KogWis 07

Fachrichtung Psychologie

Gesellschaft für Kognitionswissenschaft

Cognitive Science 2007

***Proceedings of the 8th Annual Conference of the Cognitive Science Society of Germany*.**

Zimmer, H. D.

Frings, C.,

Mecklinger, A.

Opitz, B.

Pospeschill, M.

Wentura, D. (Eds.)

Saarbrücken

URL: http://psydok.sulb.uni-saarland.de/volltexte/2008/1185/

URN: urn:nbn:de:bsz:291-psydok-11856

Cognitive Science 2007. Proceedings of the 8th Annual Conference of the Cognitive Science Society of Germany.

Content

Weiß, Petra[1]; Pfeiffer, Thies [2]; Schaffranietz, Gesche[1] & Rickheit, Gert[1] 1 Fakultät für Linguistik und Literaturwissenschaft/SFB 673 2 AI & VR Lab, Technische Fakultät Universität Bielefeld	Coordination in dialog: Alignment of object naming in the Jigsaw Map Game.
Baumotte, Henrike [1 2]; Müller, Horst M. [1] & Dogil, Grzegorz[2] 1 Experimental Neurolinguistics Group, University of Bielefeld, Germany 2 Experimental Phonetics Group, Institute for Natural Language Processing, Univ. of Stuttgart, Stuttgart	Lexical access from spectrum? Phonetic properties of proper names and common nouns in German and Mandarin Chinese
Baumotte, Henrike & Dogil , Grzegorz Experimental Phonetics Group, Institute for Natural Language Processing, Univ. of Stuttgart, Stuttgart	State of the art of phonetic language aptitude linking phonetic as well as phonological models to empirical neuroimaging (neurolinguistic) research
Lewandowski, Natalie[1]; Jilka[2], Matthias; Rota[1], Giuseppina; Reiterer[3,1], Susanne & Dogil[1], Grzegorz 1 Experimental Phonetics Group, Univ. of Stuttgart; 2 Institute for Linguistics, Anglistics, Univ. of Stuttgart; 3 Dept. of Neuroradiology, Section for Experimental MR of the CNS, Univ. of Tübingen	Phonetic convergence as a paradigm of showing phonetic talent in foreign language acquisition. Institute for Natural Language Processing,
Bermeitinger, Christina; Frings, Christian & Wentura, Dirk Cognitive Psychology – Cognition and Action Unit, Saarland Univ., Saarbrücken	Repeated masked semantic priming with new results: ERPs of a negative semantic priming effect.
Tenbrink, Thora SFB/TR 8 Spatial Cognition Univ. of Bremen, Bremen	Space, time, and the use of language: An investigation of relationships.

Strobach[1], Tilo; Frensch[1], Peter A.& Schubert[2], Torsten Dept. of Psychology, General Psychology, Humboldt Univ., Berlin; Dept. of Psychology, Experimental Psychology, Ludwig-Maximilians Univ., München	The temporal stability of skilled dual-task performance.
Aschersleben[1], Gisa & Müsseler[2], Jochen 1 Dept. of Psychology, Developmental Psychology Unit, Saarland Univ., Saarbrücken; 2 Psychology Dept., RWTH Aachen Univ., Aachen	Dual-task performance while driving a car: Age-related differences in critical situations.
Polushkina[1], Svetlana; Kühnberger[2], Kai-Uwe & Gust[2], Helmar 1 Research Training Group "Feedback Based Quality Management in eLearning" Univ. of Technology, Darmstadt, Darmstadt; 2 Institute of Cognitive Science, Univ. of Osnabrück, Osnabrück	Solving systems of equations by analogy: An experiment and a model.
Weiler[1], Daniel; Läer[1], Leonard; Engel[2], Andreas K. & König[1], Peter 1 Institute of Cognitive Science, Dept. Neurobiopsychology, Univ. of Osnabrück, Osnabrück; 2 Dept. of Neurophysiology and Pathophysiology, Univ. Medical Center Hamburg-Eppendorf, Hamburg	Unsupervised learning of reflexive and action-based affordances to model navigational behaviour.
Buechel, Eva; Vogt, Annamarie & Opwis, Klaus Dept. of Cognitive Psychology and Methodology, Univ. of Basel, Basel	Pseudoneglect as a function of hand use and line length.
Brinkmann, Thorsten & Zimmer, Hubert D. Brain and Cognition Unit, Dept. of Psychology, Saarland Univ., Saarbrücken.	Contributions of modality-specific brain regions to working memory and long-term memory: An ERP study.

Coordination in dialog: Alignment of object naming in the Jigsaw Map Game

PETRA WEIß[A], THIES PFEIFFER[B], GESCHE SCHAFFRANIETZ[A] AND GERT RICKHEIT[A]

[A]Fakultät für Linguistik und Literaturwissenschaft/SFB 673
[B]AI & VR Lab, Technische Fakultät
Universität Bielefeld[1]

People engaged in successful dialog have to share knowledge, e.g., when naming objects, for coordinating their actions. According to Clark (1996), this shared knowledge, the common ground, is explicitly established, particularly by negotiations. Pickering and Garrod (2004) propose with their alignment approach a more automatic and resource-sensitive mechanism based on priming. Within the collaborative research center (CRC) "Alignment in Communication" a series of experimental investigations of natural face-to-face dialogs should bring about vital evidence to arbitrate between the two positions. This series should ideally be based on a common setting. In this article we review experimental settings in this research line and refine a set of requirements. We then present a flexible design called the Jigsaw Map Game and demonstrate its applicability by reporting on a first experiment on object naming.

Motivation

"[H]umans are 'designed' for dialogue rather than monologue" – this is one of the central arguments brought forward by Garrod and Pickering (2004, p. 8) emphasizing the necessity for research on language use in interactive human communication. A central mechanism of successful communication is coordination. When people interact and communicate they need to tune their utterances on different levels, amongst which the lexical level plays a prominent role. The central questions are now, how coordination is achieved and how people build up shared knowledge during a dialog.

In psycholinguistics two main approaches exist addressing these questions. The collaborative model (Clark, 1996) proposes the idea of people building up a common ground, employing an explicit grounding process, e.g., by negotiation. In contrast, the notion of alignment (Pickering & Garrod, 2004) supposes that conversational partners build up an implicit common ground in an automatic and mechanistic way and fall back on utilizing full common ground only in cases of misalignment.

To develop an adequate experimental setting is of central importance for the CRC "Alignment in Communication", where besides lexical coordination also questions of coordination of sentence structure, spatial reference, prosody, gesture and gaze will be examined in a series of communication experiments. Data collected in these experiments will be used to build up a central corpus of interactive language data to function as the empirical basis of research in the CRC.

---

[1] Address: Dr. Petra Weiß, SFB 673, Fakultät für Linguistik und Literaturwissenschaft, Universität Bielefeld, Postfach 10 01 31, D-33501 Bielefeld, email: petra.weiss@uni-bielefeld.de, web: www.sfb673.org. Funded by German Research Foundation (DFG) within the Collaborative Research Center 673 (SFB 673) "Alignment in Communication". The authors want to thank Sara Maria Hellmann for her support in data analysis and Ulrike Well and the staff of the media lab (SCM) of Bielefeld University, especially Nora Johannesmann, for their technical support.

We discuss both psycholinguistic approaches and present successful experimental designs used in this line of research. We specifically highlight their particularities, but also their shortcomings with respect to our needs. This results in a set of requirements for an experimental design appropriate to investigate language coordination in interactive conversation in a natural but controlled way. Consequently, we propose the Jigsaw Map Game as a flexible design meeting these requirements and present first results from an experiment on the coordination of object naming.

Theoretical background

The collaborative model

Clark's (1996) idea is that language use is a joint action. Dialogical communication is a dynamic action. It is generated jointly by communicative partners and only successful when the action is coordinated. This way, language can only be analyzed adequately considering its social aspect. The basic setting for language use is spoken face-to-face conversation. Clark motivates this by giving three reasons: Spoken language is the most universal form of language use in all civilizations; there is no need for special skills in normal conversation and spoken language is essential in language acquisition. Clark and Brennan (1991) characterize face-to-face conversation by immediacy, no use of media leading to evanescence, recordlessness and simultaneity as well as full control of the participants what gets done and how. The less of these features are fulfilled the more demanding the communication between the interlocutors should be because they need more specialized skills and procedures to communicate adequately.

To the question, how people get coordinated to communicate successfully, the collaborative model proposes the idea of a common knowledge base, called *common ground*. "Two people's common ground is, in effect, the sum of their mutual, common, or joint knowledge, beliefs, and assumptions" (Clark, 1996, p. 93). Especially the idea of mutuality is very important. Every joint action is rooted in this common information, but only in the part people believe they share with others. What really is part of the common ground in a community can only be evaluated on a meta-level, but not by the community members themselves. Thus, there only exist individual assumptions about what the common ground comprises; this may lead to misunderstandings that have to be resolved explicitly.

In view of the fact that communication is a daily and natural human activity, the idea of conversation as a joint activity and the proposal of face-to-face interaction as the basic setting for language use seem plausible. On the other hand, the necessity to continuously update common ground information seems to be costly. The question arises, whether the explicit negotiation of information within the grounding process – though it is a typical phenomenon in conversation – really is a basic process. These and other critical points associated with the collaborative model led to the development of the alignment approach.

The notion of alignment

Pickering and Garrod (2004) suggest that dialog communication takes place easily and smoothly because basic interactive processing mechanisms permit coordination and adjustment of communicative partners in a cognitive less demanding way than explicit negotiation. They expand the preceding output/input coordination principle (Garrod & Anderson, 1987) to a mechanistic approach of interactive language processing emphasizing the necessity of aligned representations between partners on different linguistic levels.

The meaning of mechanistic in this context is twofold. On the one hand the authors propose priming as the central interactive processing mechanism leading to a coordinated knowledge base, i.e., a body of aligned representations. Simultaneous alignment on different linguistic levels of representation (phonetic, phonological, lexical, syntactic, semantic and the level of situation models) is based on the idea of percolation between the different levels, i.e., alignment at one level supports alignment on other levels (Garrod & Pickering, 2004; see also Branigan et al., 2000). On the other hand the approach is mechanistic because under default conditions coordination takes place in an automatic way without conscious control of the interlocutors.

To motivate their approach, Pickering and Garrod (2004) point out that the explicit computation of common ground is computationally too costly to be a realistic assumption for real life dialogs. Instead they propose that dialog participants usually only employ an *implicit* common ground – knowledge that is available to both speaker and hearer, but that does not include knowledge about the cognitive state of the other agent(s). Implicit common ground is usually inferred from information sources that are available to all agents (the immediate linguistic and situational context) and therefore provides a good approximation of *full* common ground, as Pickering and Garrod call by contrast the idea of common ground in the collaborative model. It is important to note that the authors do not totally deny the concept of full common ground and explicit grounding processes during conversation. But they advance the view that these concepts and processes cannot be the basic mechanisms for coordination in communication.

Based on the theoretical considerations so far our goal is to examine the question, how – under default conditions – coordination in dialog is reached. Before we will describe our own experimental paradigm developed to investigate into this question, we will give a short overview over some relevant experimental paradigms used in research on interactive language processing.

Language processing in interactive discourse

The empirical investigation of language processing in discourse has developed around a relatively small number of experimental situations (Garrod, 1999; Schober & Brennan, 2003). As language use in natural conversations cannot readily be controlled, some experimental designs have been developed to elicit semi-spontaneous dialog situations where some degree of control over the topic of conversation is possible.

Referential communication task studies

The referential communication task is primarily used to investigate lexical coordination. In an early study by Krauss and Weinheimer (1996) pairs of participants describe to each other a sequence of abstract shapes. The participants have fixed roles, one describes and the other tries to identify the figure. Krauss and Weinheimer demonstrate that the convergence on shortened names for low-codability shapes in the course of a conversation depends on the possibility of direct (verbal) interaction between the interlocutors, especially on getting concurrent feedback. Their results suggest that the referential process really depends on coordination. Clark and Wilkes-Gibbs (1986) show that the common ground between interlocutors is established as a consequence of coordinated actions among the interacting individuals across multiple linguistic exchanges and that this knowledge is consulted for the ongoing conversation.

Brennan and Clark (1996) suggest that for lexical coordination partners establish a conceptual pact during conversation, i.e., a temporary flexible agreement by partners to refer to and conceptualize an object in the same way. A conceptual pact is established by grounding; the partners mark having reached a conceptual pact by reusing the same or similar expressions in the ongoing conversation. Conceptual pacts are not restricted to a single dialog move; the more well established a pact is (the more often they have used it to refer to an object) the more likely it is to persist. In contrast to full common ground, it is not based on an explicit partner model, but may be shaped and maintained by the partner's feedback.

On the whole, results on default use of common ground information are partly inconsistent (e.g. Buhl, 2001). Therefore, the question arose under which conditions speakers are able or willing to consider common ground information. One important factor concerns the communicative situation established in the experiments. Experimental results show that speakers are more sensitive to the needs of visually co-present naïve listeners than to confederates (Schober & Brennan, 2003). This suggests, first, that the presence of a conversational partner with real needs matters and second, that people are sensitive to the behavioral naturalness of that person. These effects might be even stronger if interlocutors are allowed to interact in a real and natural way, when they are not restricted to speaking, but can nod, smile, point, gaze at each other and exhibit and place things (Clark & Krych, 2004). This being the case, employing confederates or not even using a second participant may result in findings that do not reflect the behavior in natural dialog situations.

Results gained within the referential communication paradigm illustrate that language processing is highly sensitive to the constraints imposed by interactive dialog, especially the importance of feedback and real interaction between communicative partners. However, whereas the referential communication task allows for a detailed analysis of referential processes, there always is a fixed role allocation between the communicative partners. One gives information and has a leading role and the other processes this information and has to fulfill a special task indicating correct or wrong interpretation, or one of them is even a confederate of the experimenter.

Maze game studies

In the maze game studies (Garrod & Anderson, 1987) two participants located in different rooms have to interactively get through a maze presented to them individually on a computer screen. The players, who are in audio contact with each other, can only solve the task by coordinating their moves via speech. They have to find a strategy of moving to positions where their partner has switch boxes in order to allow them a free route to the goal. As a result, the game elicits relatively free dialog sequences containing repeated location descriptions. Especially the analysis of how these descriptions develop during the course of the game leads to a number of insights about coordinated language use and interpretation (Garrod, 1999).

Though the dialogs turn out to be extraordinary varied across the sample of players, each description can be classified according to one of four basic schemes corresponding to a special combination of spatial conception of the maze together with a description lexicon. According to this classification, Garrod and Anderson (1987) analyze how the references developed during the course of the dialogs. Despite of the great variations across the whole corpus of dialogs any pair of interlocutors is very consistent in their choice of descriptions in any stretch of dialog. Conversationalists not only collaborate in establishing isolated references but also in formulating local description languages and contextual unambiguous dialog lexica. Furthermore, coordinating on a common description language is not simply a matter of sticking with the first reasonable scheme that emerges, but involves a much more extended history of development during which interlocutors explore different schemes in a coordinated fashion over a period of time. A central point of the analysis is that explicit negotiation does not seem to play an important role in either establishing a common description language or in fixing the language over the subsequent dialogs. If at all, negotiation can only be observed in later rounds of the game.

On the whole, the maze game studies highlight some of the ways in which language processing in a task allowing relatively free verbal interaction is affected by the demand of consensus in dialog. With both players sitting in different rooms, each of them presented with the maze on his/her monitor, the nature of the game demands that, though they have to cooperate, each of them pursuits his/her own goal. They have no common interaction space and no consistent conception of the world emerges during conversation. In contrast to the referential task there is no role differentiation between the partners resulting in a more interactive communication.

Map task studies

The map task was originally developed by Anderson et al. (1984) to explore properties of effective vs. ineffective communication in a pedagogic context. Subsequently, it has been used to investigate into a whole range of dialog-processing issues from reference and speech articulation to video mediated communication (Doherty-Sneddon et al., 1997).

Each participant has a map of an imaginary island containing a number of labeled landmarks. One participant has a route marked across the island, whereas the other just has the map containing the landmarks without the route. The task of one player is to communicate the route to the other player, so that she can draw the route on her map.

This basic setting offers the possibility of a lot of interesting manipulations, e.g., with respect to incompatibilities concerning the number or arrangement of the landmarks on the two maps. This allows to vary the degree of overlapping knowledge between communicators and to explore how they use different referential forms to signal their relative knowledge states. Particularly with this task an independent control of communicative success is possible by comparing the route drawn by the instruction follower with the route given on the map of the instruction giver.

The map task has contributed particularly in two ways to understanding of referential communication. First, in terms of how communicants deploy different linguistic devices to communicate about shared and unshared information (Anderson & Boyle, 1994). Second, in terms of how they combine multi-modal information in coordinating their productions and interpretations. Boyle et al. (1994) find that communicators who can see each other need much less amount of speech to convey the same amount of information as those who can not. Furthermore, communicators often look at each other during dialog segments where they are having problems in communicating, typically when there are incompatibilities between the two maps. Being able to monitor the face and the gaze of each other may help managing the exchange. Anderson et al. (1997) show that this monitoring does not reflect a moment-by-moment modeling of the listener's actual use of the visual channel, but that it seems to reflect a more general assessment of the state of the mutual intelligibility of the conversation at that time. This result is consistent with the findings from the maze game studies, where communicators cooperate to establish a mutually acceptable description scheme but without explicitly modeling their partner's knowledge state indicating that implicit processes are basic for interactive communicative coordination (Garrod, 1999).

Contrary to the maze game, in the map task again there is a clear role allocation of the partners like in the referential communication paradigm. The advantage here is that communication takes place about a special part of the world showing a landmark structure offering the possibility for structured spatial descriptions in a variety of communicative situations ranging from face-to-face dialogs to video- and computer-mediated communication. On the other hand, again no shared part of the world emerges during conversation.

Based on these considerations it becomes evident that for the examination of basic processes in natural face-to-face conversations none of the paradigms reported is sufficient on its own.

To examine the question how coordination in dialog is reached under default conditions requires a trade-off between natural communication and experimental control which is a big methodological challenge (Schober & Brennan, 2003). We try to overcome this problem by developing the *Jigsaw Map Game*, an experimental setting based on the referential communication task, combining elements of the maze game and the map task.

The Jigsaw Map Game

Following the lines of thought presented, we identify several objectives for our design of an experimental setting. The participants should

- … be naïve and not confederates.
- … be engaged in a face-to-face situation.
- … be able to directly perceive each other's behavior.
- … be able to communicate using several modalities.
- … be on equal terms regarding their roles.

The setting should allow for the most natural communication possible between the interlocutors under experimental conditions. At the same time means to control important parameters are needed. One should be able to

- … channel the course of the dialog.
- … control the task knowledge relevant for each participant.
- … induce specific conditions during the dialog.
- … assess data from several sensor devices, including eye tracking.

Encouraging natural face-to-face interaction

One of our primary objectives is to observe natural communication in dialog, so we decided to have a basic setting without any technical mediation, which is still flexible enough to allow mediation if needed. Our setting constitutes a face-to-face situation including physical objects which can be touched and handled by the participants. This promotes the presence of the participants within the setting (Lombard & Ditton, 1997). The objects should vary in shape, size, color or function to ensure a rich use of verbal and non-verbal communication. The interaction between the interlocutors and the objects happens in a distinct interaction space within a restricted area, e.g., defined by a desktop. On the technical level, the visual co-presence of objects and interlocutors in a face-to-face situation allows for multi-perspective video recordings as well as in-depth investigations using eye tracking, e.g., following the visual world paradigm.

Promoting multi-modal communication

From previous experiments (Kranstedt et al., 2006) we know, that a setting with a relatively small interaction space, may encourage a disproportionately high use of pointing or grasping gestures. This should be balanced in favor of more elaborate verbal object descriptions. For instance, the interlocutors can be made to introduce the objects into the interaction space, which is what we did in our first experiment. In the course of our setting each interlocutor has to ask her partner for specific objects she does not have access to. Each one can use verbal descriptions, iconic gestures or references to objects already in the setting sharing certain attributes.

Defining the task

With these decisions being made, most of the stated objectives have already been met. Previous work in the course of the CRC "Situated Artificial Communicators" had concentrated on complex construction tasks and tasks on object placement (Weiß et al., 2006). Inspired by the map task we decided to employ a combination of both types of tasks: the placement of objects according to a complex prescribed map (Fig. 1). The map is designed in such a way that some objects stand out as landmarks because of seize. These objects define the critical objects for the object naming task under examination in the experiment presented.

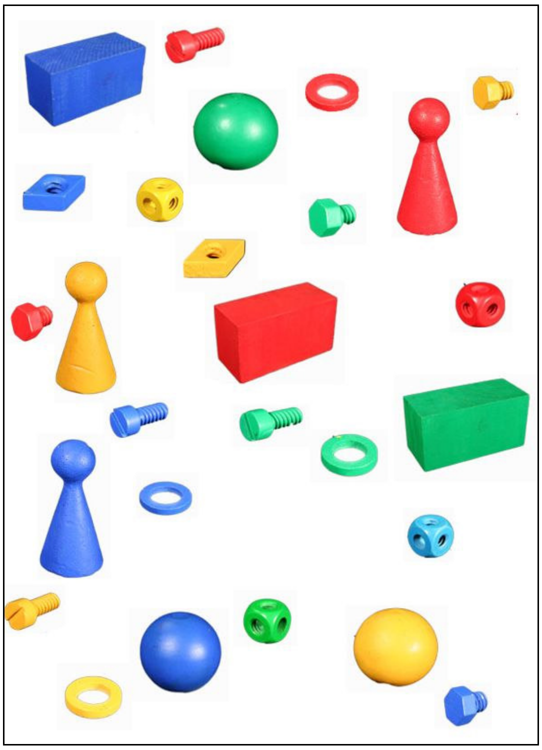


Fig. 1 Example for the full layout of the objects on the table: The critical objects for the naming task form landmarks because of their sizes and are expected to be referred to more often in the sense of spatial reference objects.

The goal of the game is that the communicators cooperatively build the common object arrangement shown on the map in Fig. 1. The cooperative character of the game emerges by the fact that none of the participants knows the complete object arrangement. Each partner only gets partial information concerning the constellation of maximal three objects (two already placed and a new one; see Fig. 2) at a time. She has to communicate to the partner what object the partner should pick next and where it has to be placed. This way the objects are related to each other, as in the construction task, but their relationship is focused on spatial relations as in the map task, leaving aside functional relations or role attributions. By placing the respective new object relatively to objects already placed we can differentiate between references identifying an object as a direct target object that has to be placed next and references identifying an object as an indirect reference object for the placement of the next target object.

Controlling the flow of the game and balancing out the roles of the participants

To control the task knowledge of each participant and at the same time channel the course of the dialog, we finally came up with the *Jigsaw Map Game*. Every participant is given an ordered set of small pieces of the map (Fig. 1), each showing only a small set of objects.

These pieces are carefully designed as in a jigsaw puzzle: They show exactly one new object in relation to several existing objects already in the interaction space (Fig. 2).

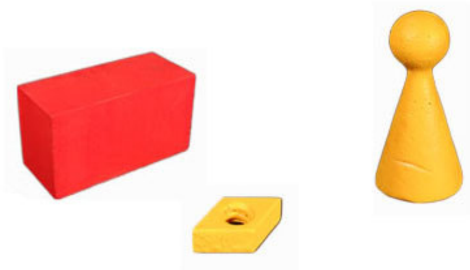
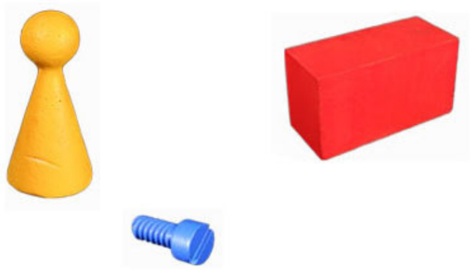
A	B
 1) A starts by drawing a piece of the puzzle showing a cuboid, asks B to fetch the “red box” from B’s object pool and to position it within the interaction space. A’s reference to the “red box” is a direct reference.	 2) Now B takes the next card and requests A to fetch and place a green bolt. In her instruction B might refer to the red cuboid as a landmark. This reference is an indirect reference.
 3) The next item A commands B to place is a yellow rhombus.	 4) Then it is B’s turn to instruct A to place the yellow token.
 5) A then requests B to put the blue bolt into position.	 6) B then asks A to situate the red bolt.

Fig. 2 The flow of the game is controlled by providing small amounts of local knowledge. Each piece of the map (examples 1 to 6) shows a maximum of two known and exactly one new object. Please note that the two sets of cards show the setting from different perspectives: While in turn 5 for A, the blue bolt is right before or below the yellow token, for B the blue bolt is to the left and behind or above the yellow token.

The participants now draw their pieces of the puzzle in turns. Each turn, the one with the new part of the map has to instruct the other one to get the new object out of a personal box of objects (each one has a set of pieces of the puzzle describing new objects to be found in the personal box of the other one) and position that new object in relation to the objects already placed in the interaction space. Once this has been accomplished, they change their roles and the one having placed the last object draws the next piece of the map. A part of the flow of the game is shown in Fig. 2.

Set-up and recording technique

We want to conclude the presentation of the design of the Jigsaw Map Game with a concrete example set-up, as it has been used and tested in our first experiment. The two participants are placed face-to-face on the longer sides of a standard office desk (see Fig. 3). In between them, an interaction space is defined by two yellow lines. The full interaction space has been carefully chosen to be totally within grasping range of both interlocutors. To each side of the interaction space there is room for the piles of new and used pieces of the map. Aside each participant there is a small box with the personal selection of objects.

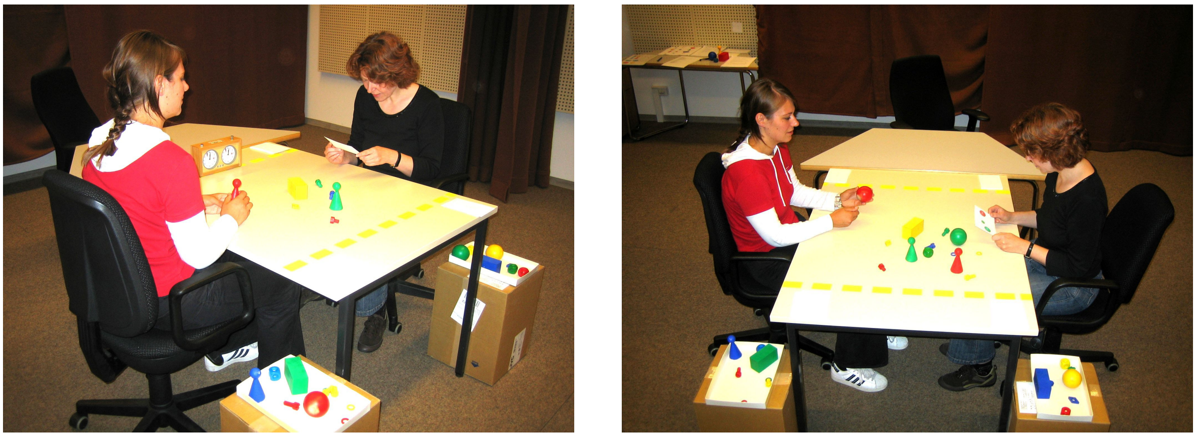


Fig. 3 The left picture has been taken from one of the cameras at the side of the setting with a frontal view on one of the interlocutors. On the right the perspective of the bird's eye camera showing the interaction space and both object pools has been taken.

The experiment has been recorded on video tape (Mini DV, Sony VX2100) from three perspectives, one corresponding to a bird's eye view (Fig. 3, right) giving a general overview of the setting and the other two from the front left and front right sides providing a uncovered view on gaze and mimics from a single participant (Fig. 3, left). Audio has been recorded separately using two microphones (KM-184), one hanging on the ceiling from above the interaction space and one placed in front of the table.

Alignment of object naming in the Jigsaw Map Game

The Jigsaw Map Game has been applied in a first experiment on the coordination of object naming (Schaffranietz, 2007). The central question of the experiment was on how coordination of object naming takes place under default conditions in an "ideal" face-to-face dialog. For such a situation we propose that coordination of object naming takes place fast and easily in a resource gentle way as put forward by Pickering and Garrod (2004). Furthermore, we were interested in the question how coordination occurs in a situation deviating from such an ideal face-to-face dialog regarding two aspects, i.e., differing *knowledge states* of the communicative partners and *cognitive load* during conversation. These considerations led us to the following design.

Experiment design

Concerning differing *knowledge states* between the communicative partners we were interested in the impact on the use of object references. The variation of the knowledge states was operationalized by priming the naming of all object shapes at the beginning of the experiment by learning. For the set of landmark objects, further referred to as the critical objects, in one condition both participants learned either the first or the second name, in a second condition A learned the first and B the second name (see Fig. 4). For all the other objects they learned one identical name.

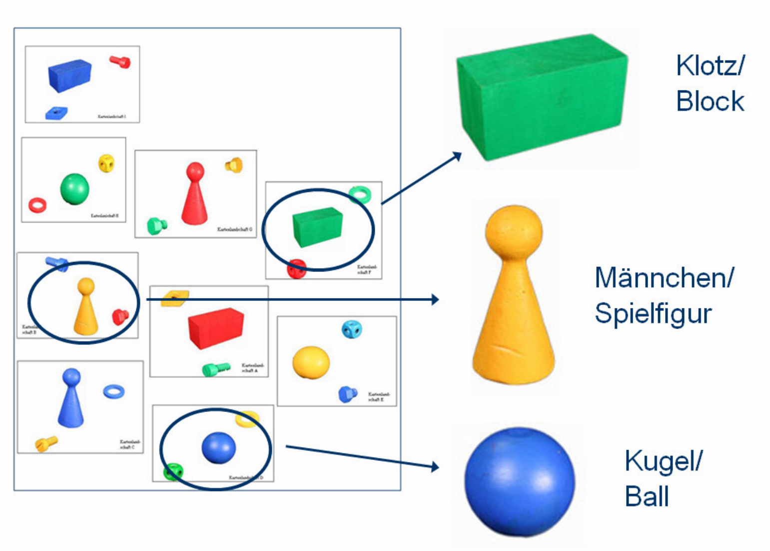


Fig. 4 Certain landmark objects define the critical objects for the object naming task. In one condition the participants learned different names (e.g. Klotz vs. Block) for the objects.

Causing the participants to explicitly learn names for the object shapes we can control the use of lexical references. Varying the learned names for the critical objects in the condition with different knowledge states we can provoke an explicit deviation of the preferred naming between the partners. Furthermore we are able to categorize the object names used during the dialogs into three categories *learned* name, name learned by the *partner* and *idiosyncratic* name, i.e., names not learned by any of the participants. With this procedure we can get to know in a controlled way which naming prevails. In this study, we chose common and therefore highly primed names for all objects so that the critical objects did not attract the attention of the participants solely because of a special or even funny naming. To vary the naming of the critical objects we chose two possible names with comparable plausibility (see Fig. 4). The plausibility to use the chosen names for the objects had been ensured by a small survey conducted with ten students which had been asked to name the objects spontaneously.

Regarding the influence of *cognitive load* on communication in dialog, Horton and Keysar (1996) report, that time pressure reduces the capabilities of speakers to take into account common ground information and behave partner oriented in their language production. On the other hand, results of Rummer et al. (2007) using the same experimental paradigm suggest, that the effects of time pressure on language production in dialog should be qualified by the degree of the interactivity of the communicative situation under consideration. On the basis of these differing results we were interested in the effect of time pressure on communication within the frame of the Jigsaw Map Game.

Cognitive load was varied by confronting one half of the participating pairs with time pressure operationalized by the experimental instruction and the presence of a chess clock, which the participants had to use during the game.

The two factors were varied between experimental pairs as between-subjects factors resulting in a 2x2 experimental design. To each of the four conditions, combining *knowledge state identical/different* and *cognitive load yes/no*, two male and two female pairs were assigned, i.e., 8 naïve participants for each condition. Thus altogether we tested 16 pairs and 32 participants respectively.

Assumptions

Our basic assumption is, that in an ideal dialog situation (represented here by identical knowledge states and no extra cognitive load) communicative partners should coordinate on common object names in an automatic and economic way. Participants should most frequently produce highly primed names for all objects, i.e., the names learned at the beginning of the experiment. Furthermore we expected nearly no explicit negotiations concerning the object references.

Furthermore, we are interested in the effect of differing knowledge states, i.e., differences in the naming of the critical objects. How do participants cope with the induced problems? According to the notion of alignment participants should frequently produce object names introduced by their conversational partner. Especially when referring to a critical object indirectly, in the sense of a spatial reference object, we expect participants to produce names previously used by the partner either for the very same object or for an object of the same shape. This effect may be modified by cognitive load imposed by time pressure.

Procedure

16 pairs of mutually unacquainted interlocutors of the same sex, mostly students, with an average age of 28, were tested in the multi-media laboratory of the Faculty of Linguistics and Literature Studies at Bielefeld University. They were paid for their participation in the experiment. One session took between 30 and 45 minutes.

The participants played the Jigsaw Map Game. Their specific task was to arrange interactively 27 unique objects (4 small bolts, 4 bigger bolts, 4 dies, 4 rings, 3 tokens with a cone shaped body and a sphere as a head, 3 cuboids, 3 spheres and 2 rhombuses) within the interaction space. The objects of the same shape and size were unique in their colors (red, yellow, green and blue). The order of the appearance of the objects on the pieces of the map and the colors of the objects were controlled and varied between sessions.

In the condition with different knowledge states, each participant learned a different name for the shape of the cones ("Männchen" or "Spielfigur"), cuboids ("Klotz" or "Block"), and spheres ("Ball" or "Kugel"). These objects formed the set of critical objects relevant for the analysis (see Fig. 4). The names had been ranked highest in a pretest where ten people had been asked to name the objects used in the experiment. The partners were separated during the instruction. Each was shown a tailored list of all objects and their naming within the game and told to familiarize with the objects.

Subsequently they were given a written instruction with a description of the game in two variants adapted to the condition cognitive load. They were also given time for questions to make sure they did fully understand their task. The instruction emphasized that the participants were requested to refrain from using pointing or grasping gestures and to concentrate on verbal descriptions. The kind of verbal communication had deliberately not been restricted. It has also been stressed, that the placement of the objects has not to be taken too picky as we wanted to reduce lengthy discussions and frustrations.

After the instruction participants were again shown the list of objects and requested to memorize their names. This was then checked in a first test where the object names had to be written down. This was immediately followed by a second test where the physical objects had to be identified verbally. The instruction was completed when the participants correctly remembered all object names. As a consequence the learned names in this experiment were primed strongly.

At the beginning of the game both participants were assigned their seats. Before they started a short example of two interaction turns was given by the experimenter, both to ensure they understood the task and to make a context switch from the object learning task.

The procedure of the game followed the description given above. After the game had been finished, the participants again were asked to complete the questionnaire about the object names. This was done to check, whether they were using the learned names or the ones established during the game. After this, they were asked to fill out a demographic questionnaire. At the end of the session participants again were asked to fill out the questionnaire about the object names, but this time they were instructed to write down the names they would normally have used.

Data preparation

We focus here solely on an annotation of how participants refer to the critical objects within their turns. First, the occurrences of the critical object names were classified according to their target. When a naming referred to the object to be placed, the reference was classified as being *direct*; when the naming referred to the object as reference object, the reference was counted as *indirect* (see Tab. 1). Second, the object references were classified according to the categories *learned* (names previously learned by the participant), *partner* (names previously learned only by the partner) and *idiosyncratic* (names learned by none of the participants).

Tab. 1 Excerpt from the session of pair number 4, the first nine turns of twenty-seven are shown. This game was started by VP8.

Pair 4	**VP7**		**VP8**	
Turn	***direct***	***indirect***	***direct***	***indirect***
1 cuboid			*“Klotz”*	
2		*“Klotz”*		
3				*“Klötzchens”*
4 token	*“Männchen”*	“Klotz”		
5				*“Kegelmännchen”, “Klotz”*
6		*“Männchen”*		
7 token			*“Männchen”*	*“Männchen”*
8		*“Männchen”*		
9				*“Männchen”*
…				

Tab. 1 shows an example how critical object references were counted and classified. In the excerpt of the dialog of pair number 4 VP8 introduces *“Klotz”* for the cuboid in a direct reference in Turn 1. VP7 then aligns to this by picking-up the naming in the indirect reference (at this moment a bolt is to be placed). Interestingly VP8 then alters the naming into *“Klötzchen”*, which is the diminutive form of *“Klotz”*. But this does not affect VP7, who consequently sticks to *“Klotz”* in Turn 4, where the name *“Männchen”* for the newly introduced token is used. VP8 at first does not adopt this name, but is more specific by using *“Kegelmännchen”* (cone-token), still in later referential acts both participants seem to have aligned on using *“Männchen”* for the token.

Results and discussion

Tab. 2 provides an overview of the total number of references to critical objects split along categories and conditions. For a statistical analysis of the frequencies of the critical object references in the categories *learned*, *partner* and *idiosyncratic* we calculated the relative frequencies by counting their proportion relatively to the total number of critical object references produced by each participant. Then these relative frequencies were analyzed separately for each variable with an ANOVA with the two factors *knowledge state* and *time pressure*.

The factor *time pressure* never shows a statistically significant effect. This result is also substantiated by an analysis of the duration of the dialogs, where *time pressure* also had no significant influence. As a consequence we will restrict our report to the factor *knowledge state*.

Though we get statistically significant results concerning the variables frequency of learned and frequency of partner names, with these variables the problem exists that in the condition where both participants had learned identical names we cannot differentiate between learned and partner names on the lexical level. All names except idiosyncratic ones can be counted either as learned or as partner. Therefore these results cannot be interpreted until we find a way to differentiate between the two categories. For the variable frequency of idiosyncratic names the analysis showed no statistically significant effects.

Tab. 2 Number of references to critical objects. Under both conditions, identical learned names and different learned names, a similar amount of references can be observed. The percentages are relative to the total number of references to critical objects observed.

number of references	***direct references***			***indirect references***			**total**
	learned	***partner***	***idiosync.***	***learned***	***partner***	***idiosync.***	
identical names learned	74 (9.8%)	0 (0%)	4 (0.5%)	287 (37.9%)	0 (0%)	22 (2.9%)	**387 (51.1%)**
different names learned	67 (8.8%)	21 (2.8%)	5 (0.7%)	172 (22.7%)	93 (12.3%)	13 (1.7%)	**371 (48.9%)**
total	**141 (18.6%)**	**21 (2.8%)**	**9 (1.2%)**	**459 (60.6%)**	**93 (12.3%)**	**35 (4.6%)**	**758 (100%)**
			171 (22.6%)			**587 (77.4%)**	**758 (100%)**

Landmark use of critical objects

Based on the overall analysis, we were especially interested whether the idea with the critical objects forming the landmarks paid of in our setting. During the course of the study, the 16 pairs of participants had to place 144 critical objects (chances for direct references). They did so using 171 (+18.8%) direct references to critical objects, or 1.19 uses per relevant instruction (see Tab. 2 for a detailed overview). This was expected, as for introducing an object the participant with the piece of the map has to ask the partner to fetch the object out of her personal pool of objects. As the box is out of view for the one asking, the referring noun phrase employs the color and the shape of the object, thus using a direct reference. This we did not count as using a critical object as a landmark. Though, within each session for 26 of the 27 objects reference objects could have been used in the placement instructions, summing up to 416 potential uses of indirect references to critical objects. Effectively, the participants used 587 (+41.1%) references to critical objects. In other words: we found 1.41 uses of critical objects as landmarks per placement instruction. Thus, the strategy that the critical objects form landmarks to elicit a large number of verbal references to those objects is successful.

