of reduced word pronunciation variants by natives and foreign language learners: Evidence from French casual speech. *Sophie Brand*

127. The neighbors will tell you what to expect: Effects of aging and predictability on language processing. *Cornelia Moers*

128. The role of voice and word order in incremental sentence processing. *Sebastian Sauppe*

129. Learning from the (un)expected: Age and individual differences in statistical learning and perceptual learning in speech. *Thordis Neger*

130. Mental representations of Dutch regular morphologically complex neologisms. *Laura de Vaan*

131. Speech production, perception, and input of simultaneous bilingual preschoolers: Evidence from voice onset time. *Antje Stoehr*

132. A holistic approach to understanding pre-history. *Vishnupriya Kolipakam*

133. Characterization of transcription factors in monogenic disorders of speech and language. *Sara Busquets Estruch*

134. Indirect request comprehension in different contexts. *Johanne Tromp*