Direct vs. indirect use of names for object shapes

An analysis of the direct-indirect use with regard to the three categories of references shows that in direct references participants use the names they have learned independently of the fact if they have learned identical (m = .20, sd = .04) or different names (m = .19, sd = .08). If different names have been learned, there is only a small number of uses of partner names in direct references (m = .06, sd = .06). Thus the participants seem to stick to the names they were primed to use during the experimental instruction when introducing new objects.

This result is remarkable, as we expected participants to use more partner names when they had learned different names. This indeed also manifests in the overall data. For the variable frequency of learned names significantly more learned names are used overall, if participants had learned identical (m = .95, sd = .07) than if they had learned different names (m = .65, sd = .26; $F_{1,30} = 16.407$, $p < .001$).

Thus, participants use partner names mainly in indirect references. When they have learned identical names they use significantly more learned names (m = .74, sd = .08) than when they have learned different names (m = .46, sd = .22; $F_{1,30} = 21.357$, $p < .000$). And this effect can directly be explained by a higher amount of partner names used in the condition where different names have been learned when they refer to critical objects indirectly (m = .25, sd = .18) than directly (m = .06, sd = .06). There is no significant difference in the use of idiosyncratic names in both conditions.

With the result on the *direct* vs. *indirect* use of the critical names the idea of alignment is supported. If an object is named directly as target object the focus of attention is concentrated on this critical object that should be placed next. This situation is a direct recall of the instruction situation where the naming of the critical objects had been learned explicitly and therefore this leads to an explicit referring to the learned name. If the object is named indirectly as reference object the focus again is on the next target object that should be placed, whereas the reference object and its learned name are out of focus which may lead to an implicit and automatic take over of the partner name.

Hints on coping behavior

Finally, we were interested in the effect of the condition where we introduced differences in the naming of the critical objects. How do the participants cope with the induced problems? A thorough investigation of this question has to be deferred until all experimental sessions are transcribed and annotated. At this point we can only provide some hints.

First, we examined, if the induced differences in lexical preferences to name the critical objects lead to longer dialog sessions, e.g., because of an increase in negotiations. This is not the case: In the condition with identical names learned the mean session duration is 12.04 min and with different names learned the mean duration is 11.58 min (sd = 2.3).

Second, we analyzed if participants use references to the critical objects more often in the condition with different learned names for the critical objects. Our analysis shows that participants use significantly more direct critical names when they have learned different names (m = .26, sd = .08) than if they have learned identical names (m = .21, sd = .04; $F_{1,30} = 5.885$, $p < .05$). Despite of these differences we found barely any explicit negotiations in a first overview of the video recordings and, if at all, only in later stages of the dialogs.

The participants seem to compensate for the differences by increased redundancy in terms of repetitions when instructing the partner to fetch the object from her box. The opposite effect can be observed for indirect references. Here participants use significantly less indirect critical names when they have learned different names (m = .74, sd = .08) than if they have learned identical names (m = .79, sd = .04; $F_{1,30} = 5.885$, $p < .05$). When differences in the lexical terms used to refer to an object shape are an issue, participants are more careful when using those objects as reference objects.

The absolute numbers of references to critical objects for both conditions, identical and different names learned, are shown in Tab. 2. While there are differences in use depending on direct or indirect references, the total number of reference uses under each condition is comparable. This again confirms that the difficulties induced by different knowledge states do not lead to an increase in negotiations.

Conclusion

All in all, the Jigsaw Map Game has proved as a new capable setting for the investigation of natural interactive language processing in a controlled experimental way. It offers the possibility of real face-to-face interaction with two (or more) communicative partners. The roles of the partners can be varied from an equal to a strongly differing status regarding their knowledge, acting or goals. Furthermore, it allows for manipulations of the nature of the game, e.g., cooperative vs. competitive, and for a wide range of experimental variations concerning the communicative situation (direct face-to-face, mediated, remote, etc.) and the experimental materials. As the order of the turns within the game can be predefined, the games of different pairs of interlocutors can be easily compared. Finally, it offers many features that allow for a detailed linguistic analysis beyond lexical references, such as the analysis of sentence structure or the use of spatial reference systems. Furthermore, it offers the possibility to analyze prosody or gesture and to employ eye tracking techniques to get insights into relevant gaze information.

In the experiment reported we could get basic insights into the way how object references are coordinated in a face-to-face dialog. Particularly the results of direct and indirect object reference uses indicate that coordination of lexical references takes place fast and easily corresponding with the notion of alignment. However, to decide how local these effects of taking over the references of the partner are, a full annotation of the dialogs and more fine-grained data analyses including transition probabilities and distance measures are necessary.

Regarding the condition where both participants learned identical names for the critical objects we cannot tell lexically whether the naming has been learned or adopted from the partner. Here, we do not know so far whether participants use a name because of the strong priming during the experimental learning phase or because they aligned to the partner's use in preceding utterances. This differentiation is necessary not only for better analyzing the data but also on a theoretical level. If an object name is primed by the partner we can interpret it as an automatic take over in the sense of alignment. If a name is used because it is learned, this indicates a more explicit use to introduce a new object in the dialog. One idea to overcome this, is investigating whether there are differences on the phonetic level between interlocutors. If there are, and if alignment happens on the phonetic level, phonetic analyses could produce measures to see, if certain names are pronounced closer the way the interlocutor has pronounced the name before. Thus, it may be possible to conclude whether the name used by the partner has been adopted or not.

Concerning the influence of cognitive load we can draw no conclusions so far, since using the chess clock did not really cause time pressure. This is supported by the analysis of the duration of the dialogs as well as by informal remarks of some participants after their experimental session. Therefore, further experimental work should establish better methods to evoke time pressure and test other ways of inducing cognitive load, e.g., by using dual-task paradigms. This will also contribute to a deeper investigation of the relation between the concept of alignment and the necessity of establishing full common ground in communicative interactions.

References

Anderson, A. H. et al. (1997). The impact of VMC on collaborative problem solving: An analysis of task performance, communicative process, and user satisfaction. In K. E. Finn, A. J. Sellen & S. B. Wilbur (eds.), *Video-mediated communication* (S. 133-155). Mahwah, NJ: Erlbaum.

Anderson, A. H. & Boyle, E. (1994). Forms of introductions in dialogues: their discourse contexts and communicative consequences. *Language and Cognitive Processes*, *9*, 101-122.

Anderson, A., Brown, G., Shillock, R. & Yule, G. (1984). *Teaching talk: Strategies for production and assessment*. Cambridge: Cambridge University Press.

Boyle, E., Anderson, A. H. & Newlands, A. (1994). The effects of visibility on dialogue performance in a co-operative problem solving task. *Language and Speech*, *37*, 1-20.

Branigan, H. P., Pickering, M. J. & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*, B13-25.

Brennan, S. E. & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1482-1493.

Buhl, H. M. (2001). Partner orientation and speaker's knowledge as conflicting parameters in language production. *Journal of Psycholinguistic Research*, *30*, 549-567.

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

Clark, H. H. & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: American Psychological Association.

Clark, H. H. & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*, 62-81.

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

Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S, Garrod, S. & Bruce, V. (1997). Face-to-face and video mediated communication: A comparison of dialogue structure and task performance. *Journal of Experimental Psychology: Applied*, *3*, 105-125.

Garrod, S. (1999). The challenge of dialogue for theories of language processing. In S. Garrod & M. J. Pickering (eds.), *Language processing* (pp. 389-415). Hove: Psychology Press.

Garrod, S. & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, *27*, 181-218.

Garrod, S. & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*, 8-11.

Horton, W. S. & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*, 91-117.

Kranstedt, A., Lücking, A., Pfeiffer, T., Rieser, H. & Wachsmuth, I. (2006). Deixis: How to determine demonstrated objects using a pointing cone. In S. Gibet, N. Courty & J.-F. Kamp (eds.). *Gesture in Human-Computer Interaction and Simulation: 6th International Gesture Workshop, GW 2005, Berder Island, France, May 18-20, 2005, Revised Selected Papers* (pp. 300-311). Berlin: Springer.

Krauss, R. M. & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology*, *4*, 343-346.

Lombard, M. & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, *3*(2). http://jcmc.indiana.edu/vol3/issue2/lombard.html.

Pickering, M. J. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169-226.

Rummer, R., Wintermantel, M., Schweppe, J. & Dirksen, R. (2007). *Wie lässt sich Common-Gound-Etablierung beim monologischen Instruieren verbessern?* Manuscript submitted for publication.

Schober, M. F. & Brennan, S. E. (2003). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. A. Gernsbacher & S. R. Goldmann (eds.), *Handbook of discourse processes* (pp. 123-164). Mahwah, NJ: Erlbaum.

Schaffranietz, G. (2007). *Zwischen Alignment und Common Ground: Ein Experiment zur Koordination von Objektbegriffen im Dialog*. Magisterarbeit. Universität Bielefeld: Fakultät für Linguistik und Literaturwissenschaft.

Weiß, P., Pfeiffer, T., Eikmeyer, H.-J. & Rickheit, G. (2006). Processing instructions. In G. Rickheit & I. Wachsmuth (eds.), *Situated communication* (pp. 31-76). Berlin: Mouton de Gruyter.

Lexical access from spectrum? Phonetic properties of proper names and common nouns in German and Mandarin-Chinese

Henrike Baumotte [1, 2 1], Horst M. Müller [1] & grzegorz Dogil [2]

1) Experimental Neurolinguistics Group, University of Bielefeld, Germany
2) Experimental Phonetics Group, Institute for Natural Language Processing, Universität Stuttgart, Germany

ABSTRACT

This paper describes the spectral analysis of spoken proper names vs. common nouns looking for possible differences between the two word classes on the phonetic level. The idea is that acoustic differences in formants and harmonics might help listeners to identify quickly proper names, used in previous experiments with German and Mandarin Chinese stimuli. Spectral parameters detectable from the harmonic structure (voice quality parameters) were not significantly different between the two lexical classes. A significant difference occurs between German proper names vs. common nouns when taking the classes originally presented to the subjects in the EEG-study as well as lexical decision task experiments, but we had to take into account that there is no vowel [i] in the category of German common nouns. Though there are no significant results after excluding the names containing the vowel [i], it is obvious that proper names differ from common nouns in the F2 which is constantly higher in proper names.

1. Introduction

Spoken word recognition is a highly complex cognitive device in which the detailed knowledge about the sound is usually normalized smoothed and obliterated on the complex route to the discovery of meaning (Culter & Clifton, 1999; Frauenfelder & Floccia, 1999). Numerous models of spoken word recognition do not take into account the phonetic detail encoded in the words to be recognized. They assume that the speech stream is prelexically coded as a normalized, language-specific phonological representation consisting of abstract units of perception such as syllables, phonemes or distinctive features. This assumption is essentially true of all traditional models of spoken word recognition (such as cohort model (Marslen-Wilson & Welsh, 1978), TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994), with notable exception of Klatt's LAFS

[1] Anschrift: Henrike Baumotte, Institute for Natural Language Processing, Universität Stuttgart, Azenbergstr. 12, 70174 Stuttgart, email: henrike.baumotte@ims.uni-stuttgart.de. This study was supported by the German Science Foundation (SFB 360).

model (Klatt 1979)). However, using abstract information as access code not necessarily excludes that other aspects of the acoustic input are used for different purposes.

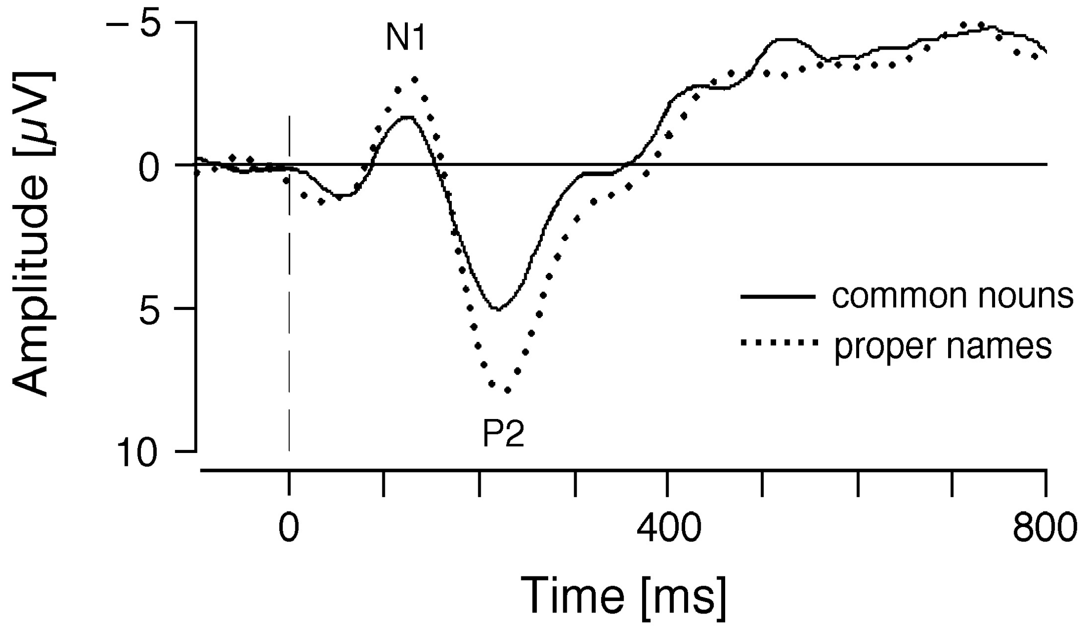
In the Lexical Access From Spectra (LAFS) model it is stipulated that word recognition is accomplished directly on the basis of spectral representations of individual words. Spectral representations of input words are compared to the prototypes stored in memory, with no intermediate levels of computation corresponding to segments, phonemes or syllables.

More recently, Goldinger and others (Lacerda, 1995; Goldinger, 1996; Pierrehumbert, 2001; Johnson, 2005) demonstrated that lexical access is direct in most circumstances. All exemplar-based models of perception stress the role of specific phonetic features, in particular those which are coded in the spectral representation of speech. According to Bybee (2005) learned phonetic detail may be associated with languages and dialects, but also with specific words in the lexicon of a given dialect, for example schwa- or t/d-deletion. Each category is represented in memory by a large cloud of remembered tokens whereas memories of highly similar instances are close to each other and memories of dissimilar instances are far apart.

Regarding the inner structure of the lexicon, there is long tradition concerning an exceptional role of proper names in language philosophy and linguistics. Experimental and patient data also underline the special status of proper names. Neurolinguistic findings give evidence for German and English spoken proper names being special in a way that listeners may be able to detect discriminatory signals in the sound patterns of spoken proper names (Nomina propria) in comparison to spoken common nouns (Nomina appellativa). Relatively few objects carry a “proper name”, and these play an important role in people’s lives. In everyday communication, they compose an essential linguistic instrument. They also exist in all natural languages and help speakers to exchange information about absent subjects or objects. Different disciplines dealing with proper names have found evidence of their exceptional position: language philosophy, traditional philology, neuropsychology (Cohen & Burke, 1993; Valentin, Brennen & Brédart, 1996), psycholinguistics (Carrol, 1985), neurolinguistics (Müller & Kutas, 1996), sociology and socioanthropology.

Searching for processing differences between English proper names and common nouns Müller and Kutas (1996) found in an EEG experiment differences in the very early stages of processing of both word classes. Investigating proper names and common nouns in natural English speech elicited ERP differences could be identified as early as 120 ms after word onset. It is reasonable to believe that the information detected by the brain so early in the processing is somehow coded in the spectra of the words belonging to the critical word classes.

Figure 1: Grand average ERPs (n=32) at the vertex (Cz) elicited by 22 different English sentences, starting with a proper name (dotted line), and 22 with a common noun (solid line) (see Müller & Kutas, 1996, Fig. 1).



In a German lexical decision experiment (Werner & Müller, 2001) the subjects could discriminate between the two critical word classes even after the first 100 ms of articulation in the case of natural speech. Apparently, the discrimination of computer-synthesized speech, which served as a control condition, was more difficult for subjects. This suggests that it is the spectral properties of natural speech that probably licence access to different semantic classes in German. A lexical decision task experiment in Mandarin-Chinese did not show a significant difference between the proper names and common nouns used (Yen & Müller, 2003). In conclusion, English and German listeners are able to dissociate significantly the first 120 ms of the tested proper names from common nouns. This is obviously an intuitive process, because the subjects at first questioned the feasibility of this task.

The question on which acoustic features this decision performance relies upon offers itself to investigation. Until now, the perception differences have been exclusively shown in the two tested germanic languages. Therefore, we used the German (Müller & Kutas, 1996) and Mandarin-Chinese (Yen & Müller, 2003) stimuli of the previous experiments to carry out phonetic analyses. This paper presents these phonetic measurements and works out which features in the stimuli presented during the earlier experiments might enable the subject to distinguish between the semantic classes. An attempt is made to find reasons for the differences between proper names and common nouns in the feature found.

2. Phonetic analyses

A phonetic analysis of spectral structure has been performed to pin down possible physical differences between proper names and common nouns. Both formants and harmonic structure of the spectra have been subject to a phonetic investigation.

2.1. Material

81 natural language German stimuli spoken by a male speaker (41 proper names, 40 common nouns) from the experiment of Werner and Müller (2001) as well as 79 natural Mandarin-Chinese stimuli spoken by a female Chinese speaker (40 proper names, 26 common nouns) of the experiment of Yen and Müller (2003) were analyzed. Only disyllabic words were selected for the study. Words containing a diphthong in the first syllable were disregarded, so that finally 39 German proper names, 32 German common nouns and 26 Chinese proper names and the same number of common nouns remained. Mean duration for the German stimuli were 383.5 ms (names) and 423.2 ms (nouns). In both cases words were matched for psycholinguistic criteria (word frequency, initial phoneme etc.).

2.2. Method

For the analysis of the voice quality the differences of the amplitudes of the first two harmonics were manually extracted from the spectra. The relation between the two first harmonics is a good indication of the type of glottal source (modal, breathy, creaky) of sound (Stevens & Hanson, 1995; Claaßen et al., 1998) and displays the voice quality characteristics. Mean corrected amplitude differences were compared using chi square-tests.

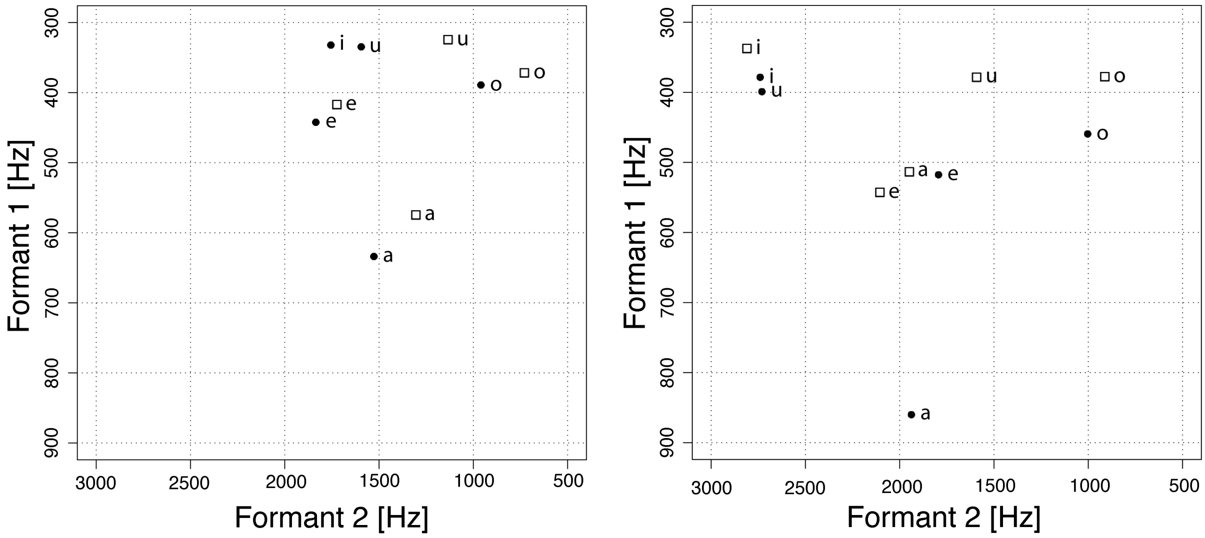
Resonance properties are immanent in the vocal tract and depend on the different vocal tract shapes during sound productions. For the analysis of the resonance properties the values of the first two formants were extracted from the first vowel of the critical words. Formants were measured manually using standard software (WaveSurfer 1.6.0). The formants were elicited from the LPC spectra taken from the middle portion of the vowel.

Mean frequency value height was analysed using *t*-tests for double sided, independent control samples. The standard deviations were obtained using the descriptive analysis-tool of the software SPSS. For these analyses the vowel [i] for the proper names in the group of German stimuli was skipped out to assure parallelism of the two groups.

3. Results

Being conscious of the fact that vowels are articulated differently depending on the position of the tongue and the rounding of the lips, vowels are grouped in categories and the mean value is calculated. The measurements of voice quality did not deliver any significant differences between the phonetic realizations of the proper names and common nouns. Neither German speaker nor the Chinese speaker differentiate the two word classes by producing them with a measurably different voice quality. The measurements of vowel quality were performed individually for each vowel type. Vowels are grouped in categories and the mean formant values for each category is calculated. Fig. 2 shows the results of the formant measurements.

Figure 2: The position of the first and second formants in proper names (dottes) and common nouns (blanked squares) for German (left) and Mandarin-Chinese (right) shown by the mean frequency value type.



On the left side of the Fig. 2 a systematic difference between the values for proper names and common nouns can be detected. The difference values for the F_1 and F_2 have been confirmed statistically with *t*-tests (table 1). The average F_1 and F_2 formant frequency values of proper names were compared to that of common nouns. Finally, there is no significant difference between the first as well as the second formant of proper names vs. common nouns in German. It is however obvious that the F_2 frequency values for proper names are generally higher in comparison to those for common nouns.

In accordance with these results, there is also no significant difference for any of the formants' mean values of Mandarin-Chinese stimuli.

Table 1: *t*-test-results after comparison of German and Mandarin-Chinese proper names with the conformable common nouns.

parameter	*t*=	significance
F_1_German	0.460	$p = 0.734$
F_2_German	3,292	$p = 0.001$
F_2_German (excluded names with vowel [i])	1.431	$p = 0.157$
F_1_Chinese	1.530	$p = 0.132$
F_2_Chinese	0.832	$p = 0.410$

4. Discussion

The measurements demonstrated that acoustic parameters might not play an important role within the differentiation of German proper names and German common nouns, i.e. the proper names and common nouns do not differ acoustically. However, in figure 2 a differentiation of German proper names and common nouns with respect to articulation is detectable. The second formant of the proper names is always higher than that of the common nouns. The high value for F_2 might show that the speaker marked proper names by using a larger constriction in the middle and frontal articulatory space than the one he used for common nouns. Furthermore, the degree of lip rounding and the lip protrusion could be fractionally smaller for proper names than for common nouns. However, these minute articulatory/acoustic differences are very difficult to relate to the semantic information which is coded within the first 100/160 ms of articulation and in the formants of the first vowel. The neurophysiological and psycholinguistical findings described by Müller and Kutas (1996, 1997), Schuth and Müller (2001), Werner and Müller (2001) as well as Yen and Müller (2003) suggest that proper names and common nouns differ not only in their processing, but also in their production. The two categories can build up by the subjects because the way stimuli are articulated is slightly different which at the same time leads to the probably perceived frequency differences.

Presumably, German stimuli differ from Mandarin-Chinese stimuli with regard to their syllable and phoneme structure, their segmental structure as well as their frequency distribution of triphons which is more complex in German. In contrast to German, which is a stress language, Chinese is classified as a tone language. In stress languages such as German, English and Dutch a variation of the position of stressed syllable can be found across words (for example: UMfahren, umFAHren), while in tone languages there is only a restricted amount of syllables connected with a large tone repertoire (e.g., four tones for Mandarin) (Dogil, 1999). That is why Mandarin-Chinese as a tone language possesses a smaller amount of distinct syllables than German. In Mandarin-Chinese the four tones are used to differentiate between words' meanings. In tone languages the prosodic channel is used for semantic differentiation and the phonological information is not available. Proper names and common nouns in Mandarin Chinese do not differ in their phonological structure which is however the case in German stimuli. Xu (1999) states that in non-tone languages, F_0 peaks come along with pitch accents or stress, while Mandarin and other Asian languages follow a gliding F_0 contour characteristic. According to House (1999) four contour tones (H, R, L and F) exist in Mandarin-Chinese which may result in more complex perceptual mechanisms containing higher order cognitive processing and short-term memory.

Whatever the reason, there was no acoustically measurable difference between the stimuli distinguishing the two word classes in Chinese. Note, that Yen and Müller (2003) who used the stimuli for the reaction time experiment could not detect any significant differences either. The lexical access seems to be critically dependent on the unfiltered spectral information.

In view of acoustic it might be claimed that the differences in articulation produce acoustic differences, which the subjects might be able to detect while hearing the German stimuli used. This raises the question why these relatively minor differences, which concern only the second formant, can be perceived as such. This remaining question leads to the subject of perceptible phonetics. The subjects of the gating study processed the stimuli differently on an acoustic-perceptible level. Goldstein (2002: 396, 398) confirms in his theory of local and temporal coding a highly differentiated perception and processing of frequencies.

It should be stressed that proper names (personal names) and common nouns in German differ not only in their processing, but also in their production. Müller and Kutas (1997) hypothesize that the classification in proper names and common nouns does not only constitute two theoretical classes, but also natural and cognitively detectable ones. Presumably, the performance of subjects can be reduced to the fact that German proper names contain recurring phoneme sequences, most often as suffixes. However, not only phonemic shapes, but also onomatopoetic and suprasegmental prosodic features can be determining for the word class decision. It might be that because of the accentuation being on the first syllable of the proper names the frequencies of the first vowel are higher than in common nouns.

Finally, further studies are necessary to analyze a more balanced data base, more speakers' data, different languages and to compare those with our previous results obtained from English, German as well as Chinese stimuli. For those studies, great care should be taken of different speech typologies as well as acoustic differences occuring from anatomical vocal tract differences. Apart from that, it would be interesting to find out if a manipulation of the second formant of the first vowel of German proper names or common nouns in German stimuli could influence the reaction of the subjects. If, for example, the second formant of proper names is lowered artificially, subjects might conclude that the stimulus presented was a N. appellativum and not a N. proprium.

REFERENCES

Bybee, J. (2005). *From usage to grammar: The mind's response to repetition.* LSA Presidential Address.

Carol, J. M. (1985). *What's in a name: An Essay on the psychology of reference.* New York: Freemann.

Claßen, K., Dogil, G., Jessen, M., Marasek, K. & Wokurek, W. (1998). Stimmqualität und Wortbetonung im Deutschen. *Linguistische Berichte*, Vol. 174, pp. 202-245.

Clark, J. & Yallop, C. (1995). *An introduction to phonetics and phonology.* 2 ed. Oxford: Blackwell Publishers Ltd.

Cohen, G. & Burke, D. M. (1993). Memory for proper names: A review. In: Cohen, G. & Burke, D. M. (eds). *Memory for proper names.* Hove: Erlbaum, pp. 249-263.

Culter, A. & Clifton, C. J. (1999). Comprehending spoken language: A blueprint of the listener. In: Brown, C. M. & Hagoort, P. (eds). The Neurocognition of Language. New York: Oxford University Press, pp. 123-166.

Dogil, G. (1999). The phonetic manifestation of word stress. In: van der Hulst, H. (ed). *Word prosodic systems in languages of Europe.* New York: De Gruyter, pp. 273-334.

Frauenfelder, U. H. & Floccia, C. (1999). The recognition of spoken words. In: Friederici, A. D. (ed). *Language Comprehension: A biological perspective.* Heidelberg: Springer-Verlag, pp. 1-40.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 22. 1166-1183.

Goldstein, E. B. (2002): *Wahrnehmungspsychologie.* 2. Aufl. Heidelberg. Spektrum Akademischer Verlag.

House, D. (1996): Differential perception of tonal contours through the syllable. *Proc. ICSLP, Fourth International Conference of Spoken Language Processing*, pp. 2048-2051.

House, D. (1999). Perception of pitch and tonal timing: implications for mechanisms of tonogenesis. ICPhS, San Francisco, pp. 1823-1826.

Johnson, K. (2005). Speaker normalization. In: Remez, R. & Pisoni, D. B. (eds). *The Handbook of Speech Perception.*

Klatt, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, 7: 279-312.

Lacerda, F. (1995). The perceptual-magnet effect: An emergent consequence of exemplar-based phonetic memory. In: *Proceedings of the 13th International Congress of Phonetic Sciences (Stockholm)*, Vol. 2, pp. 140-147.

Marslen-Wilson, W. D. & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10: 29-63.

McClelland, J. L. & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18: 1-86.

Müller, H. M. & Kutas, M. (1996). What's in a name? Electrophysiological differences between spoken nouns, proper names, and one's own name. *NeuroReport*, 8: 221-225.

Müller, H. M. & Kutas, M. (1997). *Die Verarbeitung von Eigennamen und Gattungsbezeichnungen. Eine elektrophysiologische Studie.* In: Rickheit, G. (Hrsg.). Studien zur Klinischen Linguistik: Modelle, Methoden, Intervention. Opladen: Westdeutscher Verlag, pp. 147-169.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52: 189-234.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In: Bybee, J. & Hopper, P. (eds). *Frequency and the Emergence of Linguistic Structure*. Amsterdam: Benjamins, pp. 137-157.

Schuth, A. & Müller, H. M. (2001). How fast is spoken proper name perception? Electrodermal responses in a gating study. In: Arnold, T. & Herrmann, C. S. (eds). *Cognitive Systems & Mechanisms: Abstracts of the 5th Meeting of the German Cognitive Science Society.* Leipzig: Leipziger Universitätsverlag, p. 87.

Stevens, K. N. & Hanson, H. M. (1995). Classification of glottal vibration from acoustic measurements. In: Fujimura, O. & Hirano, M. (eds). *Vocal fold physiology: Voice quality control.* San Diego: Singular, pp. 147-170.

Valentine, T., Brennen, T. & Brédart, S. (1996). *The cognitive psychology of proper names.* New York: Routledge.

Werner, R. & Müller, H. M. (2001). Processing differences between proper names and common nouns: Psycholinguistic evidence. In: Arnold, T. & Herrmann, C. S. (eds). *Cognitive Systems & Mechanisms: Abstracts of the 5th Meeting of the German Cognitive Science Society.* Leipzig: Leipziger Universitätsverlag, p. 89.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25: 61-83.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F_0 contours. *Journal of Phonetics*, 27: 55-105.

Yen, H.-L. & Müller, H. M. (2003). Processing of proper names in Mandarin Chinese. In: Schmalhofer, F., Young, R. M. & Katz, G. (eds). *Proceedings of the EuroCogSci 03.* London: Lawrence Erlbaum, p. 450.

State of the art of phonetic language aptitude linking phonetic as well as phonological models to empirical neuroimaging (neurolinguistic) research

HENRIKE BAUMOTTE [1] & GRZEGORZ DOGIL

Experimental Phonetics Group, Institute for Natural Language Processing, University of Stuttgart, Germany

ABSTRACT

Previous studies have pointed out differences in people's aptitude, capacity, success and speed of foreign language acquisition. The research project "Language talent and brain activity" at the Universities of Stuttgart and Tübingen uses a combination of phonetic and behavioural experiments with neuro-imaging techniques to be able to show differences between talented and untalented second language learners with regard to phonetic abilities. The assumption is that talented learners exhibit less of a foreign accent than untalented learners.

The first author's PhD thesis, embedded in this project, attempts to detect individual differences in the L2 speakers' production and perception skills.

We assume coarticulatory resistance to be one of the reasons for foreign accent, and demonstrate that "talented" learners of English as L2 develop strategies of breaking the language specific coarticulatory resistance.

1. Introduction

In the past many individual differences in people's aptitude, capacity, success and speed of foreign language acquisition have been documented (e.g., Saville-Troike, 2005). The research project supported by the DFG, "Language talent and brain activity" at the Universities of Stuttgart and Tübingen, attempts to uncover the differences between talented and untalented second language learners regarding phonetic abilities with the help of behavioural (linguistic and psychological) as well as neuro-imaging techniques. We consider talented learners to be those showing less of a foreign accent, whereas untalented learners exhibit a strong accent.

Speech production is obviously a very complex process demanding the activation of many muscles at the same time (Clark, Yallop & Fletcher, 2007: 11 ff). Coarticulation is described as the context-

[1] Anschrift: Henrike Baumotte, Institute for Natural Language Processing, Universität Stuttgart, Azenbergstr. 12, 70174 Stuttgart, email: henrike.baumotte@ims.uni-stuttgart.de. The authors would like to thank the Institute for Natural Language Processing, Universität Stuttgart, Germany, and especially the DFG project "Language talent and brain activity" for their support and advice.

dependent overlap of articulatory gestures (e.g., Clark, Yallop & Fletcher, 2007: 86) interacting in the production of successive phonetic segments. As a result there is an influence of vocal tract configuration at any point in time on more than one segment.

> *"Experimental phonetics studies coarticulation as a way of finding out how the brain controls the production of speech." (Roach, 1992: 22)*

Coarticulation forms part of phonological processes which are of varying types and lead to allophonic variations. In turn phonological processes vary in view of different languages. Ashby and Maidment (2005: 132) define coarticulation to be the influence of vocal tract configuration at any point in time on more than one segment. Farnetani (1997: 376) assumes that coarticulation is a universal phenomenon because it has been found in all languages analyzed. The examination of coarticulation is possible with the help of different techniques such as spectrogram analysis (Ashby & Maidment, 2005: 129) which shows changes in acoustic characteristics.

This work will be situated within the Coarticulatory Resistance (CR) theoretical framework. "Coarticulatory Resistance" can be defined as the degree of variation or similarity of the same speech sound across contexts (Bladon & Al-Bamerni, 1976: 137). In contrast a high ability to break and loose the language specific coarticulation/coarticulatory resistance leads to less of a foreign accent. In their Degree of Articulatory Constraint (DAC) model of lingual coarticulation Recasens et al. (1997) and Recasens (2004) have included quantification in a theory of coarticulation developing coefficients of CR. In the thesis the first author will try to detect individual differences in the L2 speakers' production and perception skills. In particular, we assume that coarticulatory resistance is a basic source of foreign accent. We intend to show that "talented" learners of L2 English develop strategies of breaking and loosening the language-specific coarticulatory resistance. The neuro-functional basis of this "breaking" or "loosening" procedure will be investigated. With the help of the below-mentioned tests it will be possible to categorize subjects as "proficient" as well as "unproficient" L2 learners.

Current and future experiments involve female and male native speakers of German. All speakers have an academic background and have learned English mainly at school since they were approximately 10 years old, preferably including a stay in an English-speaking environment. Phonetic experiments of production and perception skills are conducted with the help of various elicitation techniques: spontaneous speech to get an overall impression, to be able to analyze general pronunciation ability and prosody, read speech controlling coverage of the phoneme inventory and specific phonotactic constellations, pitch accent distribution and tunes, direct imitation (imitation of a model) to test constellations which are traditionally difficult to pronunciate (allophones) and suprasegmental, complex tonal constellations and delayed imitation (delayed imitation to ensure the actual linguistic competence level) in order to detect complex tonal constellations and to investigate subtle phonetic variations on the segmental and suprasegmental level. The test materials were created at the Institute of Natural Language Processing, Universität Stuttgart, Germany.

The categorization of proficient and unproficient speakers of L2 English will also consider perception and psychological experiments. Perception experiments involve comparisons of speech melody including low-pass filtered stimuli, accent identification, identification of speaker intentions and emotional colouring as well as listening comprehension tests.

The psychological experiments include tests of phonological working memory, empathy, personality (e.g., learning style and motivation), verbal as well as non-verbal intelligence (Raven-Test), musicality (Gordon-Test) and mental flexibility tests.

Conducting this phonetic and psychological test battery, the subjects which are very proficient and those considered to be unproficient will be invited to the Universität Tübingen to participate in a brain imaging study. This study will contain experiments examining the structure (brain anatomy based on magnetic resonance (MR) imaging, white and grey matter density measurements (VBM) and white matter fiber tracking (DTI) and function, (fMRI)) of subjects' brains while performing phonetic tests similar to those described above.

Coarticulatory resistance might be one of several reasons for the subjects not to be able to overcome their foreign accent. Coarticulatory resistance has been a matter of interest in many studies, e.g. French (Benguerel, Hirose, Sawashima & Ushijima, 1977a, 1977b), English (Lehiste, 1964; Bladon & Al-Bamerni, 1976; Bladon & Nolan, 1977; Majewski, Rothman & Hollien, 1977), Catalan (Recasens, 1984a, 1984b; Recasens, Fontdevila & Pallarès, 1995; Recasens & Pallarès, 2001), German (Recasens, Fontdevila & Pallarès, 1995) and Polish (Majewski, Rothman & Hollien, 1977). These findings support the assumption that coarticulation is a universal phenomenon (Farnetani, 1997: 376). However, there are differences involving phoneme sequences that occur in one language but not in the other. These investigations often compare the degree of coarticulatory resistance (CR) in two different languages looking at the same phoneme sequences or taking into account the CR within one language analyzing vowel-to-consonant(-vowel) coarticulation.

Recasens, Fontdevila and Pallarès (1995) reported for German the value for F_2 in [l] to be lower overall than in other languages. The tongue dorsum is more constrainted for German non-velarized [l] and thus less sensitive to coarticulatory effects from, e.g., [i] or [a]. Authors compared German with Catalan [ɫ] production and observed greater dorsal contact at the palatal zone for German [l] than for Catalan [ɫ]. In line with the surrounding formant frequencies for the vowel [a], consonantal effects on F_2 for [ə] are also large because no defined vocal-tract shape is necessary for the production of [ə]; this is why schwa is highly sensitive to coarticulation (Recasens, 1985). There is considerable evidence for differences in tongue dorsum configuration. The degree of variability of the surrounding formant frequency values usually decreases in the progression [n]>[l]>[d]>[t]>[s] (German: Hoole, Gfroerer & Tillmann, 1990; Kühnert et al., 1991).

Bladon and Al-Bamerni (1976) found that English [l] also includes the dichotomy of the velarized ("dark") or non-velarized ("clear") types of the consonant which agrees with the "articulatory syllable" theory of R-L coarticulation. They (1976: 146) also synchronously state their evidence of L-R coarticulation also appearing constricted by the same notion of an articulatory syllable. In this case their data do not particularly confirm the existence of different mechanisms for each coarticulatory direction (Bladon & Al-Bamerni, 1976: 148). For RP British English consonants are described to extend beyond larger regions of the vocal tract due to tongue-dorsum activity (e.g., velarized apicoalveolar, bilabiodorsoalveolar, dorsal consonants [alveolo-palatals, palatals]), which are highly resistant to vowel coarticulation. During the articulation of alveolars the tongue dorsum is not directly involved, and therefore it is free to coarticulate with the neighbouring vowels. In the case of palatals the tongue dorsum makes the vocal tract narrower, which inhibits possible coarticulatory gestures with surrounding vowels (Recasens, 1984a: 62). Bladon and Nolan (1977) found the lowest degree of coarticulatory resistance for English nasals. According to Lubker (1968) an [a] followed by an [m] was characterized by lower activity of the levator palatini (correlated with a lower velar height) than an [a] uttered in isolation. Rochet and Rochet (1991) stated that high vowels exhibit more assimilating nasality than low vowels, which gives evidence for a prevailing anticipatory effect in English. With regard to dentoalveolar consonants, laminal fricatives ([s], [z], [ʃ] and [ʒ]) appear to be more resistant than apicals ([t], [d], [n] and [l]): while laminals do not usually allow tongue tip raising effects from adjacent apical consonants, apicals tend to become laminal when adjacent to laminals (British English, Bladon & Nolan, 1977).

Until now no analysis of German subjects speaking German and additionally speaking L2 English has been conducted. In my research it is my special objective to search for coarticulation differences between German proficient vs. less proficient L2 English speakers taking into account the results of the phonetic experiments of the "Language talent and brain activity"-project.

In my PhD thesis we assume that the lack of the ability to break and loose the specific mother tongue coarticulation/coarticulatory resistance leads to a strong foreign accent.

Ziegler and von Cramon (1985, 1986), Ziegler (1989) as well as Vollmer (1993, 1997) reported differences in temporal anticipation of articulatory configurations between normal speakers and speakers suffering from apraxia and dysarthria. In apraxic speech there is a problem of phasing speech gestures appropriately which is characterized through a delay in tongue body adjustment, specifically in roundedness within the first two segments (Ziegler & von Cramon, 1986).

> *"[...] [t]he problem of phasing speech gestures appropriately is an essential constituent of apraxic speech."* (Ziegler & von Cramon, 1986: 45)

The significance of interarticulatory phasing in coarticulation is a central concept in speech production theory. Apraxic as well as dysarthric subjects prolonged the sequences considerably longer in all test words in comparison to normal speakers (Ziegler & von Cramon, 1985: 122) shown by increased durations of S_1 (a frame of 25.6 msec in the center of the preconsonantal schwa) and S_2 (a 12.8-msec frame positioned to the burst onset of the alveolar plosive). Motor speech impairments in apraxic patients can be explained by the different biomechanical properties of speech gestures which lead to articulatory, acoustic and categorical changes in the perception of speech segments. Dysarthric patients reduced the acoustic distances of the vowels dramatically which ended up in a large centralization tendency.

Speech sound variations have not only been shown in pathological speech but also depending an sex, gender and social class. For many researchers the examination of acoustic differences in the production of [s] has been a field of interest. It has been claimed that there is a certain variability in the [s]-sound depending on sex (e.g., Stevens, 1998: 398; biological explanation), gender (e.g., Strand, 1999; gender indexical explanation) as well as social class (e.g., Stuart-Smith in press).

Based on these explanations and searching for direct morphological or articulatory data, Fuchs and Toda (2007) conducted an electropalatographic (EPG) approach searching for evidence of inter-language differences in the production of [s] in German vs. English. The authors have taken the corpus built up by Brunner et al. (acc.) used to investigate the influence of the palate shape on token-to-token variability. Experiment materials consisted of /’sasa/-sequences for the German subjects and /’zasa/-sequences for the English subjects, in which the target sibilants were situated medially in an ambisyllabic post-stressed position. Fuchs and Toda (2007) found significant differences between palatal parameters of English vs. German speakers, whereas the results of Stevens (1998) and Strand (1999) could not be replicated. According to their results, German speakers realize a wider constriction than English speakers.

Looking at clusters of voiceless consonants has been one of the most common ways to investigate processes of coarticulation or coproduction at the laryngeal level (Hoole, 1999: 115). Gobl and Ní Chasaide (1999: 125) reported differences in the realization of stops between German and English, while for other phonemes those differences are less clear-cut. The VOT of English phonologically voiced stops is slightly longer in stop-sonorant sequences, such as [bl], than in the singleton case (Hoole, 1999: 111). Docherty (1992) further investigated voicing coarticulation. He considered devoicing of [l] in English words (e.g., “plead”) because of the adjacent voiceless [p]. Following Docherty (1992) place of articulation has a significant effect on VOT; [p] has a shorter VOT than [t] or [k]. This is a consequence of peak glottal opening, which is timed earlier with respect to release for [p] than for the other plosives. In general, [p] has longer occlusion duration than the other stops. Docherty (1992) concluded from this data the realization of stops is not the same. Apparently, a difference in the mode of phonation in these segments influences the following voiced segment (vowel) (Gobl & Ní Chasaide, 1999: 300; Docherty, 1992). Gobl and Ní Chasaide (1999: 141) observe that the initiation and ending of voice during unocclusion of the vowel tract are not identical (see typical pattern of laryngeal-oral coordination in Hoole 1999: 110). The authors present a technique for analyzing the voice source (Gobl & Ní Chasaide, 1999: 300 ff) making use of different parameters which allow detailed voice source measurements. Within this technique the source signal is quantified while using parameterization of the glottal waveform based on a model of differentiated glottal flow. Thereafter, the most relevant acoustically and perceptually voice source parameters (EE, RA, RK, RG, L0, L1 [see below]) are taken to qualify the observations. Some of those are thus described in more detail because they enable us to illustrate the differences between German vs. English stop segments. According to Gobl and Ní Chasaide (1999: 317) EE, the excitation strength, describes the negative amplitude when maximum discontinuity of the derived flow is reached. In consideration of speech production this characteristic parameter displaces the speed of closure with which the vocal folds vibrate and the velocity amount of air going through them. Acoustically, EE gives a description of the overall intensity of the resulting signal. RA measures the residual air flow (or dynamic leakage) coming from the excitation and ending up in complete closure. Taking into account the production level, it depends on the way in which the vocal folds work together, e.g. in a more instantaneous or gradual way with respect to their length as well as depth. Also RK indicates the symmetry of rising and falling branches of the glottal flow pulse. In the case of a large RK value the skew of glottal air flow pulses is higher.

Acoustically, RK is mostlyanalys important regarding the lower part of the source spectrum, because a high RK increases the lower harmonics (see Gobl & Ní Chasaide, 1999: 318, Fig. 15.9). At the same time the amplitude level of F0, amplitude of the first harmonic, as well as L1 (the amplitude level of F_1) are computed (Gobl & Ní Chasaide, 1999: 124, 317).

In German, at the onset of the vowel followed by [p^h] the excitation, EE, is very weak, while dynamic leakage, RA, is very high and the shape of the glottal pulse is very symmetrical, RK (Gobl & Ní Chasaide, 1999: 136). For German stimuli a gradually rising L0 has been detected which attains constant values only after 30 ms or so. Following [p^h] L1 is initially weak. In contrast an abrupt initiation of voice following [p^h] with strong excitation, EE, has been found and not always an involvement of the weak, very gradual onset in English which is typical for the German data (Gobl & Ní Chasaide, 1999: 137 f.). In English, glottal adduction is incomplete, therefore RA is supposed to be low. In correspondence to the German data, the symmetry of the vocal folds (RK) is also high as for the English stimuli. For this reason no difference in RK with regard to German vs. English can be concluded. In English L0 stays rather constant before the beginning of the last glottal pulse. Finally, L1 amplitudes are less high in English as compared to German data.

1.1. Exemplar theory

A final objective is to test phonetic models (e.g. Exemplar Theory), to reveal Foreign Language Pronunciation Talent and to look for psycho- and neuro-linguistic evidence which accounts for individual differences in phonetic talent during second language acquisition. The Exemplar Theory developed by Pierrehumbert (2001) serves as a psychological model of similarity and classification, perception and categorization (especially vowel categorization), a model of phonological experiences. A label which has more numerous or more activated exemplars in the neighbourhood of a new token is used more often. Pierrehumbert (2001) points out that stimuli are also rated according to their informative content and compared with models stored in the human brain. A considerable body of evidence has been accumulated that speakers have detailed phonetic knowledge of a type which is not readily modelled using the categories and categorical rules of phonological theory. There are systematic differences between languages in the fine details of pronunciation (exact phonetic targets and patterns of variation) which have to be learned in the course of language acquisition. According to Bybee (2005) learned phonetic detail may be associated with languages and dialects, but also with specific words in the lexicon of a given dialect, for example schwa- or t/d-deletion. Each category is represented in memory by a large cloud of remembered tokens whereas memories of highly similar instances are close to each other and memories of dissimilar instances are far apart. The volume of speech which a person processes in a lifetime is great, but not every word is finally stored. The exemplars encoding frequent recent experiences have higher resting activation levels than exemplars encoding infrequent and temporally remote experiences. When a new token is encountered, it is classified in exemplar theory according to its similarity, i.e. its distance in the parameter space to the exemplars already stored. During this process a label which has more numerous or more activated exemplars in the neighbourhood of the new token has an advantage. Exemplar theory provides us with a way to formalize the detailed phonetic knowledge that native speakers have about the categories of their language while giving a picture of the "implicit phonetic knowledge of the speaker". The acquisition of this knowledge can be understood simply in terms of the acquisition of a large number of memory traces of experiences. The assumption that people learn phonological categories by remembering many labelled tokens of these categories explains the ability to learn fine phonetic patterns of a language. It is the special objective of this PhD thesis to take into consideration Exemplar theory to explain the individual differences between L2 English aptitude of our subjects which participated in the "Language talent and brain activity"-project. Subjects categorized as untalented might not automatically be able to enlarge their exemplar clouds after having heard a sound which is not identical with those existing in their mother tongue. In turn, during L2 production not as many exemplars as in talented speakers may be activated. In the future work we would like to further unravel whether sounds which have not been stored in untalented learners make up a homogeneous group and if, can they be distinguished in their manner and degree of coarticulation/coarticulatory resistance. In which way does the manner and degree of coarticulation/coarticulatory resistance between German compared to L2 English language differ? Within this framework it might also be important to account for inter-language (Docherty, 1992; Gobl & Ní Chasaide 1999; Hoole, 1999), dialectal as well as gender and sex differences (Stevens, 1998; Strand, 1999) that might lead to speaker or speaker group specific allophonic variations (Recasens, 1999; Baumotte et al., 2007 in press). After the pilot study of Baumotte et al. (2007 in press) we expected that speaker-specific coarticulatory resistance patterns and even inter-speaker variability of coarticulation will be found. Following the above described Exemplar Theory it is not only necessary to take into consideration the production of German native speakers in their L2 English speech but also their perception (s. below further research to be done).

2. Material

In order to investigate the ability to overcome language specific coarticulatory resistance, 40 German (20 proficient as well as 20 less proficient subjects) which have been tested within the project experiments will be recorded. It is unavoidable within this research question to examine individual aspects of coarticulation/coarticulatory resistance through controlled experiments with manipulated material. Each stimulus "/gəCVtə/", i.e. each stimulus embedded in a carrier sentence, will be spoken five times to ensure a representative average value for the within and between group(s) analysis as well as to be able to observe individual variations. These test utterances contain German language phonotactic rules as well as the respective target words in a stressed position. Before starting the recordings the speakers will be asked to lay attention to accentuation (s. Cho, 2004).

Within this framework five questions offer themselves for investigation:

1.) The manner and degree of coarticulation induced within intervocalic consonants, except affricates, the glottal stop and [j] by the same vocalic context embedded in carrier sentences will be measured. In which ways do these consonants differ between German and L2 English speech? Do the consonants in German vs. L2 English speech differ in their manner and degree of coarticulation/coarticulatory resistance? The vowels used, except [y], exist in

German as well as in British English (see IPA transcriptions of the Phonetic Society [German: Kohler, 1999; Mangold, 2005; British English: Gimson, 1997]). British English containing more vowels than German, and all German vowels also exist in the British English international phonetic alphabet. The vowel [y] will be taken because we are interested in examining if it will continue to be articulated in the German specific manner considering that the English phonetic alphabet does not contain this vowel. Considering to the DAC theory proposed by Recasens et al. (1997) as well as Recasens (2004) it is especially interesting to investigate whether the same hierarchy of consonants articulated with more tongue dorsum activity being less permeable for the surrounding vowels in contrast to those articulated with less or totally without the participation of the tongue can be found. Therefore, we especially take into account several consonants which differ with regard to their participation of tongue dorsum, namely [p], [t], [k], [l], [n], [s].

2.) According to the findings of Gobl and Ní Chasaide, 1999 as well as Hoole, 1999, German voiceless plosives differ from British English voiceless plosives with regard to EE, which is very weak in German, while dynamic leakage, RA, is very high and the shape of the glottal pulse is very symmetrical, RK (Gobl & Ní Chasaide, 1999: 136). For German stimuli a gradually rising L0 has been noticed which reaches constant values only after 30 ms or so. Following [p^h] L1 is initially weak. In contrast an abrupt initiation of voice following [p^h] with strong excitation, EE, has been found and not always an involvement of the weak, very gradual onset in English which is typical for the German data (Gobl & Ní Chasaide, 1999: 137 f.). Glottal adduction, RA value is supposed to be low in English and L1 is less high in English as compared to German data. Is it possible to replicate the findings of Gobl and Ní Chasaide (1999) as well as Hoole (1999)? In which manner and degree do the German plosives differ from the English plosives taking into account the parameters mentioned above?

3.) Following Ziegler and von Cramon (1986) the comparison of temporal anticipation of articulatory configurations in Germans vs. L2 English speakers is of great interest. In the present study, subjects will produce /gətVtə/ with /i, y, u, a/ (Ziegler & von Cramon, 1986: 35 ["In the case of schwa, the formant frequencies suggest themselves as appropriate parameters to describe vowel-to-vowel coarticulation."]) embedded in a carrier phrase. Is it possible to find differences in the temporal anticipation of articulatory configurations in native vs. L2 English speakers? According to Öhmann (1966) consonant gestures are also superimposed upon vowel-to-vowel gestures in VCV sequences in German. Can this so called superimposition also be found in L2 English speech? And if so, how does it differ with regard to native German compared to L2 English speech? In L2 speech there might be a delay of anticipatory gestures in the phase of relationship of individual articulatory gestures.

4.) In which manner and degree do the formant frequency productions vary while comparing those two groups? Taking into account the Exemplar Theory developed by Pierrehumbert (2001) we follow the assumption that talented learners of L2 English produce a wider range of vowel formant frequencies than untalented learners.

3. Hypotheses

1.) F_2 and F_3 exhibit a positive correlation with dorsopalatal contact degree and, thus, with palatal constriction narrowing (Fant, 1960; Recasens & Pallarès, 1995). Following the findings of Recasens et al. (1997) and Recasens (2004) we expect to find the same hierarchy of coarticulatory resistance degree (CR) regarding consonants produced with different degrees of tongue dorsum activity.

2.) The results of the studies of Gobl and Ní Chasaide (1999) as well as Hoole (1999) will be replicated in talented vs. untalented speakers of L2 English speech while talented speakers come nearer to the British English parameter values in contrast to untalented speakers staying closer to the German language parameter values (s. the following table).

Table 1 shows the differences in voice source parameters (EE, RA, RK, RG, L0, L1) of British English vs. German.

Nr.	*Language*	*Acoustic parameters*
1	*(BE) British English*	EE -> strong; RA -> weak (incomplete glottal adduction); L0 -> rather constant; L1 -> amplitudes not high compared to German data, [p^h] follows in voiceless plosives
2	*(G) German*	EE -> weak; RA -> strong; L0 -> rises gradually attaining constant values only after 30 ms or so; L1 -> initially weak following [p^h] in voiceless plosives

3.) In L2 speech there is a delay of anticipatory gestures in the phase of relationship of individual articulatory gestures.

4.) Untalented learners of L2 English vary less in their formant frequency production than untalented learners (see paragraph about Exemplar Theory).

4. Experimental method

First, four English sentences will be presented to the subjects spoken by a standard British English speaker. In these sentences the target sequence /gəCVtə/ occurs. Afterwards they will be asked to repeat these sentences as similarly to the speech of the British English speaker as possible. After having produced free speech subjects will be exposed to the sentences under investigations which they have to read aloud from a screen in the sound-attenuated room of the Institute of Natural Language Processing, Universität Stuttgart, Germany. One after another they will fluently read the sentences “I have said /gəCVtə/ twice.” five times each in which each consonant and the vowels /i, y, u, a/ appear. The second part of the experiment is similar. German speaker will read aloud four

German sentences including the sequences under investigation. The subject will again be asked to repeat the spoken sentences. Participants will be seated on a chair in front of a computer screen (distance ca. 150 cm) in the sound-attenuated room to read aloud the same target sequences as before but embedded in the German sentence “Ich habe /gəCVtə/ gesagt.”. The reading task will begin each time with the last target sequence which subjects had to imitate during the repetition of text read by a British English and German native speaker.

5. Recording procedure

Subjects will be recorded in the sound-attenuated room of the Institute of Natural Language Processing, Stuttgart, Germany, with the help of a headset with an AKG C420 and digitized with the Yamaha O3D digital recording and mixing console at a sampling rate of 48Khz, 16 Bit quantization accuracy.

6. Method

6.1. Method of analysis

The work on coarticulatory resistance will consist of spectra analysis of the acoustic data using the linear prediction coding method (LPC). Following Recasens (1999) the first vowel (V_1) will be manually cut out of VCV-sequences and marked in the middle with the program “WaveSurfer” to get F_1, F_2 as well as F_3 to compare those frequencies produced by the subjects. After having measured those vowel formant frequencies, graphs will be plotted with the help of the free statistic tool “R“ to get a representation of the vowel formants produced and those frequencies will be statistically analyzed. Additionally, the coarticulatory effects of vowel-on-consonant (V-on-C), C-on-V1 as well as the C-on-V2-effect, will be measured while comparing the differences in length between these time periods. The temporal characterization patterns will be measured to get an overview of formant frequencies, stop articulation differences and the temporal anticipation of articulatory configurations in Germans vs. L2 English speakers. The analysis of the voiceless stop characteristics contain vowel formant frequency measurements (L1 and L2), normalization of L0 and L1 which leads to the possibility to compare between spectral level data of different speakers and languages, measurements of harmonic values (amplitude of the first harmonic, L0) and analysis of dB values to detect the excitation strength (EE). RA, RK and RG will be calculated according to Gobl and Ní Chasaide (1999: 317 f.). A comparison of German speech vs. L2 English speech will be conducted in order to investigate differences in these parameters between productions across the two language contexts. Finally, these measurements, and their differences across language contexts, will be compared across speaker groups (talented vs. untalented).

Overview of the planned investigations

Table 2 gives an overview about the objectives of study, the stimuli which will be taken to follow these special objectives and finally the measurements which have to be done to further analyze the stimuli.

Objective	***Stimuli (will be all embedded in carrier sentences)***	***Methods***
coarticulation/coarticulatory resistance, coarticulatory effects of vowel-on-consonant, temporal anticipation of articulatory configurations in Germans vs. L2 English speakers	VCV-sequences including all German consonants except affricates, the glottal stop and [j] in the sequence /gəCVtə/ with /i, y, u, a/	LPC-analysis of F_1, F_2 and F_3, measurements of C-on-V1 as well as the C-on-V2-effects looking at the time periods [ms], harmonic analysis, analysis of EE, RA, RK RG, L0, L1 (s. above), voice onset time and temporal characterization patterns

6.2. Method of statistical analysis for all analyses being conducted

Formant data will be averaged across repetitions. For the same sequence one-way repeated measures ANOVAs with 1df between groups comparison between two populations, 4df within groups (since each group inducted data of several speakers). For an effect to be deemed significant, at 0.05 level, a significant difference must be required between groups and within-groups. These calculations will be done using the free statistic tool „R“.

7. Further research to be done

Discrimination i.e. perception tasks

While looking at the results of the VOT measurements it might be a follow-up study to present different British English consonant as well as vowel allophonic variations to talented as well as untalented learners of L2 English in order to investigate differences in discrimination ability between the groups.

REFERENCES

Ashby, M. & Maidment, J. (2005). *Introducing Phonetic Science.* Cambridge: Cambridge University Press.

Baumotte, H., Lenz, M., Dogil, G. & Reiterer, S. (2007). Coarticulatory resistance as a basis for a foreign accent: V-to-V coarticulation in German VCV-sequences: a pilot study. *Belgian Royal Academy of Sciences and Arts.* in press.

Benguerel, A.-P., Hirose, H., Sawashima, M. & Ushijima, T. (1977a). Velar coarticulation in French: an electromyographic study. *Journal of Phonetics*, 5: 159-167.

Benguerel, A.-P., Hirose, H., Sawashima, M. & Ushijima, T. (1977b). Velar coarticulation in French: a fiberscopic study. *Journal of Phonetics*, 5: 149-158.

Bladon, R. A. W. & Al-Bamerni, A. (1976). Coarticulation resistance in English /l/. *Journal of Phonetics*, 4: 137-150.

Bladon, R. A. W. & Nolan, F. J. (1977). A video-fluorographic investigation of tip and blade alveolars in English. *Journal of Phonetics*, 5: 185-193.

Browman, C. P. & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology*, 3: 219-252.

Brunner, J., Fuchs, S. & Perrier, P.. On the relationship between palate shape and articulatory behaviour. *Journal of the Acoustical Society of America.* forthcoming.

Bybee, J. (2005). *From usage to grammar: The mind's response to repetition.* LSA Presidential Address.

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, 32: 141-176.

Clark, J., Yallop, G. & Fletcher, J. (2007). *An Introduction to Phonetics and Phonology*. Oxford: Blackwell Publishing.

Docherty, G. J. (1992). *The Timing of Voicing in British English Obstruents.* Berlin: Fortis.

Fant, G. (1960). *Acoustic theory of speech production.* The Hague: Mouton.

Farnetani, E. (1997). Coarticulation and Connected Speech Processes. In: Hardcastle, W. J. & Kent, R. D. (1977). *Coarticulation in recent speech production models.* Oxford: Blackwell Publishers, pp. 115-133.

Fuchs, S. & Toda, M. (2007). Do differences in male versus female /s/ reflect biological factors or sociophonetic ones? In: Fuchs, S., Shadle, C., Toda, M. & Zygis, M. (eds). *Phonetics and phonology of turbulent sounds.* Berlin: Mouton de Gruyter.

Gimson, A. Ch. (1997). Gibson's *pronunciation of English.* 5. ed. London: Arnold.

Goble, Ch. & Ní Chasaide, A. (1999). Voice source variation in the vowel as a function of consonantal context. In: Hardcastle, W. J. & Hewlett, N. (eds). *Coarticulation. Theory, Data and Techniques.* Cambridge: Cambridge University Press, pp. 122-143.

Goble, Ch. & Ní Chasaide, A. (1999). Techniques for analysing the voice source. In: Hardcastle, W. J. & Hewlett, N. (eds). *Coarticulation. Theory, Data and Techniques.* Cambridge: Cambridge University Press, pp. 300-321.

Hoole, P., Gfroerer, S. & Tillmann, H.-G. (1990). Electromagnetic articulatography as a tool in the study of coarticulation. *Forschungsberichte des Instituts für Phonetik, Munich*, 28: 107-122.

Hoole, P. (1999). Laryngeal coarticulation. In: Hardcastle, W. J. & Hewlett, N. (eds). *Coarticulation. Theory, Data and Techniques.* Cambridge: Cambridge University Press, pp. 107-121.

Kohler, K. J. (1999). German. In: International Phonetic Association (Ed). Handbook of the International Phonetic Association. A Guide to the Use of the International Phonetic Alphabet. Cambridge: Cambridge University Press.

Kühnert, B., Ledl, C., Hoole, P. & Tillmann, H. G. (1991). Tongue-jaw interactions in lingual consonants. In: Engstrand, C. & Kylander, C. (eds). *Perilus XIV: Current Phonetic Research Paradigms. Implications for Speech Motor Control*, pp. 21-25.

Lehiste, I. (1964). *Some Acoustic Characteristics of Selected English Consonants.* Research Center in Anthropology, Folklore and Linguistic. Indiana University, 34.

Lubker, J. F. (1968). An electromyographic-cineradiographic investigation of velar function during normal speech production. *Cleft Palate Journal*, 5: 1-18.

Majewski, W., Rothman, H. B. & Hollien, H. (1977). Acoustic comparisons of American English and Polish. *Journal of Phonetics,* 5: 247-251.

Mangold, M. (2005). *Duden, Aussprachewörterbuch.* 6. ed. Mannheim: Dudenverlag.

Öhman, S. E. G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, 39: 151-168.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In: Bybee, J. & Hopper, P. (eds). *Frequency and the Emergence of Linguistic Structure.* Amsterdam: Benjamins, pp. 137-157.

Recasens, D. (1984a). V-to-C coarticulation in Catalan VCV sequences: an articulatory and acoustical study. *Journal of Phonetics*, 12: 61-73.

Recasens, D. (1984b). Vowel-to-vowel coarticulation in Catalan VCV sequences. *Journal of the Acoustical Society of America*, 76: 1624-1635.

Recasens, D. (1985). Coarticulatory patterns and degree of coarticulatory resistance in Catalan CV sequences. *Language and Speech*, 28: 97-114.

Recasens, D., Fontdevila, J. & Pallarès, M. D. (1995). Velarization degree and coarticulatory resistance for /l/ in Catalan and German. *Journal of Phonetics,* 23: 37-52.

Recasens, D., Pallarés, M. D. & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, 102: 544-561.

Recasens, D. (1999). Acoustic analysis. In: Hardcastle, W. J. & Hewlett, N. (eds). *Coarticulation. Theory, Data and Techniques.* Cambridge: Cambridge University Press, pp. 323-336.

Recasens, D. & Pallarès, M. D. (2001). Coarticulation, assimilation and blending in Catalan consonant clusters. *Journal of Phonetics,* 29: 273-301.

Recasens, D. (2004). The effect of syllable position on consonant reduction (evidence from Catalan consonant clusters). *Journal of Phonetics*, 32: 435-453.

Roach, P. (1992). *Introducing Phonetics.* New York: Penguin Books.

Rochet, A. P. & Rochet, B. L. (1991). The effect of vowel height on patterns of assimilation nasality in French and English. *Proceedings of the XIIth International Congress of Phonetic Sciences, Aix-en-Provence*, 3: 54-57.

Saville-Troike, M. (2005). *Introducing Second Language Acquisition*. Cambridge: Cambridge University Press.

Stevens, K. N. (1998). Acoustic Phonetics. Cambridge: MIT Press.

Strand, E. A. (1999). Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology*, 18 (1): 86-99.

Stuart-Smith, J.. Empirical evidence for gendered speech production: /s/ in Glaswegian. In: Cole, J. & Hualde, J. (eds). *Change in Phonology. Laboratory Phonology*, 9. Berlin: Mouton de Gruyter. in press

Vollmer, K. (1993). *Untersuchung der Koartikulation bei gestörter Sprache: Experimentalstudie.* Magisterarbeit im Fach Linguistik, Universität Bielefeld.

Vollmer, K. (1997). *Koartikulation und Glottale Transparenz.* Dissertationsschrift zur Erlangung der Würde eines Doktors, Fakultät für Philosophie, Universität Stuttgart.

Ziegler, W. & von Cramon, D. (1985). Anticipatory Coarticulation in a Patient with Apraxia of Speech. *Brain and Language*, 26: 117-130.

Ziegler, W. & von Cramon, D. (1986). Disturbed Coarticulation in Apraxia of Speech: Acoustic Evidence. *Brain and Language, 29*: 34-47.

Ziegler, W. (1989). Anticipatory Coarticulation in Aphasia: More Morphology: A Reply to Sussman et al. (1988) and Katz (1988). *Brain and Language*, 37: 172-176.

Phonetic convergence as a paradigm of showing phonetic talent in foreign language acquisition

NATALIE LEWANDOWSKI[1], MATTHIAS JILKA[2], GIUSEPPINA ROTA[1], SUSANNE REITERER[3,1] & GRZEGORZ DOGIL[1]

[1]Institute for Natural Language Processing, Experimental Phonetics Group, University of Stuttgart
[2]Institute for Linguistics, Anglistics, University of Stuttgart
[3]Department of Neuroradiology, Section for Experimental MR of the CNS, University of Tübingen

This paper gives an overview of an ongoing Ph.D. project in cooperation with the DFG supported project "Language talent and brain activity" at the Universities of Stuttgart and Tübingen. Considering the socio-psychological background of communication accommodation theory and the previous research on convergent behavior, done predominantly in linguistically homogeneous situations, it wants to turn to convergence during interaction in a second language environment. Shifting the focus to native-nonnative interaction can provide valuable insights into both, second language communication behavior and the underlying mechanisms enforcing or hampering convergence. The specific aim is to determine to what extent the factor of phonetic language aptitude translates into phonetic convergence during conversation, which has been defined as an increase in segmental and suprasegmental similarities between two speakers (Pardo, 2006).

Introduction

Individuals differ to a considerable degree in the ability to learn a foreign (second) language. Those differences in the rate of learning and the overall acquired competence are the result of interindividually varying psychological and sociological features. One of them is undoubtedly the factor of language talent that has unfortunately been left out of research on second language acquisition for a long time due to the assumed equality of chances in learners. The topic has recently become more popular again, amongst other reasons due to the development of new neuroimaging methods, allowing insights into the working brain and the tracking of specialized brain areas.

This Ph.D. project is aimed at investigating the link between second language aptitude and the specific ability to accommodate to a conversational partner in a face-to-face situation. Due to the special status of the phonetic sub-skill of language talent, which appears to run separately from other subcomponents of talent, attention is focused on phonetic convergence within native-nonnative interactions.

---

[1] Natalie Lewandowski, Prof. Dr. Grzegorz Dogil, Dr. Giuseppina Rota, Institut für Maschinelle Sprachverarbeitung, Experimentelle Phonetik, Azenbergstr. 12, D-70174 Stuttgart, email: *first_name.last_name*@ims.uni-stuttgart.de.

[2] Dr. Matthias Jilka, Institut für Linguistik: Anglistik, Heilbronnerstr. 17, D-70174 Stuttgart, email: jilka@ifla.uni-stuttgart.de.

[3] Dr. Susanne Reiterer, Department of Neuroradiology, Section for Experimental MR of the CNS, University of Tübingen, Germany, email: susanne.reiterer@med.uni-tuebingen.de.

Communication Accommodation Theory

Research on the phenomenon of phonetic convergence has its origin in the Communication Accommodation Theory (CAT) that has been established in the 1970s, first under the term of “speech accommodation theory” (SAT, Giles, 1973). SA theorists began to argue that some aspects of the Labovian paradigm (Labov, 1966) concerning the formality-informality of context could be explained by interpersonal accommodation processes. They wanted to introduce a more detailed concept of *context*, which was a major term in sociolinguistic theory of that time. One of the goals was increasing the attention for specific dimensions of context, like language itself and the role of the receiver in the interaction (Giles, Coupland & Coupland, 1991). Since then SAT focused on clarifying the motivations underlying speech, the constraints imposed on it and finally their consequences for social interaction (Giles et al., 1991). After the theory has widened its focus onto nonverbal behavior and general dimensions of discourse, the theory evolved to communication accommodation theory (Giles et al. 1987, cit. after Giles et al., 1991).

CAT defines itself nowadays as an interdisciplinary approach placed at the interface of social psychology, sociology, sociolinguistics and communication. The field of interest has also evolved from a merely interindividual perspective to a broader perspective of macrovariation. The language used in interaction can be a remarkably good indicator of status differences, ethnic boundaries or can as well define ingroup or outgroup boundaries and impose conformist behaviors (Shepard, Giles & Le Poire, 2001). The crucial thesis of CAT states that language is used to achieve a desired degree of social distance between self and an interacting partner. They argue that accommodation is a complex, context-sensitive set of alternative behaviors available to interacting partners in a face-to-face situation. As mentioned earlier, it can serve to achieve solidarity with or dissociation from a partner, in a dynamic setting with online feedback (Giles et al.,1991). This distance can be negotiated by means of approximation strategies (convergence, divergence, maintenance and complementarity), interpersonal control, discourse management and interpretability (Shepard et al., 2001).

Accommodation strategies

The term divergence is used to refer to a set of behaviors allowing speakers to distance themselves in terms of verbal and nonverbal behavior from their partners. This can be achieved via an explicit accentuation of differences in speech style and/or facial expressions and gestures (Giles et al., 1991). As an outcome of later research the notion of “perceptual/subjective divergence” was introduced, covering the possible discrepancies between the performance and the actual perception of a conversational partner (Shepard et al., 2001).

Convergence, as one of the main strategies within the CAT framework, describes the adaptation of communicative behaviors towards those of a conversational partner in various verbal and nonverbal features. The characteristics being adapted to can range from gestures, smile, facial affect, head nodding, information density, voice quality, speech rate, utterance length to pausing frequencies and response latency (Giles et al., 1991).

Speakers may also adhere to their own current speaking style and wish neither to converge nor to diverge. Intended maintenance may however often be interpreted as diverging from the interlocutor (Shepard et al., 2001). Speech complementarity strategies on the other hand involve the accentuation of distinctive features between partners with different social roles, e.g. men speaking with a lower voice while talking to women (Hogg, 1985).

Accommodation can vary in many dimensions, be it direction, degree or level. Since all strategies are highly dependent on the given situational context, many variables can influence the final degree or type of accommodation. Crucial for the choice of upward or downward accommodation is the power structure within the dyad (Street, 1982). Upward convergence describes a movement towards a socially more accepted variety or style whereas downward accommodation relates to the adaptation of less prestigious forms (Giles et al., 1991). Other distinctions can be made as to the modality (unimodal vs. multimodal, i.e. occurring across several different behaviors), the direction (unidirectional vs. mutual) and symmetry. It has also been remarked (Shepard et al., 2001) that accommodation can occur partially or to a "full" extent, where the partners' behaviors match exactly.

Phonetic convergence

Although many studies since the emergence of the Accommodation Theory have investigated the linguistic properties of convergent behavior, the term of phonetic convergence is relatively new. It has been defined in terms of an increase of segmental and suprasegmental similarities between the interacting speakers (Pardo, 2006).

An important question within the research on phonetic accommodation remains the role of the perception-production link. Some accounts propose that speech perception automatically yields relevant linguistic (in this case phonetic and phonological) parameters that cause production and lead directly to imitation (Sancier & Fowler, 1997). If this holds true and there indeed is a very close link between perception and production, it should favor a fairly exact imitation of words at an articulatory and acoustic level. However, this has been proved to be extremely unlikely since even for a single speaker no two productions of the same utterance are acoustically or articulatory identical (Pardo, 2006). Far more probable is an intended imitation with a moderate degree of exactness. Interestingly, none of the accounts being currently under discussion explains the influence of obvious perception and production limitations and other factors yielding discrepancies in the imitative responses (Pardo, 2006).

So far phonetic convergence has been studied along the following features:

- speech rate (Street, 1984)
- fundamental frequency, amplitude contours (Gregory, 1990)
- voice onset time (VOT) (Sancier & Fowler, 1997)
- amplitude, utterance duration and rate (Oviatt, Darves & Coulston, 2004)
- perceptual similarity of pronunciation (Pardo, 2006).

Most of the studies, except that of Sancier and Fowler (1997), investigated phonetic convergence in a native language environment. Sancier and Fowler concentrated on the gestural drift between the two languages of a bilingual speaker. Zuengler (1991) has pointed to the crucial role of accommodation in a second language acquisition context, analyzing data on native-nonnative or fully nonnative conversational interactions. The described dissertation project wants to provide further details on second language face-to-face communication and explore possible underpinnings of phonetic convergence in an L2-environment.

Convergence and language aptitude

Considering the described biases in the imitation quality and the interindividual and context-dependent variability in the degree of convergence, we want to investigate the link between phonetic convergence and the individual differences in L2-speakers, more specifically their phonetic aptitude. Our main premise is based on the direct link between the ability to converge in pronunciation and a general phonetic language talent. When converging in pronunciation to a foreign language communication partner, the skills to do so obviously have to be given. The negative consequence of this would be a lack of convergence due to poor (language) ability (Giles & Powesland, 1997). This is even more significant when the accommodation takes place during a relatively short conversational interaction since it rules out the possibility of underlying long-term learning processes concerning a foreign accent (as it would be the case during longer stays in a foreign language environment).

Assessing phonetic aptitude

Language talent is considered to be one of the main features, in which language learners differ from one another. Studies on individual differences in the process of second language acquisition have shown that variation in language aptitude, personality traits and attitudes (motivation) leads to significant contrasts both in the rate of learning and in the eventually acquired proficiency (e.g. Dörnyei, 2005). Some factors as openness, extraversion and empathy combined with a high aptitude are assumed to translate into better communicative skills and phonetic accuracy (due to a lower socio-affective filter, Krashen, 1981).

Skehan (1998) defines talented learners simply as people with greater levels of aptitude and therefore likely to make faster progress in language learning. This language aptitude can be further divided into subareas. There is e.g. strong evidence for the phonetic subcomponent of talent to run fairly independent from other sub-skills. Well known in the literature is the so called "Joseph-Conrad-phenomenon", named after the famous Polish-born writer. It describes a situation in which despite perfect knowledge of grammar and vocabulary, a foreign accent in the second language is still prominent. Schneiderman & Desmarais (1988) provide also evidence for the phonetic sub-skills to have a special status, differing in many points especially from the grammatical competence. Some data is available from research on the maturational constraint/critical age period (Birdsong, 2006), where nativelikeness is reported less often for pronunciation than for other competences in a second language. The independence of the phonetic component of language talent may as well suggest the existence of a separate underlying neural substrate.

The independent role of a phonetic component of talent has also been recognized within frameworks for language aptitude testing. One of them is the Modern Language Aptitude Test (MLAT, Carroll & Sapon, 1959), which operationalized the following factors underlying aptitude: phonemic coding ability, associative memory, inductive language learning ability and grammatical sensitivity. Phonemic coding ability was described as the capacity to both, make sound discriminations and to code foreign sounds in a way that allows a recall at a later moment in time (Skehan, 1998).

Within our project a comprehensive approach for testing phonetic language talent was introduced (Jilka et al., 2007). Applying an extensive test battery should therefore allow covering not only the linguistic but also the psychological and sociological characteristics resulting in individual differences between the L2-learners. All subjects are German native speakers learning English as a foreign language. Other variables like educational background, brain lateralization and handedness, amount of time spent in English speaking countries are also controlled for. The linguistic part includes production and perception tests in German and English, with various elicitation techniques ranging from free speech to story-telling, direct and delayed imitation and reading tasks. The focus of the perception tasks lies on the identification of foreign accents and the prosodic features of utterances (in both native and nonnative contexts). The phonetic part of the test battery serves as a means of identifying all subskills that could underlie pronunciation talent (e.g. a sensibility for identifying prosodic variation or segmental changes) and tear them apart from mere proficiency.

After an additional series of psychological and personality questionnaires (including e.g. non-verbal intelligence, verbal intelligence, empathy, mental flexibility), the tests will be evaluated and the subjects finally classified into two groups according to their pronunciation talent. Both groups, the high-aptitude and low-aptitude learners, will be subject to further research on phonetic convergence in native-nonnative interaction. Testing for phonetic convergence is especially promising for showing phonetic *talent* as opposed to *proficiency*, since convergence in pronunciation within a short conversation rules out the possibility of underlying long-term learning effects.

Methodology

As to the accommodative behavior of conversational partners in an L2-environment the following predictions can be made:

- More talented learners will exhibit a stronger tendency to converge in their pronunciation to their conversational partners (native speakers of English).
- Less phonetically talented speakers should show considerable difficulties in accommodating to their partner.
- Certain segmental and suprasegmental features will be adapted easily and with a good outcome.
- Certain features of the English accent should be hard to take up and subjects will often fail to converge in those cases.

In order to investigate the above questions, a diapix experiment is planned (Bradlow et al. 2007). Forty participants will be asked to perform a diapix task in English together with a native speaker of English. In order to control for any unintended accent shifts, a speaker of Standard Southern British English (SSBE) and of General American (GA) will be invited, according to the subjects' accent preferences. The diapix is a dialogue based picture-matching game that elicits a wide range of utterance types and provides balanced speaker roles (Fig.1a and 1b). The participants are confronted each with one picture of a set which differ in ten details from one another. They can't see each other's picture and need to spot the ten differences while talking to their partner and describing their own picture. The native speaker will be instructed to pay special attention to avoiding any downward convergence to the nonnative accent their partners might have. After the experimental session, performed in an anechoic chamber, the recordings will be subject to acoustic analyses enabling both a global (e.g. speaking rate, f0 contour, pitch range) and a fine-grained picture (e.g. MFCCs; episodic memory store of amplitude envelope signals, Wade 2007) of the convergent features. A further step will be a perceptual judgment experiment and a search for possible correlations between these results and results of the previously performed psychological tests.

Fig. 1a and 1b: A set of pictures for the diapix task with ten differences to spot (with the courtesy of Ann R. Bradlow).

Summary

This paper gave a brief overview of the theoretical underpinnings of speech convergence, grounded in the Communication Accommodation Theory and its relation to some aspects of second language acquisition research. The planned research on phonetic convergence in native-nonnative interaction is supposed to answer a range of questions about the phonetic subcomponent of language talent and the communicative behavior of second language learners. Obtaining auditory data from a diapix task will enable a detailed analysis of the convergent behavior of talented in contrast to untalented learners of English as an L2. Especially interesting is the extent to which those two groups differ from each other in adapting native features of the English accent and which aspects are particularly easy or difficult to take up. In addition to that, assessing pronunciation talent via phonetic convergence could provide an answer to the still open questions on the functioning of the perception-production link.

References

Birdsong, D. (2006). Age and Second Language Acquisition and Processing: A selective overview. *Language Learning,* 56 (1), 9-49.

Bradlow, A. R., Baker, R. E., Choi, A., Kim, M. & van Engen, K. J. (2007). The Wildcat Corpus of Native- and Foreign-Accented English. *Journal of the Acoustical Society of America,* 121(5), Pt.2, 3072.

Carroll, J. B. & Sapon, S. M. (1959). *Modern Language Aptitude Test*. New York: Psychological Corporation.

Dörnyei, Z. (2005). *The Psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Erlbaum.

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics*, 15, 87-105.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation Theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of Accommodation: Developments in Applied Sociolinguistics* (pp. 1-68). Cambridge: Cambridge University Press.

Giles, H., Mulac, A., Bradac, J. J., & Johnson, P. (1987). Speech accommodation theory: the next decade and beyond. In M. McLaughlin (Ed.), *Communication Yearbook* (pp. 13-48), Newbury Park, CA: Sage.

Giles, H., & Powesland, P. (1997). Accommodation Theory. In N. Coupland, & A. Jaworski (Eds.), *Sociolinguistics. A Reader and Coursebook* (pp. 232-239). Basingstoke: Macmillan.

Gregory, S. W. (1990). Analysis of fundamental frequency reveals covariation in interview partner's speech. *Journal of Nonverbal Behavior,* 14, 237-251.

Hogg, M. (1985). Masculine and feminine speech in dyads and groups: A study of speech style and gender salience. *Journal of Language and Social Psychology,* 4, 99-112.

Jilka, M. (2000). The Contribution of Intonation to the Perception of Foreign Accent (Doctoral dissertation, Institute for Natural Language Processing, University of Stuttgart). *Arbeiten des Instituts für Maschinelle Sprachverarbeitung (AIMS),* 6/3, University of Stuttgart.

Jilka, M., Baumotte, H., Lewandowski, N., Reiterer, S. & Rota, G. (2007). Introducing a comprehensive approach to assessing pronunciation talent. *Proceedings of the XVI th ICPhS,* Saarbrücken, 1737-1740.

Krashen, S. D. (1981). *Second language acquisition and second language learning.* Oxford: Pergamon Press.

Labov, W. (1966). *The social stratification of English in New York City.*Washington, DC: Center for Applied Linguistics.

Oviatt, S., Darves, C., & Coulston, R. (2004). Toward Adaptive Conversational Interfaces: Modeling Speech Convergence With Animated Personas. *ACM Transactions on Computer-Human Interaction,* 11, 300-328.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America,* 119, 2382-2393.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics,* 25, 421-436.

Schneiderman, E. & Desmarais, C. (1988). A Neuropsychological Substrate for Talent in Second-Language Acquisition. In Obler, L., Fine, D. (Eds.), *The Exceptional Brain: The Neuropsychology of Talent and Special Abilities*. New York: Guildford, 103-126.

Shepard, C. A., Giles, H., & Le Poire, B. A. (2001). Communication Accommodation Theory. W. P. Robinson, & H. Giles (Eds.), *The New Handbook of Language and Social Psychology* (pp. 33-56). Chichester: John Wiley & Sons.

Skehan, P. (1998). *A Cognitive Approach to Language Learning.* Oxford: Oxford University Press.

Street, R. L., Jr. (1982). Evaluation of noncontent speech accommodation. *Language and Communication,* 2, 13-31.

Street, R. L., Jr. (1984). Speech convergence and speech evaluation in fact-finding interviews. *Human Communication Research, 11, 139-169.*

Wade, T. (2007). Implicit rate and speaker normalization in a context-rich phonetic exemplar model. *Proceedings of the XVI th ICPhS,* Saarbrücken, 765-768.

Zuengler, J. (1991). Accommodation in native-nonnative interactions: Going beyond the “what” to the “why” in second language research. In H. Giles, & J. Coupland, & N. Coupland (Eds.), *Contexts of Accommodation: Developments in Applied Sociolinguistics* (pp. 223-244). Cambridge: Cambridge University Press.

Repeated Masked Semantic Priming with New Results: ERPs of a Negative Semantic Priming Effect

CHRISTINA BERMEITINGER, CHRISTIAN FRINGS, AND DIRK WENTURA

Cognitive Psychology – Cognition and Action Unit
Saarland University[1]

In semantic priming experiments with clearly visible primes, participants usually respond faster to targets which are preceded by a related (BIRD–finch) than by an unrelated prime (FLOWER–finch). Using a new masking technique, at which prime and mask repeatedly and rapidly interchange, there is evidence for a negative priming effect, i.e. slower responses to related than to unrelated targets (e.g., Wentura & Frings, 2005, JEP:LMC). In a new experiment using this technique, we examined additionally event-related potentials. The N400 is the most prominent event-related potential correlated to semantic processing. For unrelated targets, there is usually a more negative wave than for related targets. However, in our study, there is more negativity for related than for unrelated targets which corresponds to the behavioral data. In this context, this ERP component seems to reflect a locally and temporarily reduced access to related targets or – speaking in terms of the P300 – the difficulty to discriminate prime and target representations.

Semantic Priming

The semantic priming paradigm is a traditional tool for investigating the structure of and mechanisms within semantic memory (for reviews see Neely, 1991; McNamara & Holbrook, 2003; McNamara, 2005). Usually, a prime (e.g., the word 'BIRD') precedes the target (e.g., the word 'finch') and participants have to respond to the target by naming it or by categorizing it as a legal word or non-word. Typically, responses are faster if prime and target are semantically related (e.g., BIRD – finch) compared to semantically unrelated prime-target pairs (e.g., FLOWER – finch).

There are several accounts of the semantic priming effect with the most basic distinction being the one distinguishing between strategic and automatic mechanisms (see Neely, 1991). Of course, the silver bullet to explore the basic processes of semantic memory is to prevent the prime from becoming conscious. Often, this will be realized by masking the prime. For a masked presentation of primes (i.e., primes are presented near the subjective or objective discrimination threshold), it is assumed that effects arise from automatic processes. There are several studies which yielded masked priming effects with behavioral measures (i.e., response times and/or errors). Most of them showed positive priming effects (i.e., faster reactions to related than to unrelated targets; e.g., Fowler, Wolford, Slade, & Tassinary 1981; Marcel,

---

[1] Correspondence concerning this article should be addressed to Christina Bermeitinger, Saarland University, Department of Psychology, Campus A2 4, D – 66123 Saarbrücken, Germany or via email to cbermeit@mx.uni-saarland.de.

The research reported in this article was supported by a grant of the Deutsche Forschungsgemeinschaft to Dirk Wentura and Christian Frings (WE 2284/5-2).

1983; Spence, 1983; Kiefer & Spitzer, 2000; Kiefer, 2002; Ruz, Madrid, Lupiáñez, & Tudela, 2003; Grossi, 2006). However, there are other studies that revealed negative priming effects given some specified conditions (e.g., Dagenbach, Carr, & Barnhardt, 1990; Durante & Hirshman, 1994; Kahan, 2000).

Theories of the Semantic Priming Mechanisms

The canonical models to explain positive semantic priming effects in terms of automatic processes are spreading activation models (e.g., Collins & Loftus, 1975; Anderson, 1976, 1983a, 1983b). Essentially, these models assume that the retrieval of an item activates its internal representation and that activation spreads from one concept to related concepts (i.e., from the prime to the target). Then, residual activation subsequently facilitates the retrieval of these related concepts (i.e., the targets). The second approach is represented by distributed network models which have a long history but have become influential only during than the last two decades ago for explaining semantic priming effects (e.g., Masson, 1995; McRae, de Sa, & Seidenberg, 1997). Here, priming occurs because the processing of the target starts from the pattern of activity established by processing the prime. Since related concepts have more similar patterns than unrelated prime-target pairs, the transition from prime to target pattern is accelerated. The third kind of models for explaining semantic priming effects are compound-cue models (e.g., Ratcliff & McKoon, 1988; Dosher & Rosedale, 1989). The target and elements of the context (including the prime) build the retrieval cue to memory. Together with models of memory (e.g., Theory of Distributed Associative Memory, TODAM, Murdock, 1982), compound-cue theories can account for faster reactions to related than to unrelated targets by assuming that the familiarity of a cue (containing target and prime) will be higher for related than unrelated pairs of words.

Event-Related Potentials of Semantic Priming

The most prominent event-related potential correlated to semantic processing is the N400 component (for reviews see Kutas & Van Petten, 1994; Brown & Hagoort, 2000; Kutas & Federmeier, 2000).[2] The N400 is a negative going ERP wave at approximately 400 ms after target onset and with a centroparietal maximum. It is elicited when a word occurs in an unexpected or unrelated context. This context could be a sentence (e.g., Kutas & Hillyard, 1980) or another word (e.g., Bentin, McCarthy, & Wood, 1985; Holcomb, 1988).

The underlying mechanism for the N400 priming effect (i.e., more negativity to unrelated than to related targets) are still under debate. Some favor the assumption of automatic processes like spreading activation (e.g., Deacon, Uhm, Ritter, Hewitt, & Dynowska, 1999), some favor the assumption of controlled processes (e.g., Brown, Hagoort, & Chwilla, 2000). Nevertheless, there is evidence that N400 effects occur with masked primes as well (e.g., Deacon, Hewitt, Yang, & Nagata, 2000; Kiefer & Spitzer, 2000; Kiefer, 2002; Holcomb, Reder, Misra, & Grainger, 2005; Grossi, 2006; but see e.g., Brown & Hagoort, 1993) which should be cautiously interpreted as evidence for an automatic mechanism underlying the generation of the N400.

---

[2] There are further ERP components that occur with semantic processing and semantic priming (see e.g., Martin-Loeches, Hinojosa, Gómez-Jarabo, & Rubia, 1999; Kiefer, 2002; Hill, Ott, & Weisbrod, 2005), but for the sake of conciseness, we will focus on the time window of the N400 in our remarks, hypotheses and the result section as well.

Repeated Masked Semantic Priming

Within the field of masked semantic priming research, there is one presentation technique which yields replicable negative priming effects: the repeated masked semantic priming paradigm (Wentura & Frings, 2005). By this presentation technique, the representation and processing of the prime is degraded by repeatedly interrupting the prime presentation by a consonant letter string.

There is one theory from Dagenbach and colleagues (e.g., Carr & Dagenbach, 1990), which could account for negative priming effects with masked primes. The center-surround inhibition theory assumes that the prime (i.e., the center of a category) is only weakly activated due to masking and thus the representation is too weakly activated for direct access. To strengthen the relative activation, the nodes surrounding the center will be inhibited. Thus, the contrast between center and surround is enhanced and there is a relative higher activation of the center which will be an advantage for further access (see Frings, Bermeitinger, & Wentura, in press). If the subsequently presented target falls into the inhibited surround, more activation is needed to get access to this word and responses are slowed down.

To our knowledge, there are no ERP studies yet using versions of the semantic priming paradigm that typically result in negative effects (with behavioral measures). Thus, it is a desideratum to explore whether a (reversed) N400 component will be observed given these conditions.

Method

Participants

The sample consisted of 32 students (20 women) from Saarland University which were paid for their participation. All participants were right-handed and native speakers of German; they had normal or corrected-to-normal vision. The median age was 22 years (ranging from 19 to 32 years). None of the subjects reported any neurological impairment.

Design

To guarantee a high comparability with Wentura and Frings (2005), the same design, material and procedure as in their experiments was used. Essentially, this was a 2 x 2 factorial design. The first factor was prime-target relation (related versus unrelated). The second factor was dominance of the target exemplars (high- versus low-dominance exemplar of its category). Additionally, there was a third factor (target-orthography: word versus non-word), but (as usual) analyses were focused on word trials. For the behavioral data, priming effects were used as the dependent variable; they were computed as the difference between response times for related and unrelated prime-target pairs.

Material

The same material as in Wentura and Frings (2005) was used. The prime set consisted of four category labels – INSEKT (insect), FRUCHT (fruit), VOGEL (bird), and BLUME (flower). Additionally, there were neutral primes which were created with five random letters. The neutral condition was introduced for lowering the overall rate of related prime-target pairs and was not further analyzed. All primes were written in capital letters. Three high-dominance and three low-dominance exemplars of each category served as targets. High-dominance exemplars had a mean association frequency of 67.1 % (*SD* = 10.7 %; range 55 % to 86.5 %), whereas low-dominance exemplars had a mean association frequency of 6.2 % (*SD* = 2.87 %; range 2.5 % to 11.5 %; Mannhaupt, 1983). Mean length was 5.2 (*SD* = 0.8; range 4 to 7) for the high-dominance exemplars and 5.4 (*SD* = 0.5; range 5 and 6) for the low-dominance exemplars. The average frequency was 5318 (*SD* = 10026) for high-dominance exemplars and 502 (*SD* = 727) for low-dominance exemplars (according to the German database of written language, COSMAS II). Pronounceable non-words were created by changing one letter of each target word. All stimuli were presented in light gray (about 41 cd/m²) on a black screen and were approximately 0.5 cm (0.48°) in height. Primes and masks were approximately 3.0 cm (2.86°) in width; targets were between 1.4 cm (1.34°) and 2.6 cm (2.48°) in width. All stimuli were written in the font Fixedsys.

Procedure

Participants were individually tested in a sound attenuated chamber. The experiment was run using the E-Prime software (version 1.1) with a standard PC and a 17'' CRT monitor. Instructions were given on the CRT screen from which the participants had a distance of approximately 60 cm. Participants were told that words belonging to the categories insect, fruit, bird, or flower would be presented on the screen. Some of these words, however, would be written with a spelling mistake. They were requested to quickly and correctly categorize each word with regard to orthography by pressing either the right key with their right index finger for correctly written words or the left key with their left index finger for misspelled words.

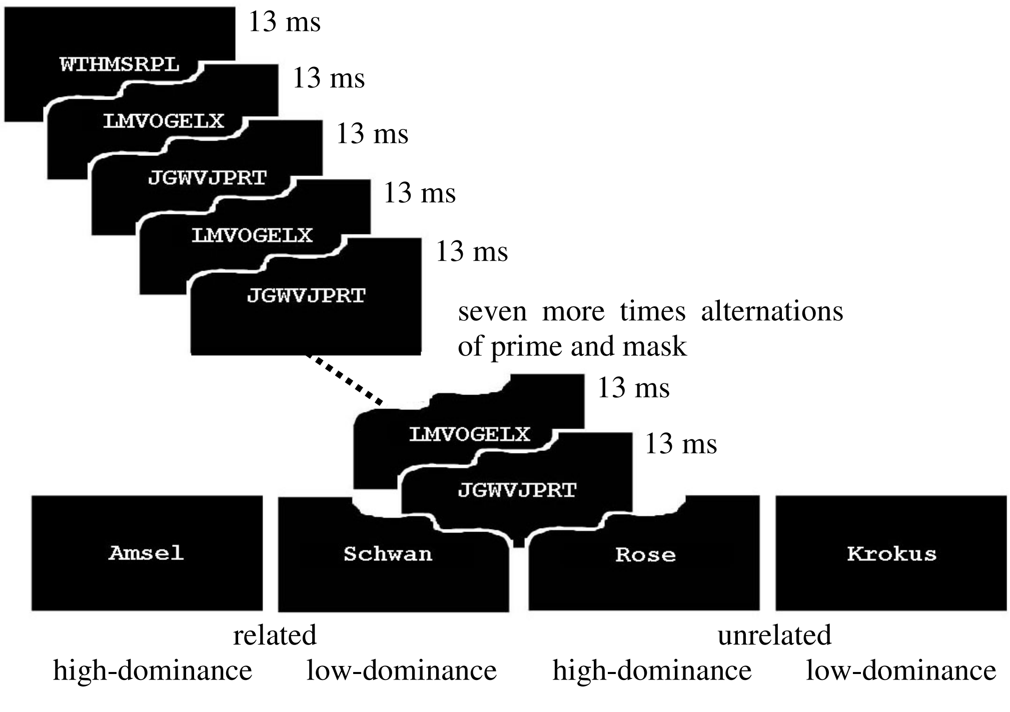


Figure 1. Procedure of one trial with examples of different prime-target relations. The target remained on the screen until a response was given. See text for further explanations..

The sequence of each trial (see also Figure 1) was as follows (see Wentura & Frings, 2005): First a fixation stimulus (+) appeared at the centre of the screen for 500 ms. It was followed by a forward mask, consisting of eight randomly generated capital consonants, which was presented for one refresh cycle (i.e., about 13 ms). The forward mask was immediately overwritten by the prime which was on the screen for the next refresh cycle. The related prime was always the category name that corresponded to the target. The unrelated prime was always INSEKT (insect) for fruit exemplars, FRUCHT (fruit) for insect exemplars, VOGEL (bird) for flower exemplars, and BLUME (flower) for bird exemplars. Random capital consonants were added to the left and to the right of the prime to create a string of eight letters. The prime was followed by a backward mask consisting of eight random consonants for one refresh cycle. Prime and backward mask were alternately presented for a total of 20 refresh cycles. Thus, the total prime duration was 133 ms. After the last backward mask, the target appeared and remained on the screen until a response was given. In the case of a false response, an error message was given on the screen until a keypress. The intertrial interval with a blank screen was 1000 ms.

The experiment comprised three blocks with 48 trials (16 related, 16 unrelated, and 16 neutral prime-target pairs; half of the trials with nonword targets). Over the course of the experiment, each target appeared in each of the three priming conditions. Within a block, each target was presented in one of the three priming conditions. The sequence of priming conditions for a given target was determined by a Latin-square design (i.e., sequence of targets and conditions was balanced over participants). Participants could make a rest after every 24 trials. Before the experimental trials, participants practiced the task with 48 practice trials.

Additionally, there were further trials which tested the prime discrimination performance of the participants.[3]

EEG Recording and Analyses

EEG activity was recorded continuously from 64 Ag/AgCl electrodes mounted in a preconfigured cap (Electro-Cap International, Inc., Eaton, OH), arranged according to the extended international 10-20 system (American Electroencephalographic Society, 1994) with a sampling rate of 250 Hz. Impedances for all electrodes were kept below 10 kΩ. Signals were referenced on-line to the left mastoid electrode. For further analysis, electrodes were re-referenced off-line to linked mastoids. Two electrodes located medially to the right eye, one above and one below, were used to monitor vertical eye movements. Electrodes placed at the outer canthi of the eyes measured horizontal eye movements. Vertical and horizontal ocular artifacts were monitored and corrected off-line (Gratton, Coles, & Donchin, 1983).

Data were digitally filtered with Butterworth Zero Phase Filters (low cutoff: 0.1 Hz; high cutoff: 30 Hz). ERPs were obtained by averaging EEG recordings time-locked to target presentation from 400 ms prior to 1500 ms after target onset. Trials with false responses or with reaction times below 200 ms or above 1500 ms or the individual Tukey criterion (see below) or were rejected. Trials containing artifacts (maximum amplitude in the recording epoch ± 100 µV; maximum difference between two sampling points 40 µV; maximum difference between any two sampling points within an epoch 150 µV) and individual channels with larger than 100 µV differences between any two sampling points within an epoch were rejected as well; with respect to this artifact correction, the average number of trials for a specific condition was 10.19 (ranging from 9.59 to 10.84 for conditions and ranging from 7.00 to 11.33 for participants) . Data were baseline-corrected with respect to the 200 ms pre-target interval (there was no significant difference between related and unrelated conditions in this time window for all electrodes, all *p*s > .05). Furthermore, ERPs were averaged for the related and unrelated condition separately for high- and low-dominance targets, respectively.

---

[3] We do not advance the view that our presentation technique realizes objective subliminality of the primes (see e.g., the high rates of participants with a non-random prime discrimination performance in the studies by Wentura & Frings, 2005, and Frings et al., in press). In the present study, performance of participants in the direct test of prime discrimination does not moderate the results. Therefore, we do not report the exact procedure and the results of the direct test.

Statistical analyses were performed by means of an ANOVA on mean voltages for each condition in the time-window between 350 and 550 ms (the choice of the exact position and expansion of the time window was based upon visual inspection). Mean amplitudes were computed for each of the following eight regions of interest (ROIs): Left-Frontal (Fp1, AF3, F7, F5, F3, F1), Right-Frontal (Fp2, AF4, F8, F6, F4, F2), Left-Central (FC3, FC1, C3, C1, CP3, CP1), Right-Central (FC4, FC2, C4, C2, CP4, CP2), Left-Parietal (P7, P5, P3, P1, PO7, PO3, O1), Right-Parietal (P8, P6, P4, P2, PO4, PO8, O2).

Results

Unless otherwise noted, all effects referred to as statistically significant throughout the text are associated with p-values of less than .05, two-tailed. If necessary, the Greenhouse-Geisser correction for nonsphericity was used (Greenhouse & Geisser, 1959); original degrees of freedom, the correction coefficient ε and corrected p-values are reported in the following. When performing post-hoc tests, significance levels were adjusted according to the Bonferroni procedures (and all p-values that are given in the corresponding result paragraphs are significant with respect to adjusted α-levels).

Behavioral Priming Effects

Mean RTs were derived from correct responses to word trials with word primes. The mean error rates to these trials were 3.78 %. RTs that were below 200 ms, above 1500 ms, or were 1.5 interquartile ranges above the third quartile with respect to the individual distribution (Tukey, 1977) were discarded. These criteria led to the exclusion of 4.43 % word trials with word primes. Mean RTs, mean error rates, and mean priming differences for word targets are shown in Table 1.

Table 1

Mean Response Time (in ms) of Word Trials as a Function of Prime-Target Relation, and Dominance of the Target Exemplars (Errors in % in Parentheses) and Priming Effects (Unrelated – Related, in ms). Slight differences are due to rounding errors.

	Related	Unrelated	Priming effect
Low-dominance	708	692	-15
	(5.7)	(6.0)	(0.3)
High-dominance	648	640	-7
	(1.8)	(1.6)	(-0.3)

Mean reaction times of word targets were subjected to a 2 (prime-target relationship: related vs. unrelated) x 2 (target dominance: high vs. low) ANOVA. There was a significant main effect of prime-target relationship, $F(1,30) = 4.42$, $MSE = 929$, $p < .05$. On average participants showed $M = 11$ ms ($SD = 30$ ms) slower reactions to related targets than to unrelated targets. Additionally, there was a main effect of target dominance, $F(1,30) = 57.12$, $MSE = 1771$, $p < .001$, indicating faster reactions to high-dominance than to low-dominance targets. The interaction effect was not significant ($F < 1$).
The same ANOVA on error rates revealed only a significant main effect of target dominance, $F(1,30) = 12.22$, $MSE = 46$, $p < .001$, with more accurate reactions to high-dominance than to low-dominance targets; all other effects were not significant (all Fs < 1.03, ps > .31).

Electrophysiological Results

Mean amplitudes of the N400 time window (between 350 and 550 ms) were subjected to a 2 (hemisphere: left vs. right) x 3 (anterior/posterior: anterior, central, posterior) x 2 (prime-target relationship: related vs. unrelated) x 2 (target dominance: high vs. low) ANOVA. There was a significant main effect of the hemisphere factor, $F(1,31) = 11.08$, $MSE = 8.92$, $p < .01$, indicating higher overall right hemisphere amplitudes than left hemisphere amplitudes, and a significant anterior/posterior factor, $F(2,62) = 10.87$, $MSE = 66.98$, $p < .001$, $\varepsilon = .794$, indicating higher overall amplitudes with increasing posteriority. Furthermore, the interaction hemisphere by anterior/posterior was significant, $F(2,62) = 7.61$, $MSE = 3.10$, $p < .01$, $\varepsilon = .612$. However, the most interesting effect is the significant interaction effect between prime-target relationship and the anterior/posterior factor (see also Figure 2), $F(2,62) = 5.52$, $MSE = 3.06$, $p < .05$, $\varepsilon = .663$. Additionally, the main effect of target dominance missed the significance criterion, $F(1,31) = 3.17$, $MSE = 22.93$, $p = .085$. Low-dominance targets showed a trend to a more negative going deflection than high-dominance targets. No other main or interaction effects were significant (all Fs < 2.12, all ps > .15).

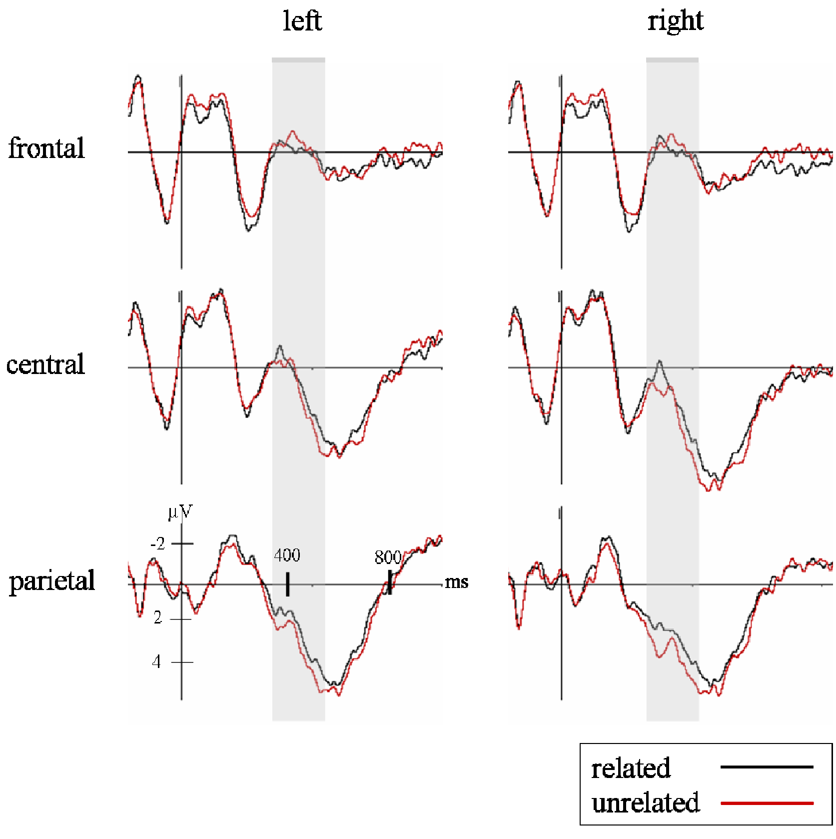


Figure 2. Grand-averages of the eight pools obtained from all subjects: comparison between related (black line) and unrelated (red line) targets. The gray bars indicate the time window from 350 to 550 ms.

Post-hoc tests revealed that ERPs to related targets were significantly more negative than to unrelated targets only at the posterior ROIs, $t(31) = 3.02$, $SE = 0.28$, $p = .005$. There are no significant differences between related and unrelated ERPs at anterior ($t(31) = 1.67$, $SE = .38$, $p = .11$) and central ROIs ($t(31) = 0.30$, $SE = .41$, $p = .76$).

Discussion

Using the repeated masked semantic priming paradigm, we replicated the behavioral negative semantic priming effect first found by Wentura and Frings (2005). However, in the original study, there was a moderation of the priming effect by dominance of the target. The negatively signed effect most markedly occurred for low-dominance targets. In the present study, this moderation is not evident (with regard to inferential statistics). Participants showed increased lexical decision response times to category exemplars preceded by their category labels independently of targets' dominance. This pattern was reported in Frings et al. (in press) as well. Obviously, for explaining the original data (Wentura & Frings, 2005) with a center-surround theory, there was some arbitrariness in the assumptions of what a category's center and what a category's surround is. Eventually, there are some interindividual differences which influenced whether an exemplar falls into the center or into the surround of a category or whether the center of a category simply consists of the category label. Of course, this *post-hoc* consideration is open to future research.

As a new aspect, additionally, we recorded event-related potentials and focused on the N400, which is the most prominent component for semantic processing. First, there was an expectable trend of targets' dominance (e.g., Stuss, Picton, & Cerri, 1988; Fujihara, Nageishi, Koyama, & Nakajima, 1998; Heinze, Muente, & Kutas, 1998; Núñez-Peña & Honrubia-Serrano, 2005). Low-dominance targets showed a more negative going wave than high-dominance targets. That finding goes with the behavioral data with faster and more accurate responses to high-dominance than to low-dominance targets. This seems to reflect an overall advantage in the processing of high-dominance words independently of the context which could be caused by the higher frequency of these words and the better accessibility by reason of this higher frequency.

However, most interesting, we found a pattern which we would call – as a first attempt – a reversed N400 effect, i.e. a more negative going wave for related than for unrelated targets between 350 and 550 ms. This corresponds perfectly to our behavioral results with a negatively signed semantic priming effect. One could interpret this as further evidence for an uncontrolled, automatic modulation of the N400 because it is very implausible that our participants had processed the masked primes on a level for strategic use (e.g., Kiefer, 2002; but see e.g., Holcomb et al., 2005). Furthermore, in this context, the reversed N400 does not reflect the *a priori* (mis)fit between a category label and its exemplars but seems to reflect a locally and temporarily reduced access to related targets caused by the interrupted prime presentation. Here, the N400 rather seems to reflect the difficulty to activate a word (e.g., by abolishing inhibition) than the difficulty to integrate a word in the foregoing semantic context. Thus, the N400 is not only modulated by the fit with the context *per se* or by word inherent properties but by special presentation techniques which modify the representation or processing of the prime as well. But with an extended comprehension of 'the context', whereby the context consists not only of the semantic content but also of perceptual features, the results can be integrated perfectly in further accounts of the N400 where it was considered an "index of the degree of semantic priming or activation that a word receives from the prior context" (Kutas, Lindamood, & Hillyard, 1984, p. 216).

As an alternative explanation, the found pattern can be interpreted in terms of the P300. The P300 is a positive going deflection starting about 250 ms after stimulus onset and is most distinct on parietal electrodes. Within the ERP trace, the P300 is the most prominent and most studied component which varies with different parameters (for reviews see e.g., Coles & Rugg, 1995; Picton, Lins & Scherg, 1995; Hruby & Marsalek, 2003; Polich & Criado, 2006). For example, the P300 is more positive for a high response certainty compared to low certainty. Related to this, the P300 correlates with the difficulty to discriminate two events/stimuli; there are higher amplitudes for lower discrimination difficulties (e.g., Andreassi, 2000). Altogether, the P300 was interpreted as a mechanism which stops a decision process (e.g., Desmedt, 1980; Rockstroh, Mueller, Heinz, Wagner, Berg & Elbert, 1996) or updates the memory representation for the stimulus context (e.g., Polich & Criado, 2006). The amplitude of the P300 is larger when an enhanced processing was required, e.g., in the case of stimuli which are more meaningful, less probable, or more disturbing (e.g., Johnson, 1993).

The found ERP trace shows a distinctive P300 and the difference between related and unrelated targets can be interpreted in terms of discrimination difficulty. There was a higher amplitude for unrelated targets which mirrors lower difficulty compared to the related targets with a lower amplitude. One could have expected that it would be more difficult to react to related than to unrelated targets, e.g., because unrelated targets are less probable in the context of the prime. Why should it be easier to process the unrelated targets? One answer could be that, with our masking technique, only a weak representation of the prime can be created. And this weak representation directly passes into the presentation of the target. In the case of related prime target pairs, it seems relatively difficult to separate and discriminate the target representation from the prime representation due to partially the same or overlapping activation patterns. Compared to a clearly visible presentation, where prime and target were easily separable, our presentation technique hampers this discrimination. In the case of unrelated prime target pairs, there are no overlapping activations and it was lower effort required to discriminate both representations. However, enhanced processing was required to segregate prime and target in the case of related pairs. Further research (e.g., with a systematic variation of the time-lag between last mask and target and a clearly visible control condition) was needed to confirm these interpretations.

In summary, the behavioral data obtained in the present study showed a negative semantic priming effect which comes along with a greater positivity for unrelated than for related targets in the time window from 350 to 550 ms. This can be interpreted in terms of a reversed N400 effect or in terms of the P300. Regarding the N400 interpretation, the center-surround inhibition theory can account for these effects by assuming that by the special presentation technique the category label is surrounded with a ring of inhibition. For the following activation of an inhibited exemplar, more activation is needed, which was reflected in a more negative going N400 wave and slower reaction times. Regarding the P300 interpretation, the pattern might reflect a larger difficulty to discriminate the representations of prime and target for related than unrelated prime target pairs. Yet, further research is needed for replicating these new findings and for obtaining the negative priming effect and the accompanying ERP pattern with other presentation techniques.

References

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1983a). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*, 261-295.

Anderson, J. R. (1983b). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Andreassi, J. L. (2000). *Psychophysiology*. New Jersey: Erlbaum.

Bentin, S., McCarthy, G., Wood, C.C. (1985). Event-related potentials associated with semantic priming. *Electroencephalography & Clinical Neurophysiology*, *60*, 343– 355.

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

Brown, C. & Hagoort, P. (2000). On the electrophysiology of language comprehension: Implications for the human language system. In M. W. Crocker, Pickering, M., & C. Clifton (Eds.), *Architectures and mechanisms for language processing* (pp. 213-237). New York, NY, US: Cambridge University Press.

Brown, C. M., Hagoort, P. & Chwilla, D. J. (2000). An event-related brain potential analysis of visual word priming effects. *Brain and Language*, *73*, 158–190.

Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 341-350.

Coles, M. G. H., & Rugg, M. D. (1995). Event-related brain potentials : An introduction. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of Mind* (pp. 1-26). Oxford, U.K.: Oxford University Press.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407-428.

Dagenbach, D., Carr, T. H., & Barnhardt, T. M. (1990). Inhibitory semantic priming of lexical decisions due to failure to retrieve weakly activated codes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 328-340.

Deacon, D., Hewitt, S., Yang, C.-M., & Nagata, M. (2000). Event-related potential indices of semantic priming using masked and unmasked words: Evidence that the N400 does not reflect a post-lexical process. *Cognitive Brain Research*, *9*, 137-146.

Deacon, D., Uhm, T.-J., Ritter, W., Hewitt, S., & Dynowska, A. (1999). The lifetime of automatic semantic priming effects may exceed two seconds. *Cognitive Brain Research*, *7*, 465-472.

Desmedt, J. E. (1980). P300 in serial tasks: An essential post-decision closure mechanism. *Progress in Brain Research*, *54*, 682-686.

Dosher, B. A., & Rosedale, G. (1989). Integrated retrieval cues as a mechanism for priming in retrieval from memory. *Journal of Experimental Psychology: General*, *118*, 191-211.

Durante, R. & Hirshman, E. (1994). Retrospective priming and masked semantic priming: The interfering effects of prime activation. *Journal of Memory and Language*, *33*, 112-127.

Fowler, C. A., Wolford, G., Slade, R., & Tassinary, L. (1981). Lexical access with and without awareness. *Journal of Experimental Psychology: General*, *110*, 341-362.

Frings, C., Bermeitinger, C., & Wentura, D. (in press). Center-surround or spreading inhibition: Which mechanism caused the negative priming effect from repeated masked semantic primes? *Experimental psychology.*

Fujihara, N., Nageishi, Y., Koyama, S., & Nakajima, Y. (1998). Electrophysiological evidence for the typicality effect of human cognitive categorization. *International Journal of Psychophysiology*, *29*, 65-75b.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artefact. *Electroencephalography & Clinical Neurophysiology*, *55*, 468-484.

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

Grossi, G. (2006). Relatedness proportion effects on masked associative priming: An ERP study. *Psychophysiology*, *43*, 21–30.

Heinze, H.-J., Muente, T.-F., & Kutas, M. (1998). Context effects in a category verification task as assessed by event-related brain potential (ERP) measures. *Biological Psychology*, *47*, 121-135.

Hill, H., Ott, F., & Weisbrod, M. (2005). SOA-dependent N400 and P300 semantic priming effects using pseudoword primes and a delayed lexical decision. *International Journal of Psychophysiology*, *56*, 209–221.

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic priming. *Brain and Language*, *35*, 66–85.

Holcomb, P. J., Reder, L., Misra, M., & Grainger, J. (2005). The effects of prime visibility on ERP measures of masked priming. *Cognitive Brain Research*, *24*, 155-172.

Hruby, T., & Marsalek, P. (2003). Event-related potentials – the P3 wave. *Acta Neurobiologiae Experimentalis*, *63*, 55-63.

Johnson, R. (1993). On the neural generators of the P300 component of the event-related potential. Psychophysiology, 30, 90-97.

Kahan, T. A. (2000). Negative priming from masked words: Retrospective prime clarification or center-surround inhibition? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1392-1410.

Kiefer, M. (2002). The N400 is modulated by unconsciously perceiced masked words: Further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research*, *13*, 27–39.

Kiefer, M. & Spitzer, M. (2000). Time course of conscious and unconscious semantic brain activation. *Neuroreport*, *11*, 2401-2407.

Kutas, M. & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*, 463-470.

Kutas, M. & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*, 99–116.

Kutas, M. & Van Petten, C. K. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83-143). San Diego, CA, US: Academic Press.

Kutas, M., Lindamood, T., & Hillyard, S. A. (1984). Word expectancy and event-related brain potentials during sentence processing. In S. Kornblum, & J. Roquin (Eds.), *Preparatory States and Processes* (pp. 217-238). Hillsdale, NJ: Erlbaum.

Mannhaupt, H.-R. (1983). Produktionsnormen für verbale Reaktionen zu 40 geläufigen Kategorien [German category norms for verbal items in 40 categories]. *Sprache und Kognition, 4*, 264-278.

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology*, *15*, 197-237.

Martín-Loeches, M., Hinojosa, J. A., Gómez-Jarabo, G., & Rubia, F. J. (1999). The recognition potential: An ERP index of lexical access. *Brain and Language*, *70*, 364–384.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *21*, 3-23.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York, NQ: Psychology Press.

McNamara, T. P., & Holbrook, J. B. (2003). Semantic memory and priming. In A. F. Healy & R. W. Proctor (Eds.), *Experimental Psychology* (pp. 447-474). Vol. 4 in I. B. Weiner (editor-in-chief), *Handbook of psychology*. New York: Wiley.

McRae, K., de Sa, V. R., Seidenberg, M. S., & (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*, 99-130.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, *89*, 609-626.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ: Erlbaum.

Núñez-Peña, M. I., & Honrubia-Serrano, M. L. (2005). N400 and category exemplar associative strength. *International Journal of Psychophysiology*, *56*, 45-54.

Picton, T. W., Lins, O. G., & Scherg, M. (1995). The recording and analysis of event-related potentials. In F. Boller & J. Grafmann (Eds.), Handbook of Neuropsychology (pp. 3-73). Amsterdam: Elsevier Science B.V.

Polich, J., & Criado, J. R. (2006). Neuropsychology and neuropharmacology of P3a and P3b. *International Journal of Psychophysiology*, *60*, 172-185.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*, 385-408.

Rockstroh, B., Mueller, M., Heinz, A., Wagner, M., Berg, P., & Elbert, T. (1996). Modulation of auditory responses during oddball tasks. *Biological Psychology*, *43*, 41-55.

Ruz, M., Madrid, E., Lupiáñez, J., & Tudela, P. (2003). High density ERP indices of conscious and unconscious semantic priming. *Cognitive Brain Research*, *17*, 719–731.

Spence, D. P. (1983). Subliminal effects on lexical decision time. *Archiv für Psychologie*, *135*, 67-72.

Stuss, D. T., Picton, T. W., & Cerri, A. M. (1988). Electrophysiological manifestations of typicality judgment. *Brain and Language*, *33*, 260-272.

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

Wentura, D., & Frings, C. (2005). Repeated masked category primes interfere with related exemplars: New evidence for negative semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 108-120.

Space, time, and the use of language: An investigation of relationships

THORA TENBRINK

tenbrink@sfbtr8.uni-bremen.de; SFB/TR 8 Spatial Cognition, University of Bremen, Germany

The investigation of linguistic expressions of space and time typically presupposes an interdependency between these two related conceptual domains, resulting in a quasi-metaphorical relationship in which time is secondary to space. The present work approaches this general idea from a range of perspectives, focusing on the question whether the interrelationship has any concrete consequences on the usage of spatial and temporal dimensional terms such as *before, after* or *left, right, in front, above*. The conclusion of this investigation is that no interdependency can be traced in the application of these terms, as they differ widely in the contextual environments and patterns in which they occur in natural discourse. Also, temporal terms do not seem to be secondary to spatial terms as far as usage is concerned.

Introduction

There is a popular and widespread assumption that temporal terms are conceptually (as well as historically) based on, spatial terms (e.g., Haspelmath, 1997). The present work investigates whether evidence for the assumed conceptual dependency can be identified with respect to application. Specifically, it deals with the German and English spatial dimensional terms *(in) front / vor(ne), behind / back / hinter, left / links,* and *right / rechts,* and their temporal counterparts *before / bevor,* and *after / nachdem,* including their various syntactic variants. These terms are used to express relationships on either a spatial or a temporal dimension.

One of the reasons for the presumption that temporal expressions depend on spatial ones is the idea that the spatial domain is concrete (i.e., perceptually accessible) while the temporal domain is abstract (i.e., less easy to perceive and to grasp). This is reflected by the notion that the entities associated with space are (concrete) objects, while the entities associated with time are (abstract) events. Now, if concrete entities are easier to handle than abstract ones, it is natural to extrapolate from the experience gained in the concrete domain in order to cope with abstract experience. Thus, events are treated in some respects in a similar way as objects, and the underlying domain of time is understood in terms of experience gained from the more accessible domain of space.

However, this very insight also provides the basis for the present assumption that, when it comes to the employment of these terms in natural discourse, fundamentally different patterns arise when comparing the localisation of objects with that of events. Spatial terms are employed in concrete settings, while temporal terms apply to the representation of abstract events. The capacity for transfer should be limited as far as application is concerned, since events are ontologically fundamentally different from objects. Therefore, the conditions for usage should be distinct in some basic and systematic respects, although parallels may well be identifiable. The present work aims to shed light on the patterns of discourse applicability for superficially (i.e., in terms of morphosyntax) similar, and semantically related, spatiotemporal expressions.

This contribution is based on a dissertation thesis (Tenbrink 2006), which has recently been published as a monograph in an abbreviated version (Tenbrink 2007).

Conceptual Background

The idea that the human concept of time is based on the concept of space is not new. For centuries, it has been in some way or other assumed that conceptions and representations of time depend on space, that time is measured by means of space, and that the past is conceived of as behind us or behind the present moment, while the future is in front. Psychological research findings support these notions, such as the fact that children mix up spatial and temporal concepts especially in cases where both are involved, as in the filling up of containers with liquid, or in tasks involving movement (Piaget, 1955).

In the field of linguistics, Clark (1973) first systematically motivated the idea that temporal language is based metaphorically on spatial language, and that many relational temporal prepositions in English are historically derived from *front* and *back*. To describe the spatial metaphor that, according to Clark, can explain all temporal concepts, he put forward two opposing but compatible notions: in the "moving time" metaphor, time is viewed as moving past us from future to past; and in the "moving ego" metaphor, we as humans are viewed as moving through time from past to future (see also Miller & Johnson-Laird, 1976:463); an example for this would be "I look forward to Monday".

However, the presence of metaphorical expressions in language does not prove the existence of an underlying pervasive conceptual metaphor that is responsible for all representations of one domain. Metaphors are fairly frequent in language, and they are often impressively consistent (Lakoff & Johnson, 1980), but this does not imply that the concepts that are represented in metaphorical language could not also be represented in an independent way. The argument holds only in the opposite direction: Metaphorical usage suggests compatibility with a different domain. Metaphors are used for highlighting similarities, but they do not embrace the whole target concept, nor does their usage imply that all aspects of the source concept can be translated to the target concept.

In an alternative view (Habel & Eschenbach, 1997), space is not seen as a concrete source domain from which the more abstract concepts of time are consistently derived. Instead, space and time share a range of representational structures, which are systematically reflected in language. But the two domains are also sufficiently different to allow for an independent conceptual representation for each of them.

Generally, the similarities between spatial and temporal concepts can be explained by an underlying metaphorical relationship only to a certain degree. While the metaphorical approach highlights many important aspects of the relationship between space and time, it neglects and obscures others. Most importantly, it does not account for the existing differences between temporal and spatial representations, both in language and other modes of representation. But in addition, it imposes a primacy on one of the two domains in order to account for the pervasive similarities. In my approach I adopt a more neutral view upon both similarities and differences with regard to spatial and temporal concepts, avoiding a bias towards a presupposed superiority of the spatial domain, or a presupposed necessity to explain all human temporal conceptions on the basis of spatial ones.

Many conceptual issues cross over the domains of space and time. For example, in both time and space, it is equally conceivable to consider the abstract domains as independent of entities that fill them, such as objects and events, or to view them as *constituted* by objects and events. Also, both temporal and spatial domains exhibit some kind of *neighborhood* structure (reflected in language) in which some entities are conceived of as closer or further apart from each other. Since such conceptualizations are relevant in both domains to similar degrees, there is no reason to assume an underlying metaphorical relationship or a dependency of time concepts on space concepts. On the contrary, in some cases similar issues are dealt with in considerably different ways. Altogether, the shared issues may well be sufficient to explain the many similarities between linguistic expressions in both domains (see Tenbrink, 2007).

But in addition to that, there are substantial differences between the conceptual domains that are reflected in language – perhaps most crucially so with respect to application. Here, one example is selected to illustrate this general idea. The typical way of referring to an object is by using a noun, while events (or processes) are usually referred to as verbs. But while there are usually no difficulties in turning verbs into nouns by nominalisation, the opposite is rare: While it is sometimes possible to use a noun like *butter* in verbal guise as a short form for the process of putting butter on bread (as in *buttering the toast*), the objects themselves cannot be expressed in verbal form. This general fact about language is directly reflected in the possible syntactic functions of the expressions under analysis. Spatial relational terms typically (i.e., with very few exceptions) cannot relate clauses to each other. In contrast, in direct correspondence to the fact that events can be expressed by verbs as well as nouns, temporal relational expressions can occur with clauses (comprising a process denoting an event) as well as with noun phrases (as long as the nouns refer to nominalised events, or else, specific metric time spans). This difference between spatial and temporal markers in their ability to appear as conjunctions is viewed as crucial with respect to the comparison between the domains of space and time. Importantly, there is a vast literature on discourse connectives (such as temporal conjunctions) pointing to the ways in which they contribute to the coherence of the discourse. Apart from temporality, (specific) connectives are capable of expressing *causality*, *elaboration*, *enablement*, etc. But while there is remarkable agreement about the fact that temporal effects are central in discourse and pervasive in language, there are no such results on spatial effects. Although all discourse is situated in space just as well as it is in time, spatial phenomena are conspicuously absent from the literature on discourse markers as a whole.

In the following, I will take a closer look at temporal as well as spatial dimensional terms in turn, summarizing what is known in the literature concerning their semantics and application, complemented by my own empirical results.

Temporal language

Language can be viewed as a central means of capturing and communicating abstract features of time. In this respect, the linguistic means available to speakers, as well as the way language is used in natural communication, are interesting as a research field: they reflect speakers' underlying conceptions of the relations between events, which are generally not purely temporal but are also perceived as connected in some more or less direct way.

The most uncontroversial aspect of the semantics of temporal dimensional terms, both in German and in English, is that they express a *relation in time*, i.e., either temporal anteriority or posteriority. However, they often do not simply express a general temporal precedence relation, but there is often an association of sequentiality or proximality. Then, a contextually dependent *time frame* is necessary to account for the semantics of such terms (Herweg, 1991).

The events described in two adjacent clauses are often conceptualised as *causally related* even in the absence of explicit causal markers. This phenomenon is explained in part by the intricate relationship between the two abstract concepts of time and causality. However, this does not imply that causality is ubiquitous, which would mean that a causal relationship can be inferred between all events related in sequence in a stretch of text. Causality is often specifically associated with usage of *after* (e.g., Heinämäki, 1974), as in:

(1) After he stumbled over a stone, he fell.

Here, the inference that the falling event was caused by the stumbling event is a natural conclusion to draw, even though it can easily be cancelled. With *before*, the conclusion gets much harder to draw, although it still seems to be possible at least for some speakers. Therefore, the association of causal relationships seems to differ between the temporal dimensional terms, and it also seems to be a matter of degree rather than a directly licensed inference. However, there are a number of further ways in which causal or quasi-causal discourse relations can be derived, both with *after* and with *before* (e.g., Schilder, 2001). Further well-discussed aspects of temporal dimensional terms concern the association with presuppositional effects of different kinds (e.g., Miller & Johnson-Laird, 1976, Lascarides & Oberlander, 1993), as well as the possibility of non-veridical interpretations together with *before* (Heinä-mäki, 1974).

To gain further insights concerning the applicability conditions of temporal dimensional terms in English and German, I investigated naturally occurring instances of temporal dimensional terms taken from various corpora, employing a pragmatic perspective. The results revealed a fairly consistent pattern, both across the two closely related languages and across text types. Speakers typically do not simply use the terms to denote the temporal relationship of unrelated events; rather, the terms support the inference of an underlying semantic relationship between the events. Additionally, the temporal relationship of known events (whatever their relation may be), or the length of the time span, might be focused upon. These results accord with the structural resources of language: The English and German languages offer many options for expressing temporal relations, such as tense and aspect, temporal adverbials, or implicitly conveying the information via clause order. Thus, if temporal dimensional terms are employed, speakers may intend to convey more than simply temporal information. Here, one simple reason for employing an explicit relational term is that the temporal order is not just conveyed but is in focus. However, there are also a number of possible conceptual relationships between the events that are linguistically related using a temporal dimensional term. Many of these are causal in some sense. This fact reflects previous findings in Cognitive Science on the relationship between time and causality: since we, as humans, experience time as a network of causally interrelated events, it is only natural that our linguistic manifestations of temporal relations are closely related to notions of causality.

However, the employment of temporal terms while implying causal relations seems to entail that the speaker wishes to leave some amount of freedom to the listener's interpretation: otherwise, they could have employed explicit causal relational terms such as *because*. This observation is in accord with the fact that speakers (in the corpora used) sometimes explicitly mention their uncertainty about the nature of the interrelationship between two events. In addition, speakers can always intend to suggest even more than what is received by the recipient (or the analyst). Some semantic relationships are supported by the application of the temporal terms in some kinds of contexts, and others are discouraged. Typically, more content is intended by – and interpreted into – an utterance than merely a temporal relationship, which is the most obvious and uncontroversial aspect of the semantics of temporal dimensional terms.

Since speakers typically do not relate two previously unknown events linguistically, the empirical investigation also addressed the question of how speakers deal with different degrees of Givenness in discourse. The two research directions are interrelated in that specific kinds of conceptual categories pose different degrees of restrictions on Givenness. In general, it can be concluded that the information conveyed in the temporal component is mostly known or inferable from the context or discourse external knowledge; but this is not a strict requirement for the application of temporal dimensional terms.

Spatial language

Locative dimensional terms denote a spatial relation between two objects. One object serves as relatum, and the other is positioned within a region surrounding a focal axis with respect to the relatum, based on the conceptualisation of a reference system. The size of the region depends on contextual factors but is at all times limited to a half plane. With unmodified dimensional terms the most likely position is on the focal axis itself; with increasing distance from the axis applicability decreases.

The area in which spatial dimensional terms have probably been investigated most intensively concerns reference systems and perspectives. It is well-known that the same spatial term can often be interpreted in many ways, sometimes leading to contradictory results. Various different classifications and models have been proposed, and the literature offers a variety of overviews and discussions of advantages and shortcomings of specific views (e.g., Levinson, 2003). Levinson elaborately describes the confusions arising with different and conflicting terminology, including various interpretations of terms like *deictic*, *extrinsic*, and *intrinsic*. Importantly, deixis is often confused with perspective, since both are based on the actual situation. I start from Levinson's account, which I extend with further insights.

Objects can be related either externally or internally; in the latter case one object is located inside another object. Levinson's account focuses on external relationships. It can be characterised schematically as recognising three different reference systems with three variations each, dependent on whether the speaker, the hearer, or a third entity serves as the origin of the perspective employed. The three different options for reference systems are labeled by Levinson as *intrinsic*, *relative*, and *absolute*. Relative reference systems use three different positions, namely, that of the located object, that of another object (the *relatum*) which it is related to (or sometimes a group of objects), and that of the perspective used. In intrinsic reference systems, the latter two positions are conflated. Unlike these two kinds of systems, absolute reference systems do not employ the dimensional terms that are in focus here, and are therefore not considered further.

Concerning the question which of these reference systems may be preferred by users, results in the literature are controversial. Very likely, features of the discourse task as well as cultural factors greatly influence speakers' choices. Much earlier research has focussed not on the distinction between conceptual reference systems, but addressed the specific perspective speakers used: typically, either their own or that of the addressee. According to Herrmann & Grabowski (1994), speakers employ the listener's point of view specifically if there are reasons for this. Obviously, processing is easier for the listener if their own perspective has been used; thus, if the speaker wishes to facilitate interpretation for the interlocutor, they use the interlocutor's viewpoint. This is especially likely if the partner is expected to have less cognitive or linguistic capabilities, e.g., because of younger age or because someone is not a native speaker of the language employed. But it can also be the case in situations where the speaker wants to be polite, e.g., if the interlocutor has a higher status. Furthermore, if actions are involved on the part of the listener, usually the listener's perspective will be used.

Objects in motion (or potentially in motion) induce a further kind of perspective that is independent of intrinsic fronts or perceptual organs. The perspective adopted can be described as though the moving object was viewed from the inside, so to speak, looking in the direction of motion. Thus, even completely symmetric objects, such as a ball, can be ascribed *front*, *back*, *right* and *left* sides when in motion:

(2) John is running to the right of the ball that is rolling down the hill.

This example can also be interpreted with a relative reference system using an external viewpoint, so that the utterance is potentially ambiguous.

Furthermore, the assignment of sides in the case of motion holds for scenarios in which the moving entity itself serves as the origin of a relative reference system. In this case the perceptual apparatus does not play a role, but rather the direction of motion, as in:

(3) If the ball continues rolling down the hill, it will fall into the hole in front of the wall.

Generally, any oriented object can serve as the origin of both intrinsic and relative reference systems; the orientation can come about by (potential) motion or by intrinsic features such as perception. Furthermore, similar effects arise by functional ordering relations such as those induced by a queue.

For the application of intrinsic (external) reference systems, the relatum needs to have internal properties, such as a front. For the description of these object parts, the same dimensional expressions (*right, left, front, back* side) are employed as for relationships between objects. Also, other objects can be located inside of other objects and then be described on the basis of the relatum's internal parts:

(4) The table is standing in the back of the room.

A typical English surface form for such a relationship is a noun in a prepositional phrase using *in*, while in German, an adverb is typically used in such cases; a prepositional phrase starting with *in* may make the relatum explicit (e.g., *hinten im Raum*).

Furthermore, regions can be partitioned into internal (relative) sections by adopting a "global perspective" (Carroll, 1993) from an observer's viewpoint. In that case, the observed region is divided into spatial sections and ascribed part regions that are described by the dimensional terms (*front, back, left*, *right*), sometimes explicitly so by referring to sides (such as "on the left/right side"). The observed region can be a specific assembly of objects that are perceived as belonging together or being relevant for the discourse situation, or any other kind of region that is within the limits of perception. For example, in German it is possible to say:

(5) Siehst du den Apfelbaum dort hinten?

[lit., "Do you see the apple tree there in the back?"]

where the visual field is partitioned into regions in relation to the position of the speaker (and hearer). Then, the area close to the observer is referred to as *vorne (front)*, and the area more distant to the speaker within the visual field is referred to as *hinten (back)*. Note that, in the case of *front*, *left*, *right* the regions coincide with the intrinsic perspective of the speaker, but not in the case of *back*, because the intrinsic back region is located behind the speaker, while in the case of partitioning the visual field the back region is a region in front of, but distant from the speaker. Furthermore, no further objects need to be involved that could provide the origin or relatum of relative or intrinsic reference systems. Thus, utterance (5) is based on a different kind of reference system, namely, the partitioning of a certain region (e.g., the visual field) into (internal) subregions of that region.

Apart from the investigations of reference systems, research on spatial dimensional terms has focused on the notion of *spatial templates* (e.g., Carlson-Radvansky & Logan, 1997) delimiting the region of application of the terms, *functional features* of objects affecting applicability (Coventry & Garrod, 2004), and *interaction-related* as well as *discourse task-related aspects* of application. In route descriptions, for example, typically a goal location is described via reference to streets and landmarks which can be easily identified in the real world. There, salience and dimensions of buildings play a role in the choice of landmarks; spatial relations are often sufficiently outlined via simple and vague expressions. Such a scenario differs fundamentally from psycholinguistic spatial localisation experiments where participants are asked to specify an entity's location relative to another.

To complement earlier results on the application of spatial dimensional terms, I conducted empirical studies that addressed referential identification tasks in which speakers employed these terms spontaneously and frequently. The relevant question in such a task is "Which" rather than "Where" (as with most earlier studies addressing limited spatial scenarios). Linguistically, there are some systematic differences between utterances relating to these two kinds of discourse tasks. For example, the German dimensional adjectives (such as *das linke*) and the English dimensional superlatives (such as *leftmost*) can only be employed as an answer to "Which", not "Where" questions: There is no way in which "Where is the object?" can be answered by "The leftmost one"; even stating "It is the leftmost one" could at best be interpreted as a fairly indirect way of answering the question. Other syntactic options, in contrast, are available for both kinds of questions: "The object to the left of the barrel" (a typical answer to "Which object do you mean?") can easily be augmented to "The object is (located) to the left of the barrel" (a typical answer to "Where is the object?"). Note that German dimensional adjectives as well as English dimensional superlatives hardly occur in the literature on spatial language, mirroring the fact that discourse tasks of referential identification have been poorly researched so far.

I used a web-based empirical study to investigate speakers' spontaneous usage of spatial dimensional terms with respect to a simple abstract spatial scenario. The data were examined with respect to choice of strategy, alternatives to locative dimensional terms in object-based instructions, and details about the application of locative dimensional terms, including linguistic variability and conceptual choices such as perspective, relatum, and axes, which together serve as the basis for a reference system. However, the reference system itself cannot be identified directly on the basis of the linguistic form, at least not for those systems that were employed spontaneously by speakers in the present scenario. This is due, on the one hand, to the overwhelming lack of explicitness with respect to either relatum or perspective, or both, and on the other hand, to the fact that there is no one-to-one correspondence between linguistic forms and underlying reference systems. This conclusion has not been recognised widely by researchers, and the consequences of the lack of identifiability of reference systems have not sufficiently been accounted for. Research has mainly focused on the distinctions of perspectives, largely ignoring the diversity caused by the possible underlying relata. If both elements are mentioned explicitly, the reference system can be identified. But as the data show, perspectives are almost never (specifically in English) given explicitly, and relata are mentioned only in a subset of cases, preferably to avoid obvious misunderstandings. Therefore, the results point to the conclusion that in the majority of situations there is more than one underlying reference system that is compatible with the linguistic representation in relation to the situation at hand.

Generally, the web study participants used a broad spectrum of variability on all scales. The analysis showed that linguistic as well as conceptual choices depend heavily on the spatial situation, i.e., the presence of other objects and (imagined) persons, and the available kinds of perspective. In spite of the high variability, however, regular patterns of usage could be identified. The systematic variations found are repeatedly explained by much the same idea, reflecting the participants' motivation to fulfill the task of providing an unambiguous description of the target object in a given situation. Three major principles, at least seem to be at work, influencing speakers' choices: The principle of *contrastivity* ensures that the goal object can be identified among the competing objects. The principle of *minimal effort* leads, on the one hand, to the omission of information that is redundant or easily inferable, and on the other hand, to linguistic and conceptual choices (including dimensional terms versus other kinds of spatial expressions, relata, and spatial axes) that enable referential identification with a minimum of additional information encoded as linguistic modifications. The principle of *partner adaptation* seems to be chiefly responsible for the choice of perspective, preferring the interaction partner's point of view to other options. In some cases, the choice of relatum is similarly affected, but this choice is also influenced by the spatial relationship between the goal object and the entities available as potential relata. These basic principles are closely related to earlier research findings in other scenarios and domains (e.g., Clark & Wilkes-Gibbs, 1986).

Comparing Temporal and Spatial Terms

In the present work, the application of spatial and temporal dimensional terms was investigated systematically with the aim of identifying any kind of interdependency that may be reflected in their usage. For temporal terms, the main conceptual aspects affecting application concern the relationships betweeen events, interrelated with the nature of the events themselves, and the knowledge status of the interlocutor. Thus, temporal dimensional terms are employed whenever two events need to be juxtaposed that are conceptually interrelated in some way, which is often causal in some sense. A purely temporal relationship between the events is the reason for use only in a subset of the cases, namely, when the temporal relationship itself is in focus, or its duration via mentioning an explicit length of time between the events. Otherwise, temporal relationships may be expressed by other, more indirect linguistic means such as clause order, tense, and aspect. Apart from these cases in which the temporal relationship itself is at stake, the presence of a temporal dimensional term usually points to the conceptualisation of some sort of interconnection between the events. However, the terms themselves only express the temporal relationship between the events, not any kinds of associated (causal or other) relationships. Therefore, a further requirement seems to be that the speaker does not wish to directly impose and represent a stronger relationship. Thus, the application of a temporal dimensional term leaves much freedom of interpretation to the interlocutor, which is sometimes reflected by the wider discourse context where implications such as the conceptualisation of a causal relationship may be made explicit.

With respect to spatial terms, the main conceptual aspects affecting application are concerned with the association of an underlying reference system, the discourse task (which influences the degree of specificity and explicitness as well as the choice of axis), and the functional relationship between the objects, which includes the nature of the objects themselves. The wider discourse context may influence the choice of reference system as well as syntactic form. The application of spatial terms is warranted in a situation where spatial relationships are to be represented for some reason, since language does not offer any other (more indirect) means of doing so. Out of the overall repertory of spatial terms, spatial dimensional terms are applied whenever a spatial axis is relevant, for example, because a spatial direction needs to be specified, or because alternative terms, such as distance expressions or those expressing in-between relations, are not contrastive in a situation requiring contrastivity, such as referential identification. In contrastive discourse tasks, perspectives are seldom mentioned explicitly (and even less in English than in German), and relata are mentioned mostly if this enhances reference. Linguistic modifications seem to be more prominent in discourse tasks involving the description of a spatial relation rather than the identification of an object out of several possible candidates, where modifications only come into play if reference is otherwise not unambiguous. Areas of applicability are furthermore affected by the functional relationship between the objects involved.

Thus, it seems that, with temporal dimensional terms, what (often) remains implicit is the conceptualised relationship between the entities, while with spatial dimensional terms, it is often one of the entities themselves – the relatum – which remains implicit, resulting in underspecification with respect to the underlying reference system. This directly reflects the ontological difference between objects and events: objects are directly perceivable and therefore in some cases do not need to be specified linguistically, while events are more abstract and must therefore be retrievable from or delimited by the discourse itself. But in both cases, the interlocutors do not necessarily differentiate between the possible interpretations if more than one works well in the context and no ambiguities arise. Thus, there is an interesting parallel between discourse relations and reference systems: in each case, there are a number of associated concepts that are easily inferable, which may or may not be reinforced by additional linguistic material, and which are often not determinate, that is, there may be more than one possible associated concept that works well in the discourse context without raising communication problems.

Conclusion

Altogether, the detailed parallel investigation of spatial and temporal dimensional terms summarized here points to the conclusion that no dependency relationship can be traced with respect to their application in discourse. This contrasts with the unquestioned close interrelationship concerning the grammatical and semantic structure of these terms. Some of the identified parallels in application between the two kinds of terms can be traced back to shared issues of the underlying conceptual domains of space and time. Other kinds of shared features are due to general principles operating in discourse, such as inference processes and the effect of ruling out possible alternatives. In addition, this work has identified features of applicability where the two kinds of terms clearly differ. With respect to these features, the relation to fundamental conceptual differences between the domains is typically straightforward: time is associated with a number of aspects that are irrelevant for space, and vice versa. Therefore, temporal dimensional terms exhibit different kinds of applicability conditions than spatial dimensional terms do.

References

Carlson-Radvansky, L.A. and G.D. Logan. 1997. The Influence of Reference Frame Selection on Spatial Template Construction. In: *Journal of Memory and Language* 37, pp. 411-437.

Carroll, M. 1993. Deictic and intrinsic orientation in spatial descriptions: a comparison between English and German. In J. Altarriba (ed.), *Cognition and Culture: A Cross-Cultural Approach to Psychology*. Elsevier.

Clark, H.H. 1973. Space, time, semantics, and the child. In: Moore, Timothy E. (ed). *Cognitive Development and the Acquisition of Language*. N.Y. Academic Press.

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

Coventry, K. R. and Garrod, S. C. 2004. *Saying, seeing and acting: The psychological semantics of spatial prepositions*. Psychology Press.

Habel, C. and C. Eschenbach. 1997. Abstract Structures in Spatial Cognition. In C. Freksa, M. Jantzen and R. Valk (eds). *Foundations of Computer Science - Potential - Theory – Cognition*, pp 369-378.

Haspelmath, M. 1997. *From Space to Time. Temporal Adverbials in the World's Languages*. München: Lincom.

Heinämäki, O. 1974. *Semantics of English Temporal Connectives*. Dissertation, Univ. of Texas, Austin.

Herrmann, T. und J. Grabowski. 1994. *Sprechen. Psychologie der Sprachproduktion*. Heidelb.: Spektrum.

Herweg, M. 1991. "Temporale Konjunktionen und Aspekt. Der sprachliche Ausdruck von Zeitrelationen zwischen Situationen." *Kognitionswissenschaft 2: 51-90.*

Lakoff, G. and Johnson, M. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

Lascarides, A. and J. Oberlander, 1993. Temporal Connectives in a Discourse Context. In *Proceedings of EACL 93*, pp260--268, Utrecht, The Netherlands.

Levinson, S.C. 2003. *Space in Language and Cognition*. Cambridge: Cambridge University Press.

Miller, G.A. and P.N. Johnson-Laird. 1976. *Language and Perception*. Cambridge University Press: Cambridge.

Piaget, J. 1955. *Die Bildung des Zeitbegriffs beim Kinde*. Zürich: Rascher.

Schilder, F. 2001. Presupposition triggered by temporal connectives. In: M. Bras and L. Vieu (eds.), *Semantic and Pragmatic Issues in Discourse and Dialogue: Experimenting with CurrentDynamic Theories*. Amsterdam: Elsevier, pp. 85-108.

Tenbrink, T. 2006. Localising objects and events: Discoursal applicability conditions for spatiotemporal expressions in English and German. *Dissertation*, University of Bremen, 2005. Mikrofiche, Staats- und Universitätsbibliothek Bremen, ej 6922/E03.

Tenbrink, Thora. 2007. *Space, time, and the use of language: An investigation of relationships*. Berlin: Mouton de Gruyter.

The Temporal Stability of Skilled Dual-Task Performance

TILO STROBACH[1]

Department of Psychology, General Psychology, Humboldt University Berlin (Germany)

PETER A. FRENSCH

Department of Psychology, General Psychology, Humboldt University Berlin (Germany)

TORSTEN SCHUBERT

Department of Psychology, Experimental Psychology, Ludwig-Maximilians University Munich (Germany)

Recent studies showed that practice reduces performance costs in dual-task compared with single-task situations remarkably. We assume that this is due to acquired skills which enable improved dual-task processing after practice. In the present study, we examined the stability of these dual-task related skills over time. Therefore, we tested the dual-task performance of a visual and auditory task in a session conducted 6 weeks after the practice was finished. Our analyses showed that the task performance was robust in single-task and dual-task situations of the visual task. The data of the auditory task revealed impaired performances after the long delay. In particular, skills of single-task processing were less robust than skills related to dual-task performance. We assume that the stability of the processing skills in the component tasks is associated to the level of task automatization at the end of practice. However, future studies have to examine the stability of dual-task performance in more detail.

Introduction

The combined performance of two concurrent tasks often leads to decreased task performance compared with the separate performances of the component tasks (e.g. Pashler, 1994; Schubert, 1999; Welford, 1952). Usually, it is assumed that specific processing mechanisms interfere with each other in dual-task situations and that the result of this interference is indicated by dual-task costs, i.e. longer processing times and/or increased error rates in dual-task compared to single-task situations. However, recent studies have shown that prolonged practice with the tasks may lead to an impressive improvement of dual-task performance (e.g., Hazeltine; Teague, & Ivry, 2002; Ruthruff, Johnston, & Van Selst, 2001; Van Selst, Ruthruff, & Johnston, 1999; Spelke, Hirst, & Neisser, 1976). Some studies even demonstrated perfect dual-task performance as indicated by the complete elimination of dual-task costs after prolonged dual-task practice (Hazeltine et al., 2002; Hirst, Spelke, Reaves, Caharack, & Neisser, 1980; Schumacher et al., 2001).

---

[1] Anschrift: Tilo Strobach, Department of Psychology, General Psychology, Rudower Chaussee 18, D-12489 Berlin, email: tilo.strobach@psychology.hu-berlin.de. Gefördert durch Deutsche Forschungsgemeinschaft (DFG), Projekt Schu-1397/2-3; Fr 1471/ 8-3.

The finding of a practice-related decrease of dual-task costs (and even of their complete elimination) suggests that learning takes place during dual-task practice. As a result of these learning processes, participants acquire a skill that allows them to improve the dual-task processing and, even more importantly, to process two tasks concurrently in some situations in a perfect manner. Both the precise learning mechanisms leading to the acquisition of this skill (e.g., Ruthruff et al., 2001; Ruthruff, Van Selst, Johnston, & Remington, 2006) as well as the stability of the acquired skill are still a matter of debate (e.g., Bherer et al., 2005). The later question in particular is of considerable interest in the present study because its investigation would allow to specify the nature of the skill that underlies improved dual-task performance. Additionally, this issue is of actual concern for a number of recent studies investigating practice-related improvements of different types of attentional mechanisms (Green & Bavelier, 2003).

The aim of the present study was, therefore, to re-assess the dual-task performance of highly-skilled participants after a longer pause. In more detail, participants performed a dual-task paradigm consisting of a visual-manual task and an auditory-verbal task according to the procedure described by Schumacher et al. (2001) for 10 sessions. The use of the Schumacher et al. procedure was thought to lead to a strong reduction of dual-task costs after practice. After the tenth session, we introduced a pause of no less than six weeks during which participants had no further contact with the dual-task situation. After the pause, we conducted a post-test session with exactly the same dual tasks used originally. The comparison between the performance of participants in the last practice session (i.e., before the pause) and the post-test session (i.e., after the pause) allows us to assess the stability of dual-task skill over time.

The stability of acquired skills has been investigated in a number of different domains and with a variety of different paradigms, though not specific for the domain of dual-task performance. Many studies suggest that forgetting of the acquired memory trace over time may represent one important factor affecting the stability of skills. A number of reasons have been proposed that may be responsible for the loss of a skill as a result of forgetting, e.g., decay (Brown, 1958), interference (Keppel, Postman, & Zavortink, 1968) or a combination of the two (Altmann & Gray, 2002). Evidence suggesting forgetting of acquired knowledge with time of a performance pause has been reported for different situations, e.g., Ebbinghaus (1885/ 1913), and more specifically for complex tracing tasks (Eysenck & Willett, 1966), recall of social information (Macrae & MacLeod, 1999), artificial grammar (Tamayo & Frensch, 2007) and other tasks. Willingham and Dumas (1997), in particular, reported findings suggesting lack of retention and of performance savings in a re-learning situation when participants performed an implicit serial response time task after a long pause.

Importantly, applied to the dual-task situation in which participants perform two choice reaction-time tasks, forgetting might be related to any part of the task components including the processing of sensory and motor information and the retrieval of the corresponding translation of stimulus to response information (e.g., Ahissar, Laiwand, & Hochstein, 2001; Van Selst et al., 1999). However, if forgetting is especially related to the dual-task specific part of the acquired skill, then a time-related deterioration should be especially strong in the dual-task compared to single-task situations after a long pause. An example for dual-task specific knowledge is the coordination of processing streams of simultaneously performed component tasks (e.g., Hirst et al., 1980).

In contrast, a number of studies showed that task performance may improve even after practice has finished. Examples of improved performance after a pause have been reported in studies investigating basic skills of visual discrimination (Karni, Tanne, Rubenstein, Askenasy, & Sagi, 1994; Stickgold, James, & Hobson, 2000), motor adaptation (Huber, Ghilardi, Massimini, & Tononi, 2004) or the performance of motor sequences (Walker, Brakefield, Morgan, Hobson, & Stickgold, 2002) after sleep. The obtained improvement is often explained by the assumption of processes of memory consolidation which cause that skill representations gradually acquire permanence (Bosshardt et al., 2005) and that they are converted into an optimal integrated memory trace (Stickgold, 2005). According to this hypothesis, we assume improved performance in single tasks and dual tasks over time when skills related to both situations are improved (e.g., processing stages of the component tasks). However, when the improvement is limited to dual-task specific skills then we should observe a pause-related performance improvement specifically for the dual-task situations.

However, two arguments challenge an assumption of improved dual-task performance over time. Findings showing memory consolidation were revealed in post-test sessions that followed only one practice session. Contrary, the post-test session in the present study was conducted after participants performed ten practice sessions. Whereas participants showed processes of memory consolidation after very minimal practice of task skills, we conducted a post-test session with highly-skilled participants. Although some studies showed ongoing processes of memory consolidation even after two practice sessions on following days (e.g., Duke & Davis, 2006; Walker et al., 2003), no study exists that showed signs of memory consolidation after extensive practice (i.e., ten practice sessions). We assume that the impact of sleep on the performance of highly skilled participants is extremely reduced and processes of memory consolidation are negligible.

Furthermore, memory consolidation was investigated in tasks including very basic procedural visual and motor skills. The present dual-task situation, instead, comprises mechanisms of regulating goal-directing behaviour in complex situations with interfering information (Schubert & Szameitat, 2003). Therefore, it is uncertain whether findings of performance gains after practice may be attributed to situations similar to the present dual-task paradigm.

A further possibility for the impact of a pause on performance is that performance in single and dual tasks is robust over time, i.e. remains stable. Stable task skills would be indicated by the lack of a difference between the performance measured before and after the pause. (Note that different reasons may be responsible for the lack of a difference between the performance measures before and after the pause; these reasons will be discussed later.).

In the literature, there are only a few studies providing empirical data that may be related to the issue of the stability of dual-task performances over time. For example, Ruthruff et al. (2001) conducted post-test sessions in which five of six participants of a previously conducted learning study performed a further dual-task situation (Van Selst et al., 1999). However, unfortunately, due to the administration of different component tasks in the post-test compared to the learning sessions it is impossible to draw exact inferences about the stability of the dual-task performance on the basis of these findings. In a further study, Ruthruff, Johnston, Van Selst, Whitsell, and Remington (2003) investigated the remaining sixth participant of the Van Selst et al. study. The authors conducted the identical dual-task situation before and after a pause of 14 months. However, the results obtained immediately after the pause were not fully reported by Ruthruff et al. (2003), which again limits the value of the study for our purposes.

In a further study, Bherer et al. (2005) analyzed the performances of groups of young and older participants in a dual-task situation after a pause of four to six weeks. Because older participants showed a general slowing in both single- and dual-task trials equal dual-task costs were observed between old and young participants after the pause (Bherer, personal communication, April 16, 2007). Although, this finding suggests robust dual-task performance over time, it needs to be treated with caution because only a few participants appeared for the follow-up session compared to the learning session (Bherer et al., 2005).

While systematic knowledge on the stability of the dual-task skill seems rather rare, findings in related research areas might provide a more revealing picture. For example, performance in the so-called task switching paradigm in which participants have to switch consecutively between two tasks is assumed to share some common processes with dual-task performance (e.g., Band & van Nes, 2006; Lien, Schweikert, & Proctor, 2003; West, 1996). Kramer, Hahn, and Gopher (1999) investigated the stability of task-switching skill in a group of young and older participants. They found that task-switching performance in both groups was robust over two months. However, it may be inappropriate to generalize findings with the task-switching paradigm to the dual-task situation because despite the number of similarities between the paradigms, there are a number of important differences as well (Pashler, 2000).

Taken together, there seems to exist a rather heterogeneous pattern of results concerning the stability of dual-task skill and prior findings with related paradigms seem to have limited generalizability. Therefore, the present study is the first to systematically investigate the stability of dual-task skill over time.

EXPERIMENT

Participants practiced a dual-task situation for 10 sessions and performed a post-test session after a pause of six weeks without practice. Participants performed identical visual and auditory tasks in single- and dual-task conditions before and after the pause.

Methods

Participants. Nine participants (undergraduates, mean age = 25.9 years, 4 females) participated in the experiment and received payment of € 80 plus performance-based bonuses. All participants had normal or corrected-to-normal vision.

Apparatus and component tasks. Participants concurrently conducted two speeded choice reaction tasks. In the visual task, participants responded manually to a white stimulus appearing at the left, central, or right position arranged horizontally on the computer screen. Whereas only circles were presented in session 1 to 8, the presentation of circles and triangles was intermixed in session 9 to 11. The stimuli appeared equally balanced in these sessions. However, only the data of trials with circle presentations are of current interest and will be described in the following sections. Three white dashes served as placeholders for the possible positions of the visual stimuli. They appeared as a warning signal 500 ms before the visual stimulus, a circle, was presented. The stimulus remained visible until the participant responded or until a 2.000 ms response interval had expired. Half of the participants responded to the stimuli by making a spatially compatible key press with the index, middle, and ring fingers of their right hand. The remaining participants responded with the ring, middle, and index fingers of their left hand. The responses were recorded with a button box connected to the experimental computer.

In the auditory-verbal task, participants responded to sinus-wave tones presented with frequencies of either 300, 950, or 1650 Hz by saying “ONE”, “TWO”, or “THREE” (German: “EINS”, “ZWEI”, “DREI”), respectively. A trial started with the presentation of three dashes on the computer screen. After an interval of 500 ms, the tones were presented for 40 ms. The trial was finished when the participant responded verbally or a 2.000 ms response interval had expired. To analyze the accuracy of each response, the experimenter recorded the verbal responses. The reaction time was recorded via a voice key connected to the experimental computer.

After correct responses in the visual and in the auditory task the RTs were presented for 1500 ms on the screen. Following incorrect responses, the word “ERROR” (German: “FEHLER”) appeared. A blank interval of 700 ms preceded the beginning of the next trial in both component tasks.

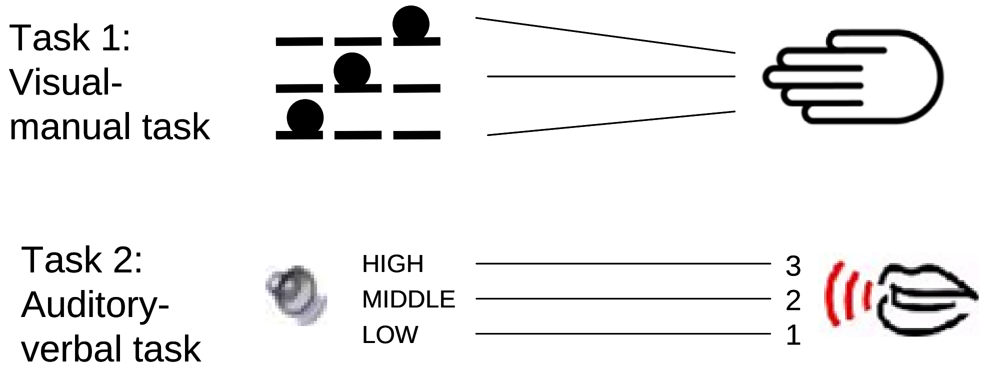


Figure 1. Illustration of the visual-manual and the auditory verbal task conducted in the present experiment.

Design and Procedure. Participants performed 11 experimental sessions. Whereas the first 10 sessions were conducted on successive days, there was a pause of approximately (i.e., in any case no less than) six weeks before the post-test session. Importantly, the sessions directly before (session 10) and after the pause (session 11) were identical.

Two different types of trials were presented during the experiment: (1) single-task trials in which only one stimulus (visual or auditory) was presented and (2) dual-task trials in which a visual and an auditory stimulus were presented simultaneously and participants were instructed to put equal emphasis on both stimuli. The trials were presented in three different types of experimental blocks: (1) blocks with the visual task as single tasks, (2) blocks with the auditory task as single tasks (*single-task blocks*), and (3) blocks with the visual and the auditory task as single and as dual tasks (*mixed blocks*). Single-task blocks consisted of 45 trials in the sessions 1 to 8 and 48 trials in sessions 9 to 11. In mixed blocks, 30 single (15 visual and 15 auditory task trials) and 18 dual-task trials were presented in sessions 1 to 8. In sessions 9 to 11, 24 single- (12 visual and 12 auditory task trials) and 18 dual-task trials were presented in mixed blocks. The stimuli were presented in random order in each block and participants were instructed to respond as fast and as accurate as possible regardless of whatever else was presented on a trial.

During the first session, participants conducted six single-task blocks of each task type in an alternate order counterbalanced across participants. Each subsequent session proceeded as follows: Participants began with two single-task blocks (1 of each task type) and subsequently performed 14 blocks consisting of four single-task blocks (2 of each task type) and 10 mixed blocks (in session 2 only eight mixed blocks). Excluding the initial two blocks, single-task blocks were alternated and separated by two mixed blocks. The procedure of the last three sessions was identical to the prior sessions, except that the 19 blocks (6 single-task and 13 mixed blocks) followed the two initial blocks. Again, single-task blocks were alternated and separated by two mixed blocks. Half of the subjects started each session with a visual single-task block and the other half with an auditory single-task block.

Table 1. Overview of the task conditions in the present experiment.

Sessions	Task conditions
1	single tasks
2 – 10 (i.e., before pause)	single and dual tasks
pause of > 6 weeks	
11 (i.e., after pause)	single and dual tasks

Results

In the main analysis of the present study, the dependent variables were the reaction time (RTs) in single tasks (limited to trials in single-task blocks) and dual tasks and the error rates in single tasks and dual tasks. The difference between both (i.e., performance in dual-task minus performance single-task trials) reflects the amount of dual-task costs. We conducted separate analyses for the RTs and for the error rates in the visual and the auditory task. As correct dual-task trials we defined trials with correct responses in the visual and in the auditory task. Only trials with two correct responses were included in the RT analysis of the dual-task data. The remaining trials were excluded from the analysis.

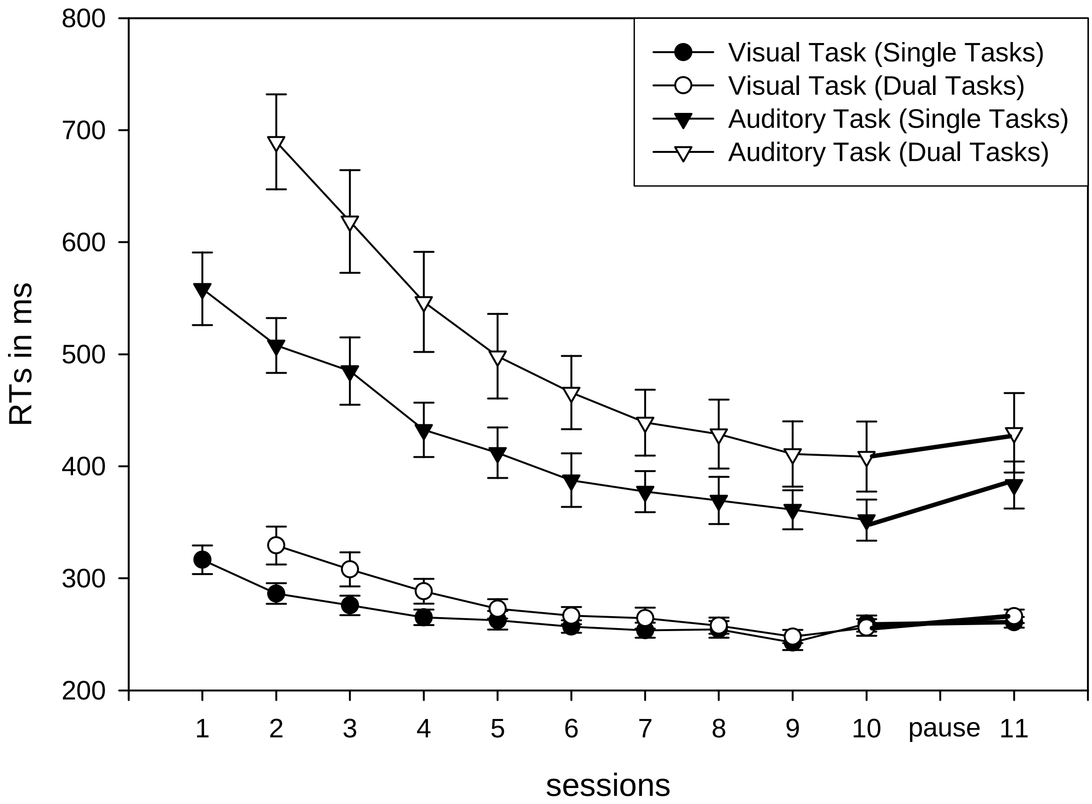


Figure 2. Mean dual-task RTs by session (session 1 to session 11) and trial type (single tasks vs. dual tasks) for the visual and the auditory tasks (bars represent standard errors).

First, we analyzed the effect of practice on the dual-task performance for the learning period before the pause, i.e., for the learning sessions 2 to 10. For that purpose, we conducted a two-way repeated measures ANOVA with the within-subject factors session (session 2 – session 10) and trial type (single tasks vs. dual tasks). As shown in Figure 2, the mean RTs in the visual task decreased during practice as reflected by an effect of session, F (8, 64) = 24.321, $p < .001$. Participants responded faster in single-task (262 ms) than in dual-task situations (277 ms), F (1, 8) = 5.927, $p < .05$. The dual-task costs in the RTs of the visual task decreased during practice as indicated by an interaction of session and trial type, F (8, 64) = 5.302, $p < .001$. The costs were reduced from the beginning (session 2: 43 ms; $p < .05$) to the end of practice where they vanished completely (session 10: -3 ms; $p > .62$). For the visual task, the analysis of errors (Table 2) showed a marginal effect of trial type, F (8, 64) = 5.192, $p < .06$, with lower error rates in single (2.3 %) than in dual tasks (4.9 %). The dual-task costs in the errors decreased significantly over practice as demonstrated by the interaction of session and trial type, F (8, 64) = 2.376, $p < .05$. Whereas the error analysis showed dual-task costs in session 2 (5.9 %; $p < .01$), the costs were eliminated in session 10 (2.0 %; $p > .18$). The factor session showed no significant main effect on the error rate.

For the auditory task, we obtained a significant effect of session on the mean RTs, F (8, 64) = 41.513, $p < .001$ (Figure 2), indicating a general improvement of the task performance during practice. Furthermore, the analysis showed an effect of the factor trial type on the RTs, F (1, 8) = 18.829, $p < .01$, with lower RTs in single (409 ms) than in dual task trials (500 ms). Participants improved their dual-task performance during practice as is reflected by the significant interaction of session and trial type, F (8, 64) = 17.588, $p < .001$. The dual-task costs were reduced extensively from session 2 (182 ms) to session 10 (57 ms) but were still present at the end of practice (session 2: $p < .001$ vs. session 10: $p < .05$) in the auditory task. As shown in Table 2, for the auditory task the error analysis showed a significant difference between single-task (3.0 %) and dual-task situations (4.9 %), F (1, 8) = 5.555, $p < .05$. There was a slight decrease of the dual-task costs during learning as demonstrated by a marginal interaction between session and trial type F (8, 64) = 1.859, $p < .10$. However, dual-task costs in the error rates were not eliminated after practice, i.e., some costs remained in session 10 (2.8 %; $p < .018$). The factor session revealed no significance in the analysis of errors.

Taken together, our results indicate a strong improvement of the dual-task performance after extensive dual-task practice for the RT and the error data in the visual task and for the RT data in the auditory task. This reduction supports results of previous studies of dual-task practice (e.g., Van Selst et al., 1999). It is important to note that despite the improvement of dual-task performance, we found no complete elimination of the dual-task costs (Hazeltine et al., 2002; Schumacher et al., 2001). Although the costs were eliminated in the visual task, they remained significant in the auditory task.

Table 2. Percent of errors (standard errors) by session (session 1 to session 11) and trial type (single tasks vs. dual tasks) for the visual and the auditory tasks.

	Visual task		Auditory task	
Sessions	Single tasks	Dual tasks	Single tasks	Dual tasks
1	0.91 (0.28)	-	5.22 (1.33)	-
2	1.23 (0.59)	7.05 (1.37)	2.80 (0.83)	7.05 (1.37)
3	1.65 (0.34)	4.44 (0.91)	3.79 (1.20)	4.44 (0.91)
4	1.81 (0.43)	4.05 (0.78)	1.81 (0.54)	4.05 (0.78)
5	1.81 (0.37)	3.96 (1.25)	2.72 (0.39)	3.96 (1.25)
6	2.30 (0.31)	4.59 (1.10)	2.72 (0.41)	4.59 (1.10)
7	3.05 (0.38)	5.65 (1.96)	3.70 (0.79)	5.65 (1.96)
8	2.80 (0.63)	4.41 (1.35)	2.96 (0.64)	4.41 (1.35)
9	2.89 (0.48)	4.38 (0.95)	3.50 (0.52)	4.38 (0.95)
10	3.47 (0.65)	5.49 (1.15)	2.72 (0.49)	5.49 (1.15)
Pause of > 6 weeks				
11	1.74 (0.35)	4.11 (0.57)	2.08 (0.43)	4.11 (0.57)

To analyze the temporal stability of skilled dual-task performance after extensive practice, we compared the performance of the participants in single- and dual-task trials between the session 10 and session 11. Both sessions were separated by a pause of more than 6 weeks. For that purpose, we conducted an ANOVA with the factors session (session 10 vs. session 11) and trial type on the RTs and the error rates in the visual and the auditory tasks.

Visual task: Importantly, there were neither significant effects nor an interaction of the pause on the RTs in the visual task (Figure 1). This finding, in particular, suggests a comparable level of dual-task performance before and after the pause and is indicative for robust skill acquisition in the visual task. The error analysis of the visual task showed an effect of session, $F(1, 8) = 9.275$, $p < .05$. Error rates were lower in session 11 (2.9 %) than in session 10 (4.5 %). Though the interaction of session and trial type revealed no significance, $F(1, 8) < 1$, we suggest that this performance improvement was rather due to an unspecific effect and not due to a specific effect of the pause on dual-task-related skills The factor trial type was also significant, $F(1, 8) = 5.896$, $p < .05$, with lower error rates in single-task (2.6 %) than in dual-task trials (4.8 %). Thus, the present RT analyses of the visual task revealed no specific effects of forgetting or consolidation over time. However, the results of the error rates suggested overall task improvement after the pause. Though the results were similar in single-task and dual-task conditions, we interpret this finding by the assumption that the observed improvement is not specific for the dual-task related skills.

Auditory task: The RT analysis of the auditory task showed a marginal effect of session, $F(1, 8) = 4.898$, $p < .06$, with slightly faster responses in session 10 (380 ms) than in session 11 (407 ms). Thus, the pause led to an increase of the RTs in the auditory task in the session after compared to the session before the pause. In addition, participants were faster in single tasks (368 ms) than in dual tasks (419 ms), $F(1, 8) = 9.369$, $p < .05$. Interestingly, the interaction of both factors revealed significance, $F(1, 8) = 7.523$, $p < .05$. The RTs in the single-task conditions increased by 31 ms ($p < .05$) and the RTs in the dual-task conditions by 21 ms ($p > .09$) from session 10 to session 11. Because of this different amount of the pause-related performance decrement in single-task and dual-task trials, the resulting dual-task costs slightly decreased after the pause from 57 ms (session 10) to 46 ms (session 11). The error analysis showed only an effect of trial type, F (1, 8) = 11.017, $p < .05$, with lower error rates in single (2.4 %) than in dual tasks (4.8 %). No effect of session or an interaction was revealed. In sum, we found decreased task performance after the pause in the RT data of the auditory task with a stronger impairment in single tasks than in dual tasks and no effect of the pause on the error rate.

Discussion

In the present study, we investigated the effects of extensive dual-task practice and the temporal stability of skilled dual-task performance in a situation requiring the simultaneous performance of a visual and an auditory task. For that purpose, we examined the performance of participants six weeks after they had completed an extensive period of 10 consecutive days of intense dual-task practice. A comparison of the dual-task performance before and after the pause reveals the degree of stability of the acquired dual-task skills over time.

The analysis of the practice-related changes during the dual-task practice showed a strong improvement of the task performance in the visual and in the auditory tasks. This improvement was more pronounced in dual- than in single-task conditions. Consequently, the additional performance costs appearing in dual-task situations were extremely reduced after 10 practice session, thus, replicating findings of prior studies on effects of practice on dual-task performance (e.g. Hirst et al., 1980; Ruthruff et al., 2003, 2006; Van Selst et al., 1999). Whereas the costs were completely eliminated in the visual task, some costs remained in the auditory task which is in contrast to other studies that showed complete elimination of costs in both component tasks in a similar paradigm (Hazeltine et al., 2002; Schumacher et al., 2001; but see Tombu & Joliceour, 2004).

The administration of a long pause of dual-task practice had different effects on task performance depending on the specific component task to be performed. For the visual task, the corresponding analysis of the data in the single-task and dual-task conditions revealed stability of skills when considering RT data and, interestingly, it revealed an improvement in the error analysis after the pause of six weeks. Thus, unlike findings of decreased task performance after a long delay (e.g., Ebbinghaus, 1885/ 1913; Tamayo & Frensch, 2007; Willingham & Dumas, 1997) the acquired skills in the present visual task do not suffer from forgetting after a long pause. Instead, the acquired skills allowed stable or improved performance even after the pause. Importantly, however, this conclusion holds true when considering performance in both, i.e., single-task and in dual-task trials. There are no signs for an effect of the pause that might depend on whether the visual task was performed either alone or together with the auditory task. So, we assume that robustness of skills in the visual task over time is not exclusively attributed to skills related to dual tasks but is associated to single-task and dual-task related skills.

The analysis of the auditory task revealed impaired performances in single-task and dual-task conditions after a long pause of six weeks. The impairment was stronger in single-task than in dual-task trials. In particular, this difference between single- and dual-task trials caused a decreased amount of dual-task costs before and after pause, with smaller dual-task costs after the pause. On first glance, this finding should be interpreted by the assumption of an improvement of dual-task skills compared to the single-task skills. However, based on the finding that RTs in dual-task trials increased in the session after the pause compared to the session before the pause we refrain from an interpretation of the related finding with the assumption of a specific memory consolidation for dual-task-related skills in the auditory task. In contrast, in our view, the findings of the auditory task show that the skills acquired in dual-task situations are impaired by the introduced performance pause but this impairment is, simply, less pronounced in dual-task compared to single-task trials.

The observed discrepancy regarding the temporal stability of the skills in visual and auditory component tasks is an interesting finding and needs further consideration. In detail, in the visual task, we did not find any dual-task costs before and after the pause. Johnston and Delgado (1993) argued that the complete reduction of performance costs in dual-task situations can be explained by an automatization of the processing of the component tasks. The automatization leads to a bypass and a parallel processing of bottleneck processes. We suggest that, once acquired, a skill that enables perfect dual-task performance seems to be stable over a long lasting pause. The finding of relatively stable dual-task performance over time is indicative for robust or even improved dual-task skills for the visual task, which is not affected by processes of forgetting. According to Shiffrin and Schneider (1977), robustness over time is a characteristic of the complete automatization of task processing. This explanation is consistent with the assumption of a highly automatic processing of the visual task in the present study after extensive practice.

It is important to note, however, that the finding of stable performance over time might be explained by a dynamic interaction of processes that are associated with forgetting and memory consolidation. Here specifically, the latter may result in some compensation of dual-task performance when forgetting of a skill over time is accompanied by a similar improvement of the dual-task skill (e.g., by memory consolidation). Future studies are necessary to disentangle the compensatory explanation from the robustness explanation.

For the auditory task, however, the findings showed decreased performances after the pause. Such impairments of performance after a long pause are often associated with processes of forgetting where acquired memory representations decay after some time. As the analysis of the practice sessions showed some dual-task costs still remaining at the end of practice after 10 sessions. Based on these findings we suggest that skills acquired during the practice of the auditory tasks were not completely automatic but still comprised non-automatic processes (Ruthruff et al., 2006). According to Shiffrin and Schneider (1977), these processes can be subject to dynamic changes over time and this, in particular, may result in forgetting of task skills of the auditory task after some time. Thereby, the differentiation of automatic and non-automatic processes provides an appropriate approach for explaining the differences in performance stability in the visual and the auditory tasks over time.

However, what may be the reasons for the different levels of automatization in the present visual and auditory tasks? According to Ruthruff et al. (2006), the different degrees of the compatibility relations between stimuli and responses in the component task may be one reason for different degrees of automatization in the two tasks. We assume that an increasing degree of compatibility of S-R translations is associated with increased automatization of the component task processing. Based on this assumption we used a spatially highly-compatible task (i.e., the visual task) where stimulus and response codes show a highly “natural” dimensional overlap and we used a task with an arbitrary S-R mapping (i.e., auditory task) where there is no pre-learned relationship between stimulus and response information. Adapted to the present task situation, we found stability in the highly spatial-compatible task but deterioration in the task with an arbitrary S-R mapping. Therefore, the present data are consistent with the assumption that an increased level of automatization may be achieved in tasks with a larger degree of S-R compatibility (as in the visual task) compared to tasks with arbitrary S-R mapping (as in the auditory task).

An important issue which needs to be discussed relates to the questions what specific skills will be improved as a result of dual-task practice and how this contributes to the overall improvement of dual-task performance. Mechanisms on two levels were proposed in previous studies to explain improved dual-task performance after practice: (1) mechanisms on the level of component tasks (e.g., Ahissar et al., 2001; Ruthruff et al., 2003, 2006) and (2) mechanisms on the level of inter-task coordination (Hirst et al., 1980; Kramer, Larish, & Strayer, 1995). The comparison of performances of single tasks and dual tasks before and after the pause might allow some conclusions whether mechanisms of the component task processing or the mechanisms related to dual-task specific skills (e.g., inter-task coordination) are stable or change over time. In detail, we assume that similar performance differences between single-task conditions and dual-task conditions over time (i.e., similar dual-task costs before and after the pause) rather indicate effects of mechanisms not exclusive for skills of dual-task performance. By contrast, different costs in dual-task situations before and after the pause show pause-related effects rather associated to dual-task skills. Most of the findings of the presented study (i.e., RTs and error rates in the visual task and error rates in auditory task) revealed similar performance differences between single-task and dual-task condition. Therefore, we assume the stability is not limited to skills related to dual tasks. The analysis of RTs in the auditory task showed that the dual-task costs are reduced over time. As indicated, however, both single and dual tasks reflect deteriorated performances after the pause. Therefore, we refrained to interpret this finding as an indicator for improved dual-task skills after pause. Further studies including the comparison of groups with dual-task practice and groups with single-tasks practice will be necessary to investigate exactly the question which mechanisms are attributed to the stability of the dual-task performance over time. That is, because the practice of component tasks in dual-task situations provides the acquisition of dual-task related skills. The separate practice, however, does not provide the acquisition of these skills exclusively related to dual tasks (Hirst et al., 1980).

In sum, the present investigation revealed that the stability of single-task and dual-task performance over time is based on the specific characteristics of the component tasks that constitute the dual tasks. The analysis of a visual task before and after a pause showed that the performance skills of single and dual tasks are highly robust. The data of an auditory task, instead, showed impaired task performance in single-task and dual-task conditions after the pause where skills of single tasks were more impaired than dual-task related skills. We assume that these task-dependent outcomes are based on the different levels of automatization of these tasks at the end of practice. However, the question whether a pause has different influences on types of mechanisms associated to improved dual-task performance remains to be considered in future studies.

Acknowledgment

We thank Hubert Zimmer and an anonymous reviewer for their very helpful comments. We are grateful to Anne-Marie Horn for support in data collection.

References

Ahissar, M., Laiwand, R. & Hochstein, S. (2001).Attentional demands following perceptual skill training. *Psychological Science, 12*(1), 56-62.

Altmann, E. M. & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science, 13*(1), 27-33.

Band, G. P. H. & Van Nes, F. T. (2006). Reconfiguration and the bottleneck: Does task switching affect the refractory-period effect? *European Journal of Cognitive Psychology, 18*, 593 – 623.

Bherer, L., Kramer, A. F., Peterson, M. S., Colcombe, S., Erickson, K. & Becic, E. (2005). Training effects on dual-task performance: Are there age-related differences in plasticity of attentional control? *Psychology and Aging, 20*(4), 695-709.

Bosshardt, S., Schmidt, C. F., Jaermann, T., Degonda, N., Boesiger, P., Nitsch, R. M., Hock, C. & Henke, K. (2005). Effects of memory consolidation on human hippocampal activity during retrieval. *Cortex*, *41*, 486–498.

Brown, J. (1958). Some Tests of the Decay Theory of Immediate Memory. *Quarterly Journal of Experimental Psychology*, *10*, 12-21.

Duke, R. A. & Davis, C. M. (2006). Procedural memory consolidation in the performance of brief keyboard sequences. *Journal of Research in Music Education, 54*, 111-124.

Ebbinghaus, H. (1885/ 1913). Memory: A Contribution to Experiment Psychology. New York: Dover.

Eysenck, H. J. & Willet, R. A. (1966). The effect of drive on performance and reminiscence in a complex tracing task. *British Journal of Psychology, 57*(1), 107-112.

Green, C. S. & Bavelier, D. (2003). Action video games modify visual attention. *Nature, 423*, 534-537.

Hazeltine, E., Teague, D. & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 527-545.

Hirst, W., Spelke, E. S., Reaves, C. C., Caharack, G. & Neisser, U. (1980). Dividing attention without alteration or automaticity. *Journal of Experimental Psychology: General, 109*, 98–117.

Huber, R., Ghilardi, M. F., Massimini, M. & Tononi, G. (2004). Local sleep and learning. *Nature, 430*, 78–81.

Johnston, J. C. & Delgado, D. F. (1993, November). *Bypassing the single-channel bottleneck in dual-task performance.* Paper presented at the 34th Annual Meeting of the Psychonomic Society, Washington, DC.

Karni, A., Tanne, D., Rubinstein, B. S., Askenasy, J. J., & Sagi, D. (1994). Dependence of REM sleep of overnight improvement of a perceptual skill. *Science, 265*, 679-682.

Keppel, G., Postman, L. & Zavortink, B. (1968). Studies of learning to learn: VIII. The influence of massive amounts of training upon the learning and retention of paired-associate lists. *Journal of Verbal Learning and Verbal Behavior, 7*, 790–796.

Kramer, A. F., Larish, J. F. & Strayer, D. L. (1995). Training for Attentional Control in Dual Task Settings: A Comparison of Young and Old Adults. *Journal of Experimental Psychology: Applied, 1*(10), 50-76.

Kramer, A. F., Hahn, S. & Gopher, D. (1999). Task coordination and aging: Explorations of executive control processes in the task switching paradigm. *Acta Psychologica, 101*, 339-378.

Lien, M. C., Schweikert, R. & Proctor, R. W. (2003). Task switching and response correspondence in the PRP-paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 692-712.

Macrae, C. N. & MacLeod, M. D. (1999). On recollections lost: When practice makes imperfect. *Journal of Personality and Social Psychology, 77*, 463-473.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*, 220-244.

Pashler, H. (2000). Task switching and multitask performance. In Monsell, S. & Driver, J. (editors). *Attention and Performance* XVIII: Control of mental processes. Cambridge, MA: MIT Press.

Ruthruff, E., Johnston, J. C. & Van Selst, M. V. (2001). Why practice reduces dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 3-21.

Ruthruff, E., Johnston, J. C., Van Selst, M. V., Whitsell, S. & Remington, R. (2003). Vanishing dual-task interference after practice: Has the bottleneck been eliminated or is it merely latent? *Journal of Experimental Psychology: Human Perception and Performance, 29*, 280-289.

Ruthruff, E., Van Selst, M., Johnston, J. C. & Remington, R. W. (2006). How does practice reduce dual-task interference: Integration, automatization, or simply stage-shortening? *Psychological Research*, *70*, 125-142.

Schubert, T. (1999). Processing differences between simple and choice reaction affect bottleneck localization in overlapping tasks. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 408–425.

Schubert, T. & Szameitat, A. J. (2003). Functional neuroanatomy of interference in overlapping dual tasks: An fMRI study. *Cognitive Brain Research*, *17*, 733-746.

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E. & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science, 12*(2), 101-108.

Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127 – 190.

Spelke, E. S., Hirst, W., & Neisser, U. (1976). Skills of divided attention. *Cognition, 4*, 215-230.

Stickgold R. (2005). Sleep-dependent memory consolidation. *Nature, 437*, 1272-1278.

Stickgold, R., James, L., & Hobson, J. A. (2000). Visual discrimination learning requires sleep after training. *Nature Neuroscience, 3*, 1237-1238.

Tamayo, R. & Frensch, P. A. (2007). Interference produces different Forgetting Rates for implicit and explicit knowledge. *Experimental Psychology, 54*(4). 304-310.

Tombu, M. & Jolicoeur, P. (2004). Virtually no evidence for virtually perfect time-sharing. *Journal of Experimental Psychology: Human Perception & Performance*, *30*, 795-810.

Van Selst, M., Ruthruff, E. & Johnston, J. C. (1999). Can practice eliminate the Psychological Refractory Period Effect? *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1268-1283.

Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A. & Stickgold, R. (2002). Practice with sleep makes perfect: sleep dependent motor skill learning. *Neuron, 35*, 205–211.

Walker, M. P., Brakefield, T., Seidmann, J., Morgan, A., Hobson, J. A., & Stickgold, R. (2003). Sleep and the time course of motor skill learning. *Learning & Memory, 10,* 275-284.

Welford, A. T. (1952). The psychological refractory period and the timing of high speed performance - A review and a theory. *British Journal of Psychology, 43,* 2 – 19.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120*, 272–292.

Willingham, D. B. & Dumas, J. A. (1997). Long term retention of a motor skill: Implicit sequence knowledge is not retained after a one year delay. *Psychological Research*, *60*, 113-119.

Dual-task performance while driving a car: Age-related differences in critical situations

GISA ASCHERSLEBEN[a] & JOCHEN MÜSSELER[b]

[a] Department of Psychology, Developmental Psychology Unit
Saarland University, Germany[1]

[b]Psychology Department
RWTH Aachen University

Dual task conditions are especially demanding for older drivers. For example, while driving a car the driver has to process additional information (e.g. from the navigation system) and to react in an appropriate way. Especially under demanding driving conditions impairments in reaction time have to be expected. One important factor is the cross-task compatibility, that is whether the information presented in the secondary task provides information, which is compatible or incompatible to the primary task. First results indicate that effects of cross-task compatibility depend on the relevance of the driving situation and on the direction of the response that has to be performed in the primary task, and not so much on the position at which the information is presented.

Background

Dual task situations are ubiquitous, at home as well as at work. Especially modern information technologies often require the coordination of different activities. For example, while driving a car the driver has to process additional information (e.g. from the navigation system or from the mobile phone) and to react in an appropriate way. This is especially the case for professional drivers like taxi drivers or lorry drivers. Under less demanding driving conditions, e.g. slow driving with low traffic density, the coordination of the primary task (driving the car) and the secondary task (manipulating the navigation system) might be relatively easy. However, under demanding driving conditions and especially in critical traffic situations, impairments in reaction times and driving errors can be expected. Thus, the aim of the present study was to analyze the impairments in performance that have to be expected under dual task conditions in critical traffic situations. We were especially interested in the age-related differences in performance.

Since the 1930s dual-task performance is an important topic in cognitive psychology (Telford, 1931). In a typical dual-task experiment, participants are required to react as fast as possible to two stimuli being presented in close succession. For example, participants are instructed to press a left/right key (R1) in response to a high/low tone (S1) and to say „blue“ or „yellow“ (R2) in response to a blue or yellow stimulus (S2). Under conditions with a short time interval (stimulus-onset asynchrony, SOA) between the presentation of S1 and S2 (i.e. S2 is presented

[1] Address: Prof. Dr. Gisa Aschersleben, Department of Psychology, Developmental Psychology Unit, Saarland University, P.O. Box 15 11 50, D-66041 Saarbrücken, email: aschersleben@mx.uni-saarland.de, web: www.uni-saarland.de/fak5/entwicklung, sponsorship: German Science Foundation (DFG Mu 1298/5).

when response preparation and production of R1 has not yet been completed), dramatic increases in reaction time and errors in the secondary task (S2-R2) are observed.

Until recently these impairments in performance were mainly explained by capacity limitations at the response selection stage (for an overview see Pashler, 1994). It had been assumed that two reactions could not be selected simultaneously because the central stage of response selection is limited to process one event at a time. As a consequence, response selection of the secondary task has to wait until response selection of the primary task had been completed (so-called „locus-of-slack" logic, see Fig. 1). More recent dual-task studies found evidence that processing S1 not only impairs the selection of R2, but also impairs perceptual encoding of S2. Using a go/no-go manipulation in the primary task Müsseler and Wühr (2002) could show that performance in the secondary task is impaired even under conditions under which only identification of the stimulus but no response selection is required in the primary task (see also Jolicœur, 1999). Moreover, by introducing compatibility relationships between tasks (for an overview see Lien & Proctor, 2002), it was shown that compatibility information could successfully bypass the response-selection bottleneck (Hommel, 1998; see also Müsseler, Koch, & Wühr, 2005; Müsseler, Wühr, & Umiltà, 2006). Based on these findings, it has to be assumed that sensory, cognitive as well as motoric processes on a central level are susceptible to interference in dual-task situations.

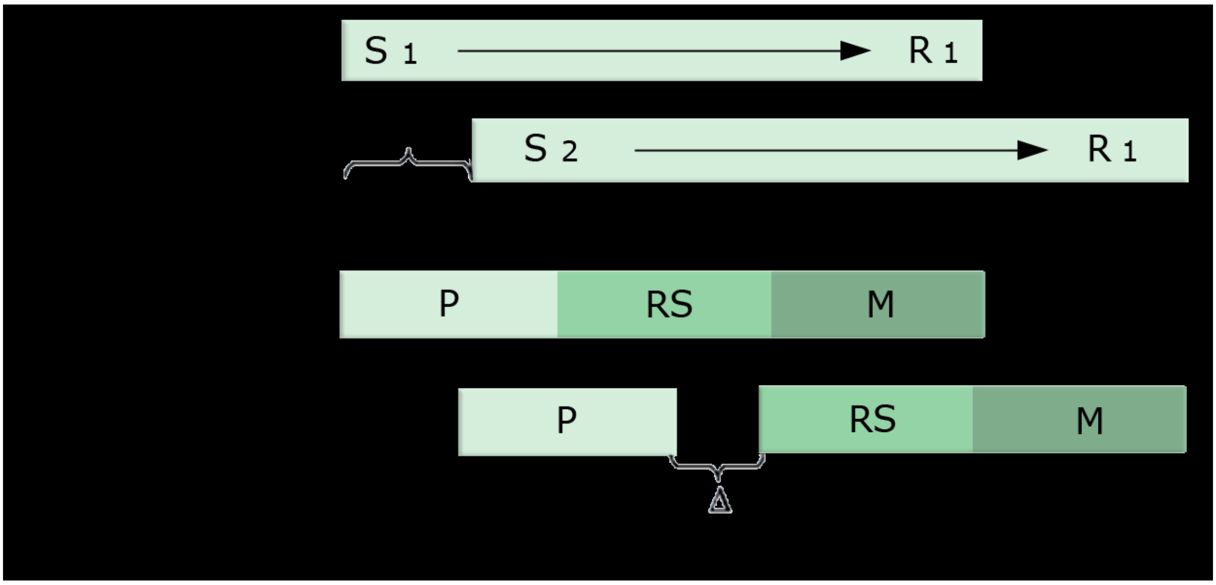


Fig. 1 (A) Basic dual-task situation in the psychological refractory period paradigm (prp task). Participants are required to react to two successively presented stimuli (S1 and S2) with two different responses (R1 and R2). (B) Typical explanation based on the locus-of-slack logic. Response selection (RS) of the second task has to wait until RS of the first task terminated. As a result reaction times in the second task are lengthened by Δt (P = perceptual processing, RS = response selection, M = motor processes, SOA = stimulus onset asynchrony).

Concering dual-task effects in driving situations, it has been shown in various studies that, for example, using a mobile phone when driving distracts the driver (for a recent overview see Dragutinovic & Twisk, 2005): It not only causes physical distraction (when drivers have to simultaneously operate their mobile phone and operate their vehicle) but also cognitive distraction. Cognitive distraction occurs when a driver has to divert part of his/her attention from driving to the telephone conversation. However, it has been shown in laboratory research (see previous paragraphs) that the ability to divide one's attention between two simultaneous tasks is limited. Mobile phone use while driving could, therefore, negatively affect driving performance. Consequently, the results of several studies strongly suggest that using a mobile phone while driving can increase the risk of being involved in a road crash up to four times (Redelmeier & Tibshirani, 1997). In a simulated driving task study, Strayer and Johnston (2001) showed that telephone conversations resulted in a significant increase in reaction time to simulated traffic signals. Moreover during conversation, participants missed twice as many signals. Similarly, Tornros and Bolling (2005) reported impaired reaction time as well as increased amount of errors both during mobile phone dialing and conversation in a peripheral detection task, in which participants had to respond to a light stimulus that appeared in the participant's periphery with respect to the main driving focal point. In general, the conclusions of these studies are that the use of mobile phones negatively affects different aspects of a driver's performance: Reactions to traffic signals are slower, braking reactions are slower with shorter stopping distances, drivers miss more important traffic signals, they are inclined to riskier behaviour like accepting shorter gaps or making fewer speed adjustments or adjustments to dangerous road conditions.

In the present context, it is important to note that performance in dual-task situations decreases and susceptibility to interference in compatibility tasks increases in elderly people (Verhaegen, Steitz, Sliwinski, & Cerella, 2003; Pick & Proctor, 1999; Spieler, Balota, & Faust, 1996). It has been demonstrated in many experimental studies that dual task conditions are in particular demanding for elderly people. The age-related increase in interference has been commonly explained by a general slowing of cognitive processes, especially of attentional processes and executive functioning (e.g. Hartley & Little, 1999; Reuter-Lorenz, 2002; Salthouse & Somberg, 1982; but see also Lindenberger & Baltes, 1994). Based on a meta-analysis of 34 studies on the effects of age on dual-task performance Riby, Perfect, and Stollery (2004) conclude that with increasing age tasks with a substantial controlled processing component showed greater dual task impairment than tasks that were relatively simple or relied on automatic processing. The latter tasks could be performed without a substantial interfering influence from the secondary task.

In many cognitive functions that are important for driving a car a reduced performance can be found with increasing age. A significantly reduced rate of information processing and, thus, an increase in general reaction time can already be observed at the age of 45 years (Stenneken, Aschersleben, Cole, & Prinz, 2002; Stern, Oster, & Newport, 1980). Especially with increasing complexity of the tasks disproportionate impairments in the performance of elderly people are obtained (Dobson, Kirasic, & Allen, 2002; Kliegl, Krampe, & Mayr, 2003; Li, Lindenberger, Hommel, Aschersleben, Prinz, & Baltes, 2004). The ability to discriminate relevant from irrelevant information is most impaired in elderly people if they have to perform the task under time pressure (Plude & Hoyer, 1986). However, the ability to select information in a fast and efficient way is especially important when driving a car.

Thus, a whole range of psychological (e.g. reduced memory span, reaction time, divided attention, sustained attention) as well as physiological (reduced vision and hearing abilities, reduced fine motor skills) functions are impaired with increasing age, which should have a negative influence on the roadworthiness. Nevertheless, accident statistics draw a different picture (Cohen, 2001). The risk to cause an accident is rather high immediately after the successful application for a driving license, then declines (between the age of 18 to 25 years) and is constant until the age of 65 years. Only after the age of 65 years, risk to cause an accident slightly increases again (HUK-Verband, 1994). The reasons for this somewhat unexpected pattern might be found in the use of compensatory strategies in elderly people. Older drivers avoid critical situations like rush-hour traffic, driving at night or under bad weather conditions (Hartenstein, 1995; Stutts, 1998). They do not drink and drive, use well-known routes and drive at slower tempi (Cox & Cox, 1998). Finally, a shift towards more serial operation of controls can be observed that probably represents a compensatory mechanism allowing older drivers to maintain their level of performance (Hakamies-Blomqvist, Mynttinen, Backman, & Mikkonen, 1999). In sum, situations requiring an increased level of cognitive effort are avoided. The most frequent causes of an accident in older drivers are driving errors made at intersections and junctions, rear-end collisions and overlooked traffic signs (Cox & Cox, 1998; Praxenthaler, 1995). This again suggests an increase of accidents in situations with cognitive overload. On the other hand, older drivers only rarely cause accidents as a result of driving while intoxicated, because of speeding or inadequate overtaking (Praxenthaler, 1995).

It seems that older drivers develop compensatory strategies that allow them to compensate for their sensory, motoric as well as cognitive shortcomings (e.g. Sommer, Falkmer, Bekiaris, & Panou, 2004). Older drives do not necessarily drive worse, however, as it is not always possible to avoid problematic and demanding traffic situations (e.g. driving at night) elderly people drive with higher cognitive load. Thus, it is important to study the influence of this fact on reactions, in critical situations in which a coordination of different tasks is required. As critical situations cannot always be avoided and, therefore, elderly people might also get into driving situations, in which their compensatory strategies do not work, the aim of the present research is to, first, determine those situations that are especially impairing for elderly people and, second, to make these situations less dangerous by suggesting appropriate designs.

First Results

While driving a car the driver often has to immediately react to visual information that is spatially localized (e.g. avoiding obstacles that suddenly appear at the right side of the road). From compatibility research it is well known that speed as well as accuracy of a spatial reaction is influenced by the spatial location of the imperative stimulus. A stimulus that is presented ipsilateral (at the same side at which the reaction is required) results in faster responses and less errors than a stimulus that is spatially non-corresponding with the required reaction. In classical compatibility theories, this pattern of results is explained by an automatic activation of the corresponding reaction that in the case of an incompatible situation has to be inhibited first before the correct response can be performed (Kornblum, Hasbroucq, & Osman, 1990). The question about the influence of compatibility effects in driving situations has been mostly ignored in research. Exceptions are studies by Wang and colleagues. Wang, Proctor, and Pick (2003) analyzed the influence of warning signals in a collision avoidance system on steering responses. Participants were told that tones presented to the left or right ear were warning signals from a collision avoidance system. They were instructed that the signals indicated the location of the danger source, from which they were to turn away, or the escape direction, toward which they were to turn. Spatial compatibility effects predict that it would be most beneficial to have the tone correspond to the desired response direction. Due to the stimulus valence of warning signals - you typically turn away from sounds created by hazards – an avoiding behavior, i.e., turning away from the warning tone, might be more compatible than responding toward it. Wang et al.'s (2003) results showed a typical spatial compatibility effect, suggesting that spatial compatibility was the primary factor influencing performance. More recently, though, Wang, Pick, Proctor, and Ye (2007) found that when participants were engaging in a simulated driving scenario, they were faster at turning away from the collision avoidance signal rather than toward it. This latter result suggests that, when driving, humans' responses to warning signals may follow valence compatibility principles – and not the principles of spatial compatibility. However, response latencies in this study were long – in the range of 4 seconds – because participants typically waited until they perceived the encroaching vehicle to respond, implying that the results may have been due to participants directing attention in the direction of the impending threat to facilitate its detection.

The aim of our first experiments was to analyze whether a critical situation while driving a car (e.g. a pedestrian suddenly entering the road from the left or right side) results in a reversal of classical compatibility effects, similar to what had been reported by Wang et al. (2003). Classical compatibility theories would imply that first an automatic activation of the corresponding reaction (i.e. in the direction of the dangerous stimulus) takes place, which then has to be inhibited before the correct response can be selected. The question is whether this automatic activation of the ipsilateral response is the same in driving situations.

In the first five experiments 144 participants aged between 18 und 65 years took part. In a darkened room participants were seated in a car seat in front of a steering wheel and foot pedals. Short videos were presented via beamer. In all experiments a taxi driver scenario was realized. In a simulated driving situation, participants watched short videos, in which they approached an intersection and a pedestrian entered the road from the left or the right side either calling the taxi by waving the arm or causing a critical situation by turning the back towards the driver. Participants were instructed to react as fast as possible by turning the wheel clockwise or counterclockwise either towards the location of the person or away from the person. Each trial was started by pressing a foot pedal.

In the first experiment, we tested whether the assumption of automatic ipsilateral response activation also holds in critical driving situations or if the compatibility effects depend on the meaning of the stimuli. Reaction time in the driving situation was compared with a control condition, in which neutral stimuli (diamond and square) were presented laterally. Here, participants also had to react as fast as possible by turning the wheel clockwise or counterclockwise. This first experiment was conducted with a group of young participants (mean age 21 years). Different to what had been predicted by classical compatibility theories reaction time in the critical situation (avoiding reaction; incompatible condition) was shorter than reaction to the person calling a taxi (compatible condition). In the control condition with neutral stimuli the usual compatibility effect (shorter reaction times with ipsilateral reaction) was observed. Thus, in critical situations the compatibility effects seem to be inverted indicating that the meaning of the stimuli is important (Müsseler, Aschersleben, Arning & Proctor, 2007; Exp. 1). This interpretation was supported by the following experiments. Under conditions under which the pedestrians were presented peripherally, they were probably no longer interpreted as being dangerous stimuli and, as a consequence, the usual compatibility effect was observed (Müsseler et al., 2007; Exp. 2). In a third experiment, we were able to replicate the results obtained in the first experiment and also compared them with the results of a group of older drivers (mean age 62 years; Aschersleben, Arning, & Müsseler, in prep.). As expected we observed longer reaction times in the elderly participants, however, there was no interaction between age and compatibility indicating that the compatibility effect was not increased in the group of older drivers (see Fig. 2). One possible explanation for this somewhat unexpected result might be the fact that older drivers also are more experienced in driving.

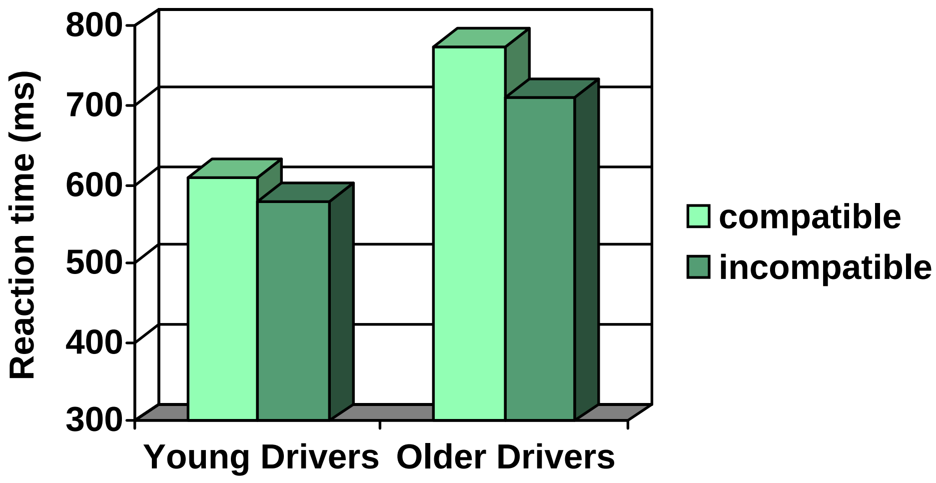


Fig. 2 Reaction times (in ms) for the two tested aged groups and the compatible vs. incompatible condition.

Next, we extended the experimental design to analyze the influence of dual-task situations in car driving. Participants heard a spoken message from the navigation system just before the imperative stimulus (pedestrian) was presented. Thus, the primary task required to react with a left or right steering-wheel response to a pedestrian suddenly entering the street whereas the secondary task required listening (and later react) to the message from the navigation system. As this message consisted in an instruction concerning the driving direction („turn left (right)“) and therefore also contained spatial information, we were mainly interested in cross-task compatibility effects between the primary and the secondary task. Results indicate that when the pedestrian enters from the right (left) side, a right (left) message from the navigation system impairs performance as it conflicts with the left (right) steering-wheel response to avoid hitting the pedestrian. In other words, superior performance is observed when the spoken message and the direction of the steering-wheel response corresponded than when they do not. When testing a group of elderly people, we observed increased reaction times as well as the cross-task compatibility effect - but again no interaction between age and compatibility, indicating that the compatibility effect was not increased in the group of older drivers (Aschersleben et al., in prep.). To test the assumption that the decreasing performance with age in the dual-task situation is compensated by expertise in the elderly people, we tested professional drivers. First results indicate a reduced cross-task compatibility effect in a group of elderly professional drivers indicating the indeed two factors might be interacting here: age and expertise.

Future Prospects

The aim of the present project is to analyze the specific cognitive load of drivers in dual task situations. First results indicate that in critical situations a clear influence of the secondary task (e.g. a message from the navigation system) on the reaction time in the primary task. In further experiments we will analyze the influence of different stimulus modalities (e.g. auditory vs. visual) of the secondary task. Moreover, until now we only studied so-called functionally dependent tasks, that is the information provided by the secondary task was relevant for the primary task. If the same pattern of results can be obtained in functionally independent tasks (e.g. manipulation of the car radio) is an empirical question, which will be analyzed in further studies.

Literature

Aschersleben, G., Arning, K., & Müsseler, J. (in prep.). *Age-related differences in critical driving situations: The influence of dual-task situations and S-R compatibility.* Manuscript in preparation.

Cohen, A. S. (2001). Leistungsanforderungen und –möglichkeiten der Senioren als Fahrzeuglenker. In A. Flade, M. Limbourg, & B. Schlag (Hrsg.), *Mobilität älterer Menschen* (S. 241-258). Opladen: Leske & Buldrich.

Cox, A. B., & Cox, D. J. (1998). Compensatory driving strategy of older people may increase driving risk. *Journal of the American Geriatrics Society, 46*(8), 1058-1059.

Dobson, S. H., Kirasic, K. C., & Allen, G. L. (1995). Age-related differences in adults' spatial task performance: Influences of task complexity and perceptual speed. *Aging & Cognition, 2*(1), 19-38.

Dragutinovic, N. & Twisk, D. (2005). *Use of mobile phones while driving – effects on road safety*. R-2005-12, SWOV Institute for Road Safety Research, The Netherlands.

Hakamies-Blomqvist, L., Mynttinen, S., & Backman, M. (1999). Age-related differences in driving: are older drivers more serial? *International Journal of Behavioral Development, 23*(3), 575-589.

Hartenstein, W. (1995). Das "Älterwerden" der Autofahrer-Population: Größenordnungen, Fahrgewohnheiten, Einstellungen, Auswirkungen. In J. Lindlacher (Ed.), *Ältere Menschen im Straßenverkehr. Schriftenreihe Straßenverkehr, Bd. 34* (pp. 16-24). München: ADAC.

Hartley, A. A., & Little, D. M. (1999). Age-related differences and similarities in dual-task interference. *Journal of Experimental Psychology: General, 128*(4), 416-449.

Hommel, B. (1998). Automatic stimulus-response translation in dual-task performance. Journal of Experimental Psychology: Human Perception & Performance, 24(5), 1368-1384.

HUK-Verband. (1994). *Fahrzeugsicherheit 90*. München: Büro für Kfz-Technik der Haftpflichtversicherer, Autoversicherer und Rechtschutzversicherer e. V.

Jolicœur, P. (1999). Dual-task interference and visual encoding. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 596-616.

Kliegl, R., Krampe, R. T., & Mayr, U. (2003). Formal models of age differences in task complexity effects. In U. Staudinger & U. Lindenberger (Eds.), *Understanding human development: Life-span psychology in exchange with other disciplines* (pp. 289-313). Dordrecht: Kluwer.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility: A model and taxonomy. *Psychological Review, 97*, 253-270.

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science, 15*(3), 155-163.

Lien, M. C., & Proctor, R. W. (2002). Stimulus-response compatibility and psychological refractory period effects: implications for response selection. *Psychonomic Bulletin & Review, 9*(2), 212-238.

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging, 9*(3), 339-355.

Müsseler, J. & Wühr, P. (2002). Response-evoked interference in visual encoding. In Prinz, W. & Hommel, B. (Eds.) *Attention and Performance XIX: Common Mechanisms in Perception and Action* (pp. 520-537). Oxford, UK: Oxford University Press.

Müsseler, J., Koch, I., & Wühr, P. (2005). Compatibility effects with irrelevant locations in dual task: A cross-task Simon effect. *European Journal of Cognitive Psychology, 17*, 708-72.

Müsseler, J., Wühr, P., & Umiltà, C. (2006). Processing of irrelevant location information under dual-task conditions. *Psychological Research, 70,* 459-46.

Müsseler, J., Aschersleben, G., Arning, K., & Proctor, R. W. (2007). *Spatial compatibility and affective valence compatibility in natural scenes*. Manuscript submitted for publication.

Pashler, H. (1998). *Attention*. Hove, UK, Psychology Press.

Pick, D. F., & Proctor, R. W. (1999). Age differences in the effects of irrelevant location information. In M. W. Scerbo & M. Mouloua (Eds.), *Automation technology and human performance: current research and trends* (pp. 258-261). Mahwah, NJ: Lawrence Erlbaum.

Plude, D. J., & Hoyer, W. J. (1986). Age and the selectivity of visual information processing. *Psychology & Aging, 1*(1), 4-10.

Praxenthaler, H. (1995). Besitz und Nutzung des Führerscheins in verschiedenen Altersgruppen. In J. Lindlacher (Hrsg.), *Ältere Menschen im Straßenverkehr. Schriftenreihe Straßenverkehr, Bd. 34* (pp. 24-36). München: ADAC.

Redelmeier D.A. & Tibshirani R.J. (1997). Association between cellular-telephone calls and motor vehicle crashes. *The New England Journal of Medicine, 336*(7), 453-458.

Reuter-Lorenz, P. A. (2002). New visions of the aging mind and brain. *Trends in Cognitive Sciences, 6*(9), 394-400.

Riby, L. M., Perfect, T. J., & Stollery, B. T. (2004). The effects of age and task domain on dual task performance: A meta-analysis. *European Journal of Cognitive Psychology, 16*(6), 868-891.

Salthouse, T. A., & Somberg, B. L. (1982). Isolating the age deficit in speeded performance. *Journal of Gerontology, 37*(1), 59-63.

Sommer, S. M., Falkmer, T., Bekiaris, E. & Panou, M. (2004). Toward a client-centred approach to fitness-to-drive assessment of elderly drivers, *Scandinavian Journal of occupational therapy, 11,* 62-29.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception & Performance, 22*(2), 461-479.

Stenneken, P., Aschersleben, G., Cole, J. & Prinz, W. (2002). Self-induced versus reactive triggering of synchronous movements in a deafferented patient and control subjects. Psychological Research, 66, 40-49.

Stern, J. A., Oster, P. J., & Newport, K. (1980). Reaction time measures, hemispheric specialization, and age. [References]. In L. W. Poon (Ed.), *Aging in the 1980s: Psychological issues* (pp. 309-326). Washington, DC: American Psychological Association.

Strayer, D.L. & Johnston, W.A. (2001). Driven to distraction; Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science*, *12*(6), 462-466.

Stutts, J. C. (1998). Do older drivers with visual and cognitive impairments drive less? *Journal of the American Geriatrics Society, 46*(7), 854-861.

Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology, 14*, 1-37.

Tornros, J. & Bolling, A. (2005). Mobile phone use – Effects of handheld and hands-free phones on driving performance. *Accident Analysis and Prevention*, *37*(5), 902-909.

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: A meta-analysis. *Psychology & Aging, 18*(3), 443-460.

Wang, D.Y.D., Proctor, R. W., & Pick, D. F. (2003). Stimulus-response compatibility effects for warning signals and steering responses. *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 226-230.

Wang, D,-Y, Pick, D. F., Proctor, R. W., & Ye, Y. (2007). Effect of a side collision avoidance signal on simulated driving with a navigation system. *Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*, Stevenson, WA, 206-211.

Solving Systems of Equations by Analogy: An Experiment and a Model

SVETLANA POLUSHKINA
Research Training Group "Feedback Based Quality Management in eLearning"
University of Technology Darmstadt[1]

KAI-UWE KÜHNBERGER, HELMAR GUST
Institute of Cognitive Science
University of Osnabrück

In this case study, the Heuristic-Driven Theory Projection framework of analogical reasoning was tested for its applicability to problem solving in a sub-domain of mathematics. The predictions of a computational model were checked against the results of an experiment with human participants. The study supports the viability of heuristic anti-unification as a powerful mechanism for analogy making.

Introduction

Analogical reasoning has a long-standing tradition as a research field for both psychologists and computer scientists (for a review, see e.g. O'Donoghue, 2004). This work is a case study in a sub-domain of mathematics, testing the applicability of the Heuristic-Driven Theory Projection framework (HDTP; Gust, Kühnberger & Schmid, 2003) to analogical problem solving. HDTP proposes a generalisation-based model for analogy, which employs the anti-unification algorithm to produce mappings and transfer suggestions between a source and a target. In the present study, the source and the target contained systems of higher-order polynomial equations in two variables, as shown in Figure 1.

SOURCE *(solution is given)*	TARGET *(solution is to be produced)*
$4a^2 + 4ab + b^2 = 0$ && $a - b = 6$	$9m^4 + 24m^2n + 16n^2 = 0$ && $m^2 - n = 7$

Figure 1: The working example

The sub-domain of systems of polynomial equations was chosen due to the highly structured representations of the expressions involved and possibilities to construct an infinite number of equivalent expressions, or equations. Clearly, some special domain knowledge is indispensable to solve problems in such a domain. The more knowledge, however, is available to the problem solver, the more search is required to produce the solution - and the more helpful is the use of analogy to reduce the search space. Importantly, the two systems of equations used in the working example are of different order and not fully isomorphic. As obvious from Figure 2, the solution trace of the target system of equations, though, practically

[1] Address: Svetlana Polushkina, Institute of Mathematics, University of Technology Darmstadt, D-64289 Darmstadt, email: polushkina@mathematik.tu-darmstadt.de. The Research Training Group "Feedback Based Quality Management in eLearning" is founded by the German Research Foundation (DFG).

totally parallels the solution trace of the source system. Investigating analogical problem solving on such an example seems reasonable and promising.

After step	Source (given)	Target (to be produced)
1.	$4a^2 + 4ab + b^2 = 0 \&\& a = 6 + b$	$9m^4 + 24m^2n + 16n^2 = 0 \&\& m^2 = 7 + n$
2.	$(2a + b)^2 = 0 \&\& a = 6 + b$	$(3m^2 + 4n)^2 = 0 \&\& m^2 = 7 + n$
3.	$2a + b = 0 \&\& a = 6 + b$	$3m^2 + 4n = 0 \&\& m^2 = 7 + n$
4.	$2 * (6 + b) + b = 0 \&\& a = 6 + b$	$3 * (7 + n) + 4n = 0 \&\& m^2 = 7 + n$
5.	$12 + 2b + b = 0 \&\& a = 6 + b$	$21 + 3n + 4n = 0 \&\& m^2 = 7 + n$
6.	$12 + 3b = 0 \&\& a = 6 + b$	$21 + 7n = 0 \&\& m^2 = 7 + n$
7.	$3b = -12 \&\& a = 6 + b$	$7n = -21 \&\& m^2 = 7 + n$
8.	$b = -4 \&\& a = 6 + b$	$n = -3 \&\& m^2 = 7 + n$
9.	$b = -4 \&\& a = 6 + (-4)$	$n = -3 \&\& m^2 = 7 + (-3)$
10.	$b = -4 \&\& a = 2$	$n = -3 \&\& m^2 = 4$

Figure 2: Source and target solutions to the working example

The following research questions were investigated: How does an analogical problem solver work? What is the relation between analogy and structural similarity? What role do generalisations play? What background knowledge and heuristics are useful? What kind of additional information is helpful and how it should be presented to promote successful analogical problem solving? Methodologically, an experiment with human participants was carried out and a computational model of a domain-specific problem solver was constructed. Both, in the experiment and in the model, the amount and kind of provided information were manipulated and the influence of those manipulations on problem solving was analysed. The experimental and modelling results were finally compared to check the validity of the model.

Theoretical Framework: The Heuristic-Driven Theory Projection

The Heuristic-Driven Theory Projection (HDTP) is a formally sound theoretical framework for analogical reasoning. HDTP is based on the so called anti-unification (Reynolds, 1970), which provides an algorithm for computing generalisations of the source and the target for establishing an analogical relation (Schmid et al., 2003). The anti-unification algorithm maps the entities in the source and the target via the roles they play in their respective domains, or in the structures of their respective expressions. Importantly, the most special, i.e. the least general generalisation is computed as a result of the application of the anti-unification algorithm. Such a generalisation preserves the necessary maximum of the structural commonalities between the source and the target. Apart from the generalisation, substitutions allowing for reproducing the source and the target from their generalisations are also determined by the anti-unification algorithm. In a computational model, such an approach can be combined with explicitly storing the common structure of the source and the target domains in the 'working memory of the problem solver'.

Additionally to the classical anti-unification of terms, HDTP also allows for anti-unifying predicates, formulae and even entire theories. In solving a realistic mathematical problem, such a capacity is certainly advantageous.

Furthermore, the anti-unification algorithm can be extended to promote the order of the underlying language, namely toward higher-order generalisation. Such an extension, however, is not trivial. Obviously, the goal of introducing the second and higher order into the anti-unification algorithm is to provide for adaptation of non-isomorphic structures in a way, preserving the maximum of their commonalities. Such an adaptation can include changing the order of sub-terms of those structures. Moreover, many-to-one mappings should be possible in an adaptation process like this. In general, the anti-unification algorithm producing a generalisation with many-to-one correspondences of its elements to the respective elements in the source and / or the target can include both operations of deletion and insertion of sub-terms. In a second- or higher-order case, anti-unification allowing for both deletion and insertion of sub-terms can lead to infinite sets of the most special generalisations. To avoid this effect, the variant of the higher-order anti-unification algorithm employed in HDTP allows for insertion, but not for deletion of expressions and sub-expressions (Hasker, 1995). Such restricted higher-order anti-unification yields a finite set of the most specific generalisations. To reduce this set to just one most specific generalisation, additional heuristic criteria, specifying particular preferences on the choice of generalisation, can be integrated into the algorithm.

Thus, another direction of extending the classical anti-unification within HDTP concerns integrating various kinds of background knowledge, or heuristics into the algorithm. Such knowledge can relate to equivalence classes of differently represented expressions. Moreover, an ontology providing information on the sorts of entities contained in the anti-unified domains can be included. Such ontology in the case of mathematics can involve the sorts like a number, a variable or an operator. HDTP makes use of flat ontologies (Gust et al., 2005).

Additionally to the mapping and transfer (problem solving) phases of analogy, HDTP also covers the evaluation and the consolidation (learning) phases. In the evaluation phase, a so called oracle validates the transferred facts and inferences for the target domain and by doing this detects possible inconsistencies (Gust and Kühnberger, 2006). In mathematics, especially in solving systems of equations, some sort of testing the obtained results is vital: Obviously, in such a domain there is no use in analogy, unless it leads to correct solutions.

With all its properties, Heuristic-Driven Theory Projection lends itself as a framework for modelling an analogical problem solver. HDTP has been already applied to modelling creativity and learning in naïve physics and to modelling human understanding of metaphors. In this work, HDTP is applied to problem solving in mathematics.

Experiment

54 male and female pupils of the 8th grade of the gymnasium took part in the classroom experiment. They received handouts with the solution of the source example problem and were asked to solve the target example problem. There were two variants of presentation of the source problem solution, with or without explanations on the derivations of the source solution steps (*Action Explanations Factor*). For example, for the seventh solution step (see Figure 2) the explanation that subtraction by 12 took place could be given additionally to the presentation of the equations themselves. Moreover, some pupils were provided with additional hints on mappings between corresponding equations in the source and target problems (*Hint Factor*). There were two sorts of such hints: first, hierarchical mapping hints containing explicit generalisations of the mapped terms, and second, flat mapping hints, establishing direct relations between mapped terms (*Flat / Hierarchical Mapping Hint Factor*). Both kinds of hints concerned the pairs of analogous equations in the source and the target. For the shorter pair of equations those hints are given as they were presented to the students, but in English translation in the Figures 3 and 4. In total, there were six experimental conditions, to which the participants were assigned taking into account equal distribution of math grades and gender over the groups.

Hint 1

Compare the shorter equations of the two problems. Parts of those equations exactly correspond to each other. Do you see, how? Here are the correspondences from the left to the right:

Example	Problem to Solve
a – b = 6	m² - n = 7
a	m²
b	n
6	7

Use these correspondences while solving the problem!

Figure 3: The Flat Mapping Hint for the shorter pair of equations

Hint 1

Compare the shorter equations of the two problems. Those equations are both of the form

X – Y = Number,

where X and Y stay for some letters. Do you see, how the parts of those two equations respectively correspond to this general form? Here are the correspondences from the left to the right:

Example	General Form	Problem to Solve
a – b = 6	X – Y = Number	m² - n = 7
a	X	m²
b	Y	n
6	Number	7

Use these correspondences while solving the problem!

Figure 4: The Hierarchical Mapping Hint for the shorter pair of equations

Data on time consumption and solution correctness were obtained and evaluated. In any experimental group, there were participants who produced correct solutions. The only effect on time consumption detected was its significant reduction under the influence of the *Hint Factor* (t (50) =1.8183, p<.05). The absence of the effect of the *Action Explanations Factor* on the time consumption suggests practically automatic inference of the derivations of the solution steps in the source example. Interestingly, there were significantly more structurally correct solutions (77.78%), than just correct answers (48.15%; χ^2 = 10.1647, p<.005). Participants provided with the hierarchical mapping hints made more miscalculation mistakes in solving the target problem, whereas those provided with the flat mapping hints made more conceptual mistakes in their solutions. The latter qualitative result suggests that the generalisation-based mapping hints helped the participants to discover the essence of the example problems, but put more information load on them.

Furthermore, the participants were asked to make similarity preference choices immediately before learning about the solution of the source problem and after solving the target problem. For this, they were presented with one of the example systems of equations, the source or the target one, respectively, and two other systems of equations, a structurally equivalent one and one containing the same numbers and variables. An example of such a similarity assessment question is depicted in Figure 5. Structural similarity was preferred much more often (87.96% of choices), than superficial similarity (12.04%). Changing similarity preferences from structural to superficial correlated with solving the target problem incorrectly.

Which of the two systems of equations at the bottom is more similar to the upper one? Please tick off your choice!

$$4a^2+4ab+b^2=0$$
$$a-b=6$$

$25s^2+10st+t^2=0$ $s-t=3$	$4+a^2-4ab-b^2=0$ $ab=6$

Figure 5: An example for the similarity assessment question

The results of the experiment emphasise the role of structure-preserving mappings with minimal level of abstraction (information load) for successful analogical problem solving. In educational contexts, generalisation-based hints on corresponding parts of the source and target promote analogy making by the students. Additionally, showing solution path derivations, even without details on the specific applied mathematical operations, helps induce inferences on step transitions, which supports the transfer of the solution structure from the source to the target problem.

Computational Model

To test the predictions of the HDTP framework, a computational model of the problem solver for the sub-domain of mathematics was implemented in PROLOG. The model employs many-sorted higher-order anti-unification to produce structure-preserving least-general generalisations of source and target systems of equations. The model obeys to the systematicity principle (Gentner, 1983), allowing at the same time for flexible many-to-one mapping (Owen, 1990). Moreover, it favours sort-preserving mappings and produces abstractions of arbitrarily high, but minimal order to reduce the ‘information load’ on the analogy maker.

The basic architecture of the model is depicted in Figure 6. Principally, the model contains two executive modules, the Anti-unifier and the Problem Solver, and a data base for the Solution History. The executive modules are responsible for different sub-process of analogical problem solving, like mapping and transfer. When needed, they contact the Solution History to require or update the contemporarily valid anti-unification results and preferably a list of already undertaken actions. The contents of the model components and the interaction flow among them are sketched in Figure 6.

Importantly, the model accounts for equivalence classes of algebraic expressions, which allows for problem re-representation. This means that the model is capable of representing expressions in a variety of equivalent ways, for example, if a change in the problem ‘perception’ is required for the problem solving process. In case the Problem Solver encounters hard obstacles, while trying to transfer a step from the source to the target problem, i.e. if there is no applicable step producing rule in the knowledge base, a suggestion of how to re-formulate the involved expressions to make one of the rules apply is generated by the Problem Solver and forwarded to the Anti-unifier with a request to generalise over these new expressions. In the working example such a re-representation procedure gets necessary in the second solution step, where the first binomial formula should be applied. The Problem Solver runs into trouble, because the initial anti-unification of the longer equations has produced the following source-to-target mappings for the left most pair of terms: [4 & 9], [a & m] and [2 & 4]. The application of the binomial formula, however, would get possible with the other mappings: [4 & 9], [a^1 & m^2] and [2 & 2]. Thus, the attempt to apply the solution step requires for the re-representation, which then gets carried out by the model.

By this, some dead ends can be avoided. Furthermore, the model keeps track of many aspects of the problem solving process, such as solution steps' inputs and outcomes, order and producing operations. In utilises domain-specific background knowledge and effective heuristics. The employed similarity measure rewards establishing sub-structures shared between source and target and penalises introducing new structural elements.

The model is capable of solving the example target problem, given the solution of the source problem. Additionally, the Anti-unifier can autonomously work on a number of classes of algebraic expressions.

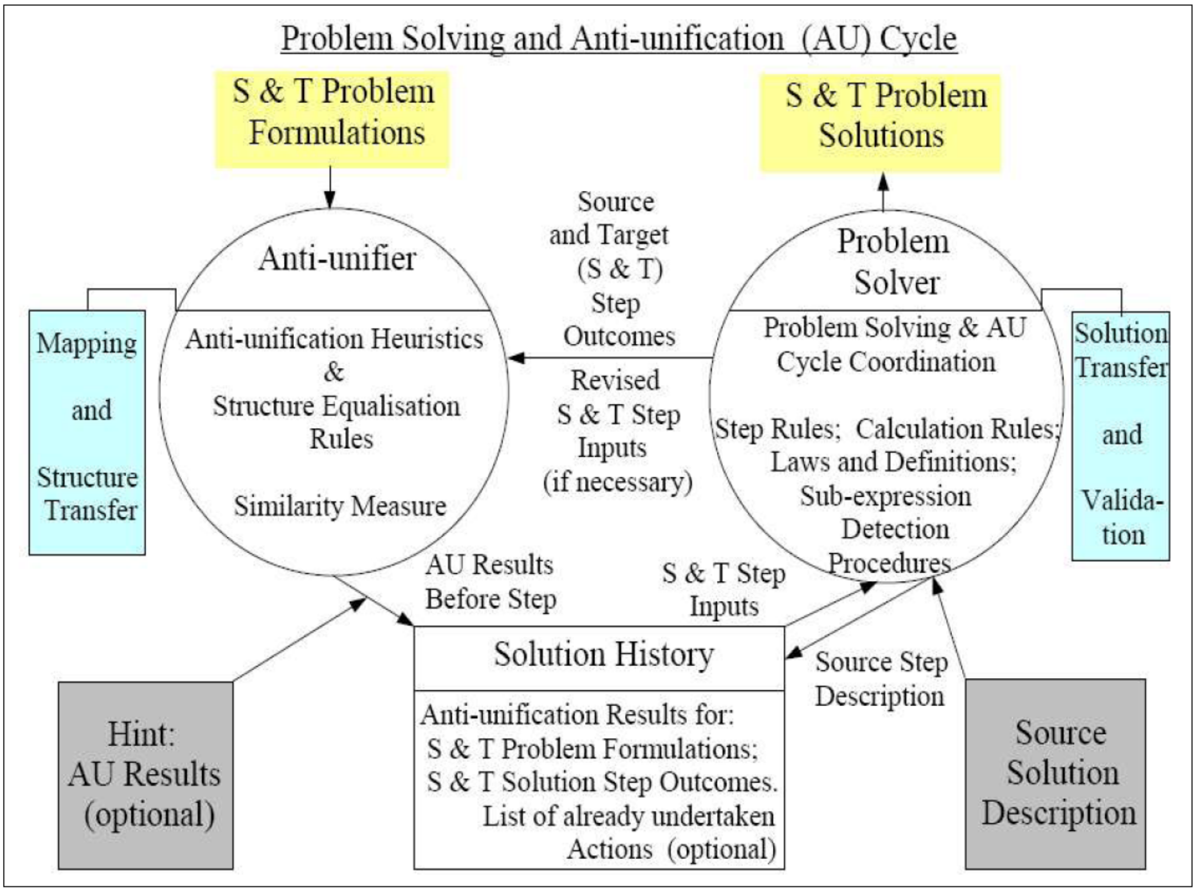


Figure 6: A model of the analogical problem solver

Information provided to the model was manipulated in a way equivalent to the manipulations in the experiment with human participants. First, some anti-unification results were provided to the model (*Hint Factor*). Second, information on the step derivations (applied mathematical operations, not complete step producing rules) was sometimes excluded from the descriptions of the source solution steps (*Actions Input Factor*). By this, the information provided on the source solution steps was reduced from [Input, Outcome, Operator] to just [Input, Outcome].

Finally, the storage format for the results of anti-unification was varied (*Flat / Hierarchical Generalisation Outcome Factor*). In total, six versions of the model were constructed, five of which were capable of mastering the working example.

Time consumption and number of inferences necessary for solving the target example problem were compared for different versions of the model. Significant reduction in time consumption was detected for the *Hint Factor* ($t(18) = 176.79$, $p<.0005$), which was the only effect on time consumption. Moreover, providing mapping hints seemed to reduce the number of necessary inferences (*Hint Factor*), especially if those hints were given in a minimal-information form (*Flat / Hierarchical Generalisation Outcome Factor*). No significant effect of the *Actions Input Factor* could be found, suggesting that describing concrete mathematical operations in the solution path derivation is less essential for promoting analogical problem solving, than establishing correspondences between the source and the target.

Computationally, the application of the model and the results of its manipulation emphasise the role of producing structure-preserving mappings and storing results in a format minimising information load. Keeping track of the problem solving history and utilising hints on analogical mapping appear to dramatically reduce the computational cost for problem solving by analogy. The capacity to re-represent the involved source and target expressions, during the problem solving or already in the course of the anti-unification process, substantially broadens the applicability area of the model.

Results

In this study, the results of the computational modelling parallel those of the experiment with human participants. In both cases, hints on analogical mapping significantly reduced the duration of successful problem solving. In both cases, no effect of providing information on the particular mathematical operators applied to produce the consecutive source solution steps could be detected. The miscalculation mistakes, made relatively more often by the pupils who received the hierarchical mapping hints, as compared to those with the flat mapping hints, can be set in relation to the increased number of inferences necessary for the model with hierarchical generalisation outcome format, as opposite to the flat one.

The latter would support the interpretation of “more abstraction as more information load”. Minimal abstraction, however, seems important, as suggested by a reduced number of conceptual mistakes in the solutions produced by pupils given generalisation-based mapping hints, as compared to the pupils provided with hints on direct mapping between pairs of terms.

This work contributes to the broad research fields of analogical reasoning and problem solving. On an example of a particular case of analogically solving systems of higher-order polynomial equations, a model of analogical problem solving within the Heuristic-Driven Theory Projection was put forward in this work.

The particular focus of this study lay on the mapping and transfer phases of analogical problem solving. The proposed model can be characterised by a number of remarkable properties. The models developed within other approaches to analogy usually feature one or more of those properties, but their combination appears interesting. Being based on the anti-unification algorithm, the presented model uses generalisations for analogical mapping. The generalisations produced by the model obey to the systematicity principle in preserving as much of the structure shared by the source and the target, as possible. Such generalisations help, for example, reducing the information load put on the problem solver. The model is capable of producing flexible many-to-one mappings and is able to anti-unify terms, predicates, formulae and entire theories. The model accounts for equivalence classes of algebraic expressions and features a number of rewriting rules. Those rules allow, for example, for re-representing a problem formulation, if its ‘perception’ has been changed due to the problem solving process. The model keeps track of the problem solving process. It utilises all the information contained in a derivational path of a source solution, from the step inputs, over their order, over the operators that have produced the solution steps, to their outcomes. The model uses some heuristic background knowledge and a flat ontology to make the search for the problem solution more efficient. Finally, due to all these properties, the model can analogically solve relatively complex problems in a sub-domain of mathematics.

Correspondence between the empirical and computational results suggests the adequacy of using the model, and HDTP in general, for providing insights into the analogical problem solving in humans. Essentially, structure-preserving mappings between source and target support successful analogical problem solving. Furthermore, anti-unification producing least-general generalisations of source and target offers an effective mechanism for analogy making.

The comparison of the results of the experiment with human participants with those of the computational modelling shows the applicability of the Heuristic-Driven Theory Projection framework to analogical problem solving, as exemplified for a sub-domain of mathematics. The results of the case study support the utility of anti-unification as a mechanism and a part of a model for analogical mapping and transfer, the two central processes of analogical reasoning.

Acknowledgements

This work was carried out in preparation of the Master’s Thesis of the first author. The studies of the first author were supported by the German National Merit Foundation (Studienstiftung des deutschen Volkes).

The results of this study, especially the empirical ones, have been partly presented in (Polushkina, Kühnberger & Gust, in press).

References

Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. *Cognitive Science*, 7, 155-170.

Gust, H., Kühnberger, K.-U., & Schmid, U. (2003). Solving predictive analogy tasks with anti-unification. In *Proceedings of the Joint International Conference on Cognitive Science 2003 (AMCS / ICCS 2003)*, Sydney.

Gust, H., Kühnberger, K.-U., & Schmid, U. (2005). Ontologies as a Cue fort the Metaphorical Meaning of Technical Concepts.

Gust, H., & Kühnberger, K.-U. (2006). The Relevance of Artificial Intelligence for Human Cognition. In *AAAI Spring Symposium Series 2006.* Symposium: Between a Rock and a Hard Place: Cognitive Science Principles Meet AI-Hard Problems, Stanford, pp. 64-69.

Hasker, R. W. (1995). *The replay of program derivations.* Doctoral dissertation, University of Illinois at Urbana Champaign.

O’Donoghue, D. (2004). *Finding Novel Analogies*. PhD Thesis. Department of Computer Science, University College Dublin.

Owen, S. (1990). *Analogy for Automated Reasoning*. Academic Press, Inc.

Polushkina, S., Kühnberger, K.-U., & Gust, H. (in press). Solving Systems of Equations: A Case Study in Analogical Problem Solving. In *Proceedings of the Second European Cognitive Science Conference (EuroCogSci07)*, Delphi, Greece.

Reynolds, J. C. (1970). Transformational systems and the algebraic structure of atomic formulas. *Machine Intelligence*, Vol. 5, pp. 135-151. Edinburgh University Press.

Schmid, U., Gust, H., Kühnberger, K-U., & Burghardt, J. (2003). An Algebraic Framework for Solving Proportional and Predictive Analogies. In *Proceedings of the European Conference on Cognitive Science (EuroCogSci03)*, Osnabrück, Germany.

Unsupervised learning of reflexive and action-based affordances to model navigational behavior

DANIEL WEILLER[1] , LEONARD LÄER[1] , ANDREAS K. ENGEL[2] , PETER KÖNIG[1]

[1] Institute of Cognitive Science
Dept. Neurobiopsychology
University of Osnabrück
49069 Osnabrück
Germany

[2] Dept. of Neurophysiology and Pathophysiology
University Medical Center Hamburg-Eppendorf
20246 Hamburg
Germany

Abstract

Here we model animat navigation in a real world environment by using place cell as a sensory representation. The cells' place fields divided the environment into discrete states. The robot learns knowledge of the environment by memorizing the sensory outcome of its motor actions. This was composed of a central process, learning the probability of state-to-state transitions by motor actions and a distal processing routine, learning the extent to which these state-to-state transitions are caused by sensory-driven reflex behavior (obstacle avoidance). Navigational decision making integrates central and distal learned environmental knowledge to select an action that leads to a goal state. Differentiating distal and central processing increases the behavioral accuracy of the selected actions. We claim that the system can easily be expanded to model other behaviors, using alternative definitions of states and actions.

Introduction

Navigation refers to the practice and the skill of animals as well as humans to find their way and to move from one place to another by any means (Wilson and Keil, 1999). The ability of animals to navigate in essentially two-dimensional maze environments has been studied extensively (Olton and Samuelson, 1976; Morris, 1984). Navigation involves cognitive processes like sensory processing, actions execution and decision-making. Here we propose a cognitive model of these processes implemented on a robot faced with the task to navigate in a four-arm-maze environment. To solve this task the robot learns the sensory outcome of its actions, thus acquiring of environmental properties. This knowledge was used to plan and execute actions to solve the navigational task. We claim that the introduced cognitive model is not restricted to a navigational behavior and can easily be generalized to model other behavior.

Navigation can be defined by executing the appropriate action at the right location in an environment to move towards a goal. In the rodents hippocampus O'Keefe (O'Keefe et. al., 1971) found place cells, which encode the position of the animal. These cells fire only when the animal is located in a certain region of the environment, defined as the cell's place field. Although the contribution of these cells to the animal's behavior has still not been fully understood, it is assumed that these cells constitute a cognitive map (O'Keefe and Nadel, 1978) of the environment and are, thus, the bases of navigation. Wyss and coworkers (Wyss et al., 2006) have recently shown that place cells can be understood as an optimally stable representation of the visual input of a behaving robot in a hierarchical network. This implies that unsupervised learning of the sensory input results in a reorganization of the sensory space, spanned by its visual input, which has a spatial meaning. We used this fact by using place cells to locate the robot in its environment. We simplified this task by approximating the firing properties of place cells by Gaussian function and distributed the corresponding place fields in the whole environment. These place fields correspond to the robots internal states and represent the positions between which it is able to differentiate. Hence, in order to enable the cognitive model controlling the robot to perform navigational behavior, we chose place cells as the representation of the environment.

To navigate in the environment the robot first has to learn the sensory outcome of its actions and second to plan its actions according to its knowledge. Learning and planning are both done in its state space, spanned by the place fields. The robot learns local state transitions caused by its action execution. Because the execution of the same actions in a state can result in a transition to different states, the information gained from these local transitions is stored as transition probabilities in a probabilistic directed graph. The robot has also to avoid obstacles. We implemented a reflexive obstacle avoidance behavior controlled by the robots proximity sensors. In case the robot used its reflexive behavior during a transition between two states, we memorized the occurrence of such an event in so called reflex factors. Here the architecture of the cognitive model differentiated between central processing, responsible for state transition memorization and distal processing, responsible for reflex factor learning. The transition probabilities and the reflex factors reflect the environmental properties in relation to the robot's actions. Thus by random action execution, the robot learns an approximation of the environmental affordances (Gibson, 1977), defined the action possibilities afforded by the environment. The robot plans goal directed actions by integrating the information gained by central and distal processing in a local decision-making process. This integration results in a quantitative measure how reliably each executable action leads towards the goal. Overall, the key components of our cognitive model are (i) a high-level representation (place fields) of sensory input space, (ii) the knowledge of environmental properties acquired by active exploration of local state transitions and (iii) a decision-making process driven by this knowledge.

Here we show that using the described cognitive model, a robot can successfully navigate to different goals within a four-arm-maze environment. As expected, the differentiation between central and distal processing reduces the negative effect of the obstacle avoidance behavior on navigational performance. We claim that by redefining the states and actions the introduced model can be expanded to model other behaviors.

Methods

Overview of the architecture

Our cognitive model learns the properties of the environment and plans its action to move towards a goal, based on the state space which is spanned by the spatial representation of place field (state). We divided the four-arm-maze environment (Fig. 1A) into compact discrete states (Fig. 1B) similar to place fields. The architecture of the cognitive model consisted of central and distal components. The central component captures the transition between states, caused by the robots action execution in these states. In contrast the distal component accounts for the usage of distal sensors, like infrared-sensors, facilitating obstacle avoidance during the robots state transition. Here the obstacle avoidance behavior is defined as reflexive behavior. While the central component accounts for any of the robots transition, the distal component constitute only transitions combined with reflexive behavior. Thus the transitions and the transitions combined with reflexive behavior represent the robots locally learned environmental properties according to the robots actions. To navigate to a particular target within the environment, the model chooses during the decision-making process the action that maximally increases the probability of reaching the respective spatial position.

Sensory processing

We chose place cells as a representation of the environment. In a previous study it has been shown how such place cell properties can be acquired by mobile robots using unsupervised learning in a hierarchical network (Wyss et al. 2006). Because here our main purpose is to model behavior we deliberately used predefined place cells to simplify this task. We approximated the firing properties of place cells as a function of the robots position by Gaussian functions (standard deviation: 0.04 m). To cover the whole four-arm-maze environment we randomly distributed 72 of these Gaussian functions. Hence, for each of the robots possible position we obtain the activity of each place cell. A winner-take-all process extracted the robots position in the state space from the activity of the place cells. Accordingly the robot was located in the state (place fields) corresponding to the most activated place cells. (The used distribution of these states is shown in Figure 1B.) In order to determine the robots current state we had to extract its position and calculate the place cells activity using the distributed Gaussian functions as described above. Hence the robot was tracked by a Color Cmos Camera 905C (Analog Camera), which was attached above the environment as shown in Figure 1A. The analog camera signal was digitized by a Hauppauge WinTV Express card. With the help of the camera and the color code attached on top of the robot, its position and orientation were calculated. Thus, the place cell represents a mapping from the position space where the robot is navigating, to the state space of the agent, controlling the robot. The agent uses only positional information provided by this state space (place fields).

Action execution

The robot was able to execute eight different actions in order to restrict the number of transition needed to learn the environmental properties. Each of these actions consisted of a certain orientation followed by a straight movement of the robot. The corresponding orientations were equally spaced from 0 to 325 degrees. As a result of executing such an action in a state (source), the robot will reach a different state (endstate) and thus results in a transition between states. The endstate is defined by the winner-take-all process calculating the current state, being dominated by another place cell. The position within a place field a transition is completed is defined by the place cell's activity not increasing anymore as the robot moves further and thus a local maximum of the activity is reached. The local maximum is defined by the derivative of the robots obtained activity of the place cells being zero. The frequencies of the transitions from source i with action k to endstates j is stored in the experience matrix $EM_{i,j,k.}$.

Reflexes

To prevent the robot hitting one of the mazes boundaries, a reflexive obstacle avoidance behavior was implemented. The proximity sensors (Fig 1C) were used in order to perform this behavior. If the robot had to use its obstacle avoidance behavior during action execution, the system associated the preceding state and action with the occurrence of a reflex event. The frequencies of co-occurrence of the reflexive event and a particular state (j) – action (i) combination is stored in the reflex matrix $RM_{j,i}$.

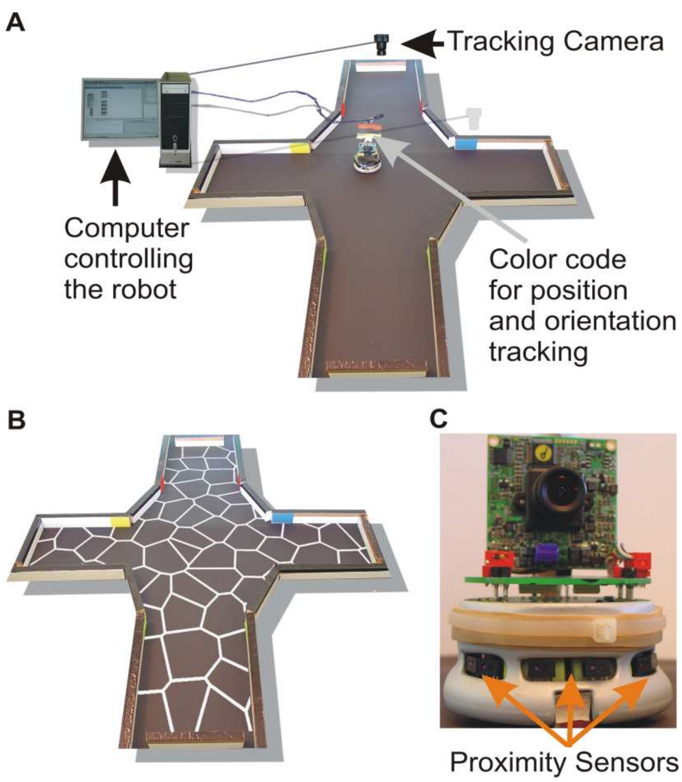


Figure 1: (A) We chose a four-arm-maze environment in order to test the model. A computer controlled the Khepera robot and extracted the robot's position and orientation, using the over head camera and the color code attached on top of the robot. (B) We subdivided the sensory space (position) into states. In the experiments we used this state distribution. The white boundaries assign for the region, one place cells is most activated. (C) Khepera robot used in the navigation experiments.

Decision making

The properties of the environment (boundaries, obstacles, etc.) determine how likely it is that a certain state transition will occur given a chosen action. These state transitions are approximated and learned by the agent as it explores its maze environment and are stored in a transition matrix (Figure 2A). The transition matrix consists of a 2D matrix for each action i TM_i. The row index determines the source j, the state where an action was executed and the row index represents the endstate k of this action. Thus the transition probability defined by source j, endstate k and action i is stored in the transition matrix $TM_{i,j,k}$ shown in Figure 2A. Hence summing of the transition matrix over the endstates k (rows) is normalized to one for each action and sources. For the experiments described below, the robot learned the transition probabilities based on 240 minutes of random exploration.

Next we address the problem of choosing the action with the most desirable outcome to move towards the goal. To accomplish this task an iterative reverse flooding approach was introduced, which integrates the environmental properties (Fig. 2B). The properties gained from the central component of the architecture are stored in the transition matrix. This matrix consisted of eight 2D matrices TM_i, one for each action, which share similarities with a directed graph. The vertices of this graph correspond to the states, the edges correspond to the transitions and the edges weight to the transition probabilities. This results in 8 directed graphs equivalent to the eight possible actions. In each of the iteration steps of reverse flooding the activation of the state corresponding to the goal states is one. The activation of activated states is propagated through the graph by passing the activity weighted with the corresponding transition probability to the states with transitions to the activated one (reverse direction of the directed edges). Technically spoken the activation is propagated from the endstates to the sources weighted by their transition probability, representing a backward flooding. This process gives rise to 8 different activity values for each state. Thus up to now only the learned environmental properties resulting from the central process were considered during the flooding process. To integrate also the learned properties caused by distal processing we introduced reflex factors. The reflex factor is proportional to the percentage of actions i combined with a reflexive event at source j:

$$rf_{j,i} = 1 - \left(\frac{RM_{j,i}}{\sum_k EM_{i,j,k}} \right) \bullet \frac{5}{6}$$

The weighting factor of 5/6 was introduced to prevent zero activation at an action which is combined only with obstacle avoidance behavior. Thus during each iteration step the eight activations of state j corresponding each to one of the eight actions i were multiplied by the corresponding reflex factor $rf_{j,i}$. After each of the iteration steps the maximum of the eight activations of a state were accounted as the states activation for the next iteration step. This iterative process was continued until the states activity converged. In order to select an action on a state to move towards the goal we considered the eight different incoming state activation values which resulted from the activation propagation of the eight actions. The robot chose the actions which resulted in the highest incoming activation of a state.

Furthermore we introduced a decay factor df which was here 0.9. After each iteration step, the states activation was multiplied by this factor. As more transitions are needed to reach the goal states as more the decay factor is taken into account and thus decreases the states activity. Hence, the decay factor penalized these trajectories to the goal state with more transitions to the goal.

Here the flooding algorithm defined in the last section was implemented with the help of matrices.

$$act_j(0)\begin{cases} 0 & j \neq l \\ 1 & j = l \end{cases}$$

represented the activation at the 0'th activation propagation, where the goal was located at state l.

$$\overrightarrow{act}(t+1) = \max_i((\mathrm{TM}_i \cdot \overrightarrow{\mathrm{act}}(\mathrm{t})) \cdot \mathrm{rf}) \cdot \mathrm{df} + \overrightarrow{\mathrm{act}}(0)$$

where $\overrightarrow{act}$(t) is the vector of activation values for the states after t iteration steps. df represents the decay factor.

Robot Setup

To test the model in a real-world environment we used Khepera II robots (K-Team). The robot was equipped with 8 proximity sensors, which emitted infrared light and measured the strength of its reflection, and two wheels, each controlled by one motor (Fig. 1C). For implementation and flexible programming, we used MicroPsi (Bach, 2003; Bach and Vuine, 2003), an Eclipse-based Java programming environment, as an interface to the robot. The agent that controlled the robot's behavior was implemented in this framework. The real-world environment was a four-arm maze with boundaries built from white wooden pieces (Fig. 1B). Each arm had a width of 0.21 m and a length of 0.28 m. The four-arm maze environment fitted into an area of one meter squared.

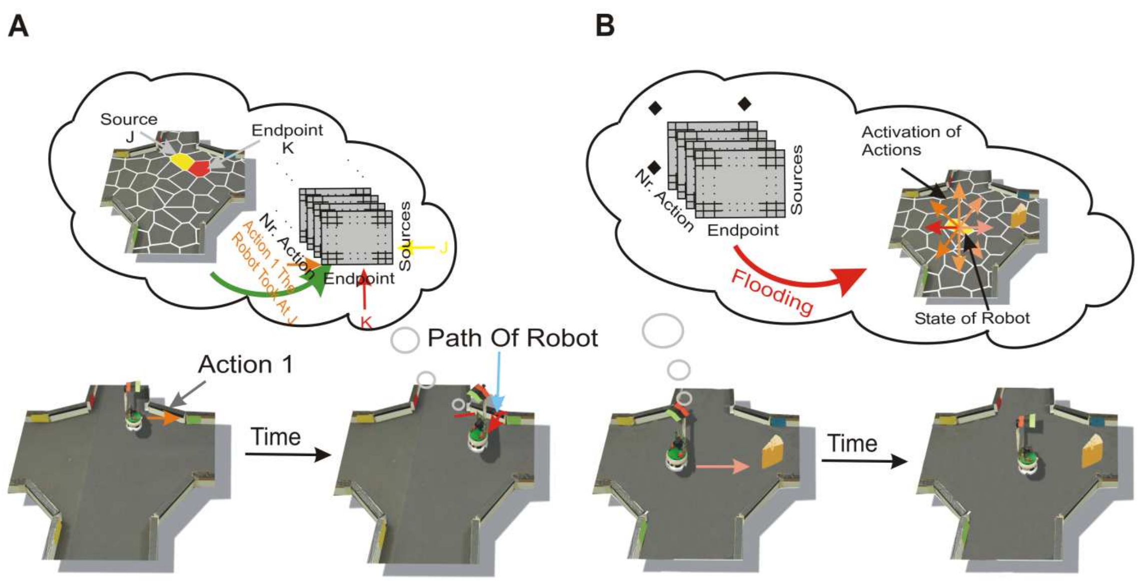


Figure 2**:** (A) Learning of the properties of the environment. The robot is on a certain state, defined here as Source J (yellow labeled) and randomly chooses an action (Action 1). The execution of the action results in another state, defined as endstate K (red labeled). This transition was stored in a three dimensional matrix, called the experience matrix, with the dimensions sources, endpoints and actions. The number of action executions combined with obstacle avoidance from a source was stored separately. (B) The robot moving to a goal (the "cheese" for the artifical "rodent"). His choice is a consequence of the flooding of the transition matrix, resulting in an activation of the different actions, shown as colored arrows. The action with the strongest activation was chosen.

Analysis

As a means of comparison, a simulated robot was implemented using MATLAB (Version 7.0 (R14), Mathworks). The same navigational and experience algorithm was used as described above. The obstacle avoidance behavior was implemented by setting the angle of reflection equal to the angle of incidence to the boundary, with a random scatter of 10 to –10 degrees added.

To compare the navigational behavior and the learned transition of the robot we introduce the geometrical transition matrix. It takes into account only the topographical properties of states in the environment. In order to experience the transition probabilities, based on the topographical properties, we let the simulated robot execute every action on each x/y position within the state given by the resolution of the tracker. Thus the geometrical transition matrix only takes the topographical distributions of states into account. Because the robot chose a new action according to the local maximum of the place cells activity, we weighted an actions transition by the probability of the robot choosing an action at the corresponding cell's activity. The execution of the different actions on each position within a state is due to transition probabilities resulting from an infinite experience time of the robot and thus represents the true underlying transition probabilities.

In order to compare the outcome of the different actions of one state learned by the robot, we measured the correlation coefficient of the transition probabilities of these actions on each state. We correlated the transition probabilities represented by a row vector of the Transition matrix of action i, TM_i, with the same row vector of the Transition matrix of action j TM_j. Before calculating the correlation coefficients between the two vectors we reduced the transition probabilities in the row vector by the average of these transition probabilities. This average was calculated by averaging over the transition probabilities of the topographical next neighbors. Thus two actions are equivalent when their correlation coefficient is 1.0; they are linearly uncorrelated when the correlation coefficient is 0.0.

To characterize the predictability of an actions' transition to a state we defined a second measure: The predictability of action i in state j is given by the maximum transition probability stored in the row vector j of the Transition Matrix TM_i. This maximum transition probability was reduced by the probability of transferring to one of the connected states by chance.

$$Pr_{i,j} = max_k\left(TM_{i,j,k}\right) - \frac{1}{conn_{i,j}}$$

$Pr_{i,j}$ corresponds to the predictability of action i in state j and $conn_{i,j}$ is the number of states the robot can transfer by executing action i on state j.

In order to evaluate the decision making process we analyzed the activation of each action calculated by the flooding process. We chose the normalized activity as an appropriate measure to characterize the selection of an action during navigating to a goal. This activity is defined as the most activated action on a state normalized by the sum of the incoming activity and the decay factor.

$$NormAct_j = \frac{\overrightarrow{act_j}}{\left(\sum_i((\mathrm{TM_i} \cdot \overrightarrow{\mathrm{act}}(\mathrm{t})) \cdot \mathrm{rf}) \cdot \mathrm{df} + \overrightarrow{\mathrm{act}}(0))_j\right) \cdot (1 - df)}$$

act_j represents the converged activity of state j after the flooding process. The denominator corresponds to the sum of the activation of a state j over all actions; the activity of state j as well as the sum of activities is given by the converged activity resulting from the flooding process. In order to reduce the dependency of the normalized activation onto the decay factor we multiplied the denominator by this decay factor. Thus normalized activity ranged from 0 to 10.

Results

Here we investigated the robots navigational performance and how the central processes, namely the transition probabilities, as well as the distal processes, defined by the reflex factors, contributed to the decision-making process.

Navigation performance

We investigated the navigation performance of the robot by analyzing its path to a number of different target sites in the environment. In each of the measured trials, the robot was placed on one of five possible starting positions and given one out of four target locations. In order to obtain a comparable measure we normalized the length of the robot's path by the *direct path*. The direct path represented the shortest traversable distance from the robot's starting point to the goal state. Figure 3 shows a path traveled by the robot (yellow line) and the corresponding direct path (light gray line). Overall the robot's median path length across 20 trials was 1.71 with a standard deviation of 0.47. This represents an increase of 71% (± 47%) when compared to the direct path. For all configurations of the start positions and targets, the robot was able to reach the target in a reasonably short amount of time.

This relative increase of the robots path length might have multiple causes: the division of the environment into discrete states (place fields), the robots learned transitions and the robots behavior while navigating through the environment. First we investigated the contribution of the discrete states in the lengthening of the robot's path to the targets. To provide a first approximation of this increase, we simulated the robot's behavior using the same navigational algorithm as described in the Methods section. The simulation used the geometrical transition matrix to navigate from the same start positions to the same goal states as the real robot. The transition probabilities of the geometrical transition matrix take only the topographical distribution of states into account (see Method section). Figure 3 shows a path of the simulated robot to a goal (red line). This simulation resulted in a median increase of 19% (±9%) compared to the direct path. Thus, the discrete states used here to represent the environment did not greatly contribute to the lengthening of the robot's path to a goal.

How can we interpret the robot's navigational behavior? Approximately a quarter of the increase of the robot's path to a goal was caused by the usage of discrete states as a representation of the environment. Another quarter of the lengthening can be explained by the differences between the robots learned and the geometrical properties, stored in the robot's experienced and the geometrical transition matrix (data not shown). Further we analyzed the effect of obstacle avoidance onto the robots navigational performance. The agent engaged its obstacle avoidance behavior in 60% of the trials independent of the particular combination of starting and goal states. Analyzing only the trials in which the agent did not engage obstacle avoidance we obtained a median of 1.36 (±0.23). Thus the largest share of the lengthening of the robot's path compared to the direct path is due the obstacle avoidance behavior. In all trials, the robot was able to find its goal in a reasonably short amount of time, with the main increase in path length arising from the necessity of navigating through the narrow arms of the maze, where obstacle contact occurs most frequently.

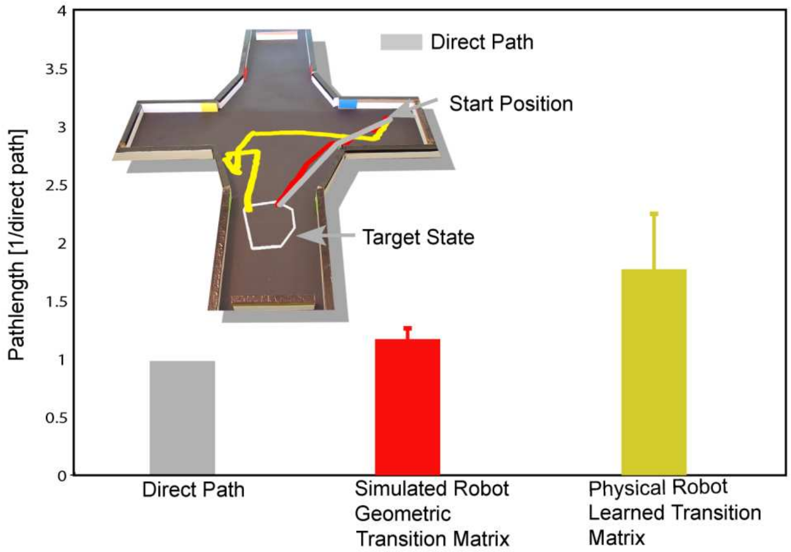


Figure 3**:** Navigational behavior of the robot was investigated by measuring the length of the path to different goals. The direct path, defined as the shortest traversable path from the start point to the goal state (shown as the gray line in the upper part), was used to normalize the length of the robot's path (yellow line) to the goal. The red line corresponds to the length of a path by a simulated robot by taking the topographical distribution of states (geometric transition matrix) into account. The bars represent the median length between different starting and goal states and their standard deviation.

Characteristics of the learned transition matrix

The robot's navigational performance results from the decision-making process. This process is based on the learned transition and the learned reflex factors, representing the learned environmental properties. Here we investigated the characteristics of the robots learned transitions. We investigated first, the differences between the transitions of different actions on a state; second the influence of the used topographical distribution of states on the learned transitions and third the predictability of the state to which one action execution transit to. Fourth we examined the effect of the robots limited time to learn the environmental properties on the learned transition probabilities. For the most part we analyzed the transition matrixes characteristics by comparison to the simulation, based on the geometrical transition matrix (see Method section), representing the transitions based on the used topographical distribution of states. By this comparison we investigate the extent to which the topographical distribution of states gives rise to the investigated characteristics of the transition matrix.

Here we analyzed the similarity between the transitions of different actions, defined as the redundancy of the robot's possible actions on a state, by comparing the transition probabilities associated with these actions. For this purpose we computed correlation coefficients (see Methods and Figure 4A,B) between the transition probabilities of the different actions on each state. Higher correlation coefficients (>0.5) were more frequently observed in the experienced transition matrix (44%) than in the geometrical case (25%), (Figure 4A). Thus the robot's real world action execution resulted in more similar outcomes and thus resulted in a higher redundancy of the actions compared to the geometrical case. Most (93%) of the highly correlated actions in the experienced case were obtained for states at the boundaries of the environment, and so were primarily due to the robot's obstacle avoidance behavior elicited by wall contact. The robot's action execution resulted in more similar transitions compared to the transitions based on the topographical distribution of states.

Next we investigated the influence of the topographical distribution of states on the robot's learned state transitions. Because the used topographical properties of states are fully represented by the geometrical transition matrix (see Method section), we calculated for each state and action the correlation coefficient between the transition probabilities stored in the geometrical and the robot's experienced matrix. Across all actions and states a mean correlation coefficient of 0.56 (±0.52) was obtained. Although these correlation coefficients are low it should be considered that these coefficients are calculated only for neighboring states and thus a conservative estimate. While different actions executed by the robot resulted in similar transitions more often than expected when only the topographical properties of the states are taken into account, the topographical state distribution nevertheless had an influence on the robot's learned transitions.

We then analyzed the predictability of action outcomes. Predictability defines the ability to predict the state to which one action execution makes a transition. In order to evaluate the actions' predictability we introduced predictability values (see Method section), proportional to maximum transition probability of an action. Figure 4D shows the occurrence of predictability values for the experienced and geometric transition matrices. Lower predictability values (<0.3) of the actions occurred more often in the experienced case (37%) compared to the geometric one (13%). Thus in general the robot's actions are equally likely to reach a number of spatially adjacent states. This is due to the actions transition probabilities characterized by a non-sparse probability distribution. Furthermore we investigated the influence of the obstacle avoidance behavior on the action predictability of the experienced transition matrix. Most (84%) of the low predictability values are due to actions for which the robot had to use its obstacle avoidance at least once. Thus obstacle avoidance reduced predictability of the action result. In most cases we obtained a lower predictability of the robot's resultant state, than we would have expected by the topographical distribution of place fields.

Are the differences between the robot's learned and the geometrical properties due to the robot's limited experience time? We generated the transition probabilities of the geometrical transition matrix by simulating the execution of each action on each position within a state (see Method section). Devolving this procedure to the robots learning of the transition probabilities, it has to experience its environment for an infinite time. In contrast, the robot's experienced transition matrix is based on executing each action on each state 11.54 times on average. Here we investigated the influence of the robots limited experience to the mean correlation coefficient between the geometrical and the robots experienced transition probabilities (0.56 ± 0.52). In order to investigate the influence of the robot's limited experience time on the difference between the geometrical and learned matrix, we compared generated geometrical transition matrices to the geometrical transition matrices. The generated geometrical transition matrices were calculated like the geometrical transition matrix; the only difference in the generated case is the number of actions executed on each state restricted to the one of the robots and thus was less than for the geometrical transition matrix. We simulated 300 generated transition matrices. In order to compare these matrices we correlated the transition probabilities for each action and state of the generated matrices with the geometric one. We averaged these correlation coefficients for each generated transition matrix. This yielded a distribution of averaged correlation coefficients with a mean value of 0.86 (± 0.1). Thus, the averaged correlation coefficient of 0.56 (± 0.52) between the geometric and the robots learned transition probabilities were lower than the correlation coefficients between generated and geometrical transition matrix. Thus the difference between the robot's experienced transition matrix and the geometrical transition matrix is dominated by the robots behavior and not due to limited time the robot experienced the environment.

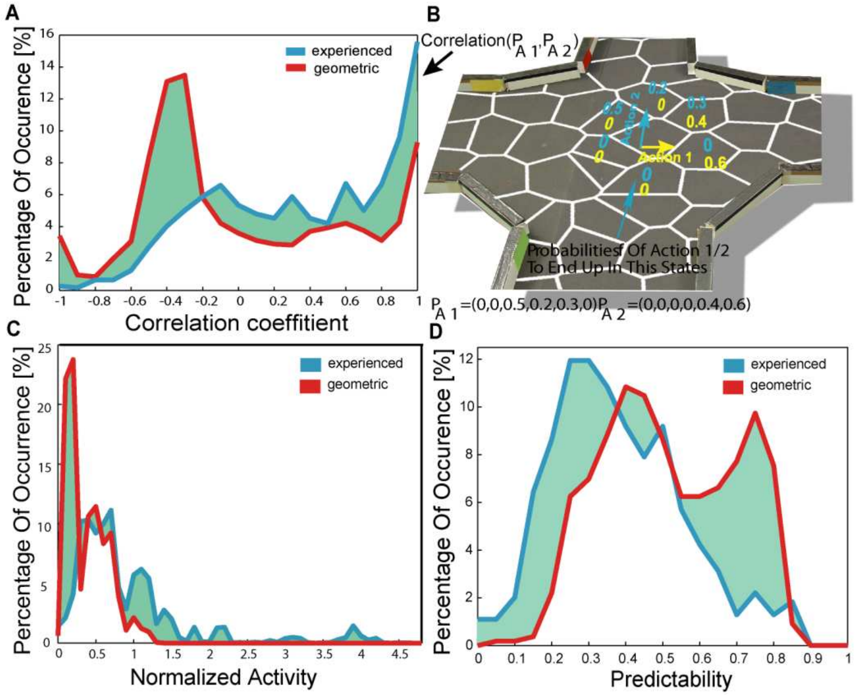


Figure 4: (A) The percentage of occurrence of the different correlation coefficients between the transition probabilities of different actions of a given state. The coefficients for the geometrical and the robot's experienced transition matrix are shown. (B) An example of calculating the correlation coefficients displayed in A. PA1/2 represents the probabilities of the action 1/2. (C) Here the ratio of the most activated action to the sum of all incoming activation (normalized activity (see Method section)) of a state is shown. The higher relfex factors in the experinced case (data not shown) increase the normalized activity compared to the geometric case. (D) Distribution of maximum transition probabilities of all experienced actions and states.

Here we have investigated the properties of the robots learned transition probabilities in the Transition matrix. We obtained a similarity between the transitions of different actions in the robot's experienced compared to the geometrical transition matrix. Also in general a lower predictability of the transition of the robot's actions to a state was obtained compared to the geometrical case. Thus both properties of the robots transition matrix are not fully caused by the topographical distribution of states. The obstacle avoidance behavior gives rise to this lower predictability as well as the similarity of the action results. Neglecting the reflex factors during the decision making process, which navigational behavior would result by only taking the learned transition into account? We would expect that it is not important for the robot to choose a precise action while moving towards a goal, caused by low action predictability as well as the high similarity between the transition probabilities of the different actions. However the transition matrix is influenced by the geometrical distribution of the place fields, while the obstacle avoidance behavior causes a similarity between the actions and a low predictability of an action's resultant state.

Reflexes

Next we investigate the impact of the distal processing on the agent's decision making process, which involves the selection of actions in order to move to a goal. The flooding processes integrates the distal components in the decision making process with help of the reflex factors. After flooding (see Method section), the agent selects the action most highly activated at the state corresponding to the robot's location. Here we investigate the impact of the reflex factors on this process by analyzing the normalized activation (see Method section). This measure is proportional to the ratio of the highest action's activation to the sum of the other action’s activation on a state. Thus a low normalized activation describes a decision making process with the execution of different actions would result in a similar navigational performance. In contrast, high values define a decision making process in which the agent chooses a precise action in order to move to the goal and thus executing different actions than the most activated one would result in different navigational performances. Taking only the transition matrix during the flooding process into account and thus neglecting the reflex factors, based on the properties of the transition probabilities investigated above, we would expect lower normalized activities compared to the geometrical transitions. In contrast, taking the reflexes into account, this normalized activation was higher for the experienced than for the geometric transition matrix (Figure 4C). This implies that the robot chose a precise action in order to move to a goal and thus executing a different action than the highest activated one result in a worse navigational performance. The higher normalized activations for the experienced transition matrix are due to higher reflex factors compared to the geometrical transition matrix (data not shown). Thus taking the reflexes into account reduces the effects of the obstacle avoidance behavior on the learned transitions during the decision-making process and results in a more precise action selection in order to successfully reach a goal.

How do the different components of the algorithm influence the behavior of the robot? Here we analyzed the contribution of the central processes and distal processes to the robots decision-making process. Taking only the central processes, namely the state transitions, for the decision-making into account, different actions executions would result in similar navigational performances, although navigation in the narrow arms required a precise action in order to reduces hits against the walls and thus reduce the path length to goals. In contrast, integrating the distal learned environmental properties, namely reflexes into the decision-making process the robot has to execute one precise action to navigate towards the goal. Thus as we expected, taking the distal processing into account reduces the effects of reflexive behavior and allows the robot to successfully navigate in the environment.

Discussion

Here we have introduced and implemented a model that allows a robot to navigate through an environment. The model learns the environmental properties, in an unsupervised manner by randomly executing the robot’s actions possibilities. Because the robots learning process was done in a finite time period, the robots knowledge of its actions possibilities only approximates its environmental affordances. The architecture of this model differentiated between central processing versus distal processing. The distal processing is defined by the state transitions where reflexive behavior of the sensory-driven obstacle avoidance occurred. The central processing is represented by all learned transitions between the states. The reflexive behavior acts upon the robots learned transitions, resulting in uniformly distributed and less predictable actions outcomes than we would have expected by looking at the used topographical distribution of place fields. However, as expected the integration of the information gained by the reflexive and central processing in the decision-making process reduced the impact of sensory-driven behavior on the navigational performance. Consequently the robot was able to successfully navigate in the environment in a short amount of time.

The cognitive model is based on a sensory representation composed of discrete states. In this state space First the robot learned the sensory outcomes of its actions execution, namely the state transition and the reflex factors. Thus, the robot learned the environmental properties with respect to its actions. Based on these results, the robot planned its action in order to move to the goal state in its internal state space. We defined the states such that they are equivalent to place cells place field, providing a representation of body position within the external space. These place cells can be understood as an optimally stable sensory representation of the visual input given by a robot moving in an environment (Wyss et al., 2006). The unsupervised learning resulted in a reorganization of the sensory space spanned by the robots visual input, leading to a low dimensional representation of the sensory input with a spatially meaning. In order to model other behavior, we have to choose an appropriate organization of the sensory space. On this sensory representation states can be defined, resulting in a state space. Further a definition of actions has to be done, which is adapted to the behavior to be modeled. Corresponding to these actions the sensory outcome in the state space can be learned. Differentiating between distal, namely the transitions influenced by the sensory driven behavior, and the central processes, the state transitions, would result in a better performance of the system to reach a certain goal state. Using a different sensory representation and other definition for the possible actions, different behaviors can be modeled.

Different studies have modeled navigational behavior by using place cells as a representation of the environment. Here we divide the different approaches into two group characterized by the type of learning used: Hebbian learning or reinforcement learning. The first type of learning exploits the fact that while moving in the environment, more than one place cell is active at the rodent's location, caused by the overlapping place fields of the corresponding cells. The Hebbian learning approach takes this fact and applies the biologically motivated principles of LTP and LTD, resulting in a strengthening of the connections between place cells which were active in a certain time interval. These cells and their connections between each other represent a cognitive map (Gerstner and Abott, 1996; Blum and Abott, 1996; Gaussier et al., 2002). Other studies introduced a cell type - goal cells - representing the goal of the navigational task (Burgess et al., 1997; Truellier and Meyer, 2000). The connections between the place and the goal cell encode the place cell's direction to the goal. The strength of connections between these two cell types was also modulated by Hebbian learning. In contrast to our model, the mentioned approaches rely on a global orientation and a metric, measuring the directions and distance to the goal at a given location within the environment. The global orientation used by these studies is defined using the same frame of reference over the whole environment. In contrast, we wanted the robot to learn the topology of the environment and thus did not introduce global variables as orientation or a metric. Furthermore, some of the mentioned studies (Stroesslin et al., 2005; Forster et al, 2000; Gerstner and Abbott, 1997; Burgess et al., 1997; Truellier and Meyer, 2000) used population coding to encode the position or direction to the goal. The population vector approach is based on the assumption of place fields and rodent's orientations having separate topologies. Thus to decode the robot's position or orientation the weighted average of place cells or orientations has to be calculated. This incorporates knowledge of the topology in the decoding scheme and impedes a generalization to other action repertoires. In contrast we defined the actions independently of each other so that the action repertoire can easily be expanded, for example including the action of lifting an object. Other branches of studies (Forster et al., 2000; Aleo and Gerstner, 2000; Stoesslin et al., 2005) used reinforcement learning (Sutton and Barto, 1997) to perform a navigational task. The concepts of Markov Decision Process and value iteration (Sutton and Barto, 1997) are commonalities between reinforcement learning and our approach, while in our model the value iteration was expanded by reflexes. A pure reinforcement learning approach involves learning the properties of the environment by using an explicit reinforcement signal, given by a goal state; while in the presented model these properties are latently learned (Tolman, 1948), resulting in a global strategy for navigation in this environment. In contrast to other studies, here we presented a cognitive model that is able to learn the topology and properties of the environment in a latent manner and can be expanded to model other behaviors by redefining the meaning of the actions and states.

We introduced a cognitive architecture in order to model animal-like behavior and tested it in a navigational framework. The navigational performance given by this architecture is not constrained to a specific setup because the behaviorally interpreted properties of the environment are self-learned and not predefined. Here we showed that differentiation between central and distal processing routines resulted in a better navigational performance. We argued that this cognitive model can be expanded to model other behavior.

References

Arleo A, Gerstner W. (2000) Spatial cognition and neuro-mimetic navigation: a model of hippocampal place cell activity. *Biol Cybern,* 83(3):287-99.

Bach, J. (2003) The MicroPsi Agent Architecture *Proceedings of ICCM-5, International Conference on Cognitive Modeling, Bamberg, Germany* (pp. 15-20)

Bach, J., Vuine, R. (2003) Designing Agents with MicroPsi Node Nets. Proceedings of KI 2003, Annual German Conference on AI. LNAI 2821, Springer, Berlin, Heidelberg. (pp. 164-178)

Blum KI, Abbott LF. (1996) A model of spatial map formation in the hippocampus of the rat. *Neural Comput*. 8(1):85-93.

Burgess N, Donnett JG, Jeffery KJ, O'Keefe J. (1997) Robotic and neuronal simulation of the hippocampus and rat navigation. *Philos Trans R Soc Lond B Biol Sci*. 29;352(1360):1535-43.

Foster DJ, Morris RG, Dayan P. (2000) A model of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus. 10(1):1-16*

Gaussier P, Revel A, Banquet JP, Babeau V. (2002) From view cells and place cells to cognitive map learning: processing stages of the hippocampal system. *Biol Cybern*. 86(1):15-28.

Gerstner W, Abbott LF. (1997) Learning navigational maps through potentiation and modulation of hippocampal place cells. *J Comput Neurosci.* 4(1):79-94.

Gibson James J. (1977) The Theory of Affordances. In Perceiving, Acting, and Knowing, *Eds. Robert Shaw and John Bransford*

Morris R (1984) Developments of a water-maze procedure for studying spatial learning in the rat. *J Neurosci Methods* 11 (1): 47-60.

Olten, D.S., & Samuelson, R.J. (1976) Remembrance of places passed: Spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes,* 2, 97-116.

O'Keefe J, Dostrovsky J. (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* 34(1):171-5.

O'Keefe J. Nadel L. (1978) The Hippocampus as a Cognitive Map. *Oxford University Press*.

Sutton, Richard S.; Andrew G. Barto (1998) Reinforcement Learning: An Introduction. *MIT Press*.

Strosslin T, Sheynikhovich D, Chavarriaga R, Gerstner W. (2005) Robust self-localisation and navigation based on hippocampal place cells. *Neural Netw*. 18(9):1125-40. Epub 2005 Nov 2.

Tolman EC. (1948) Cognitive Maps in Rats and Man. *Psychological Review* 55: 189-208.

Trullier O, Meyer JA. (2000) Animat navigation using a cognitive graph. *Biol Cybern.* 83:271-85.

Wilson R.A., Keil F.C. (1999) The MIT encyclopedia of the cognitive sciences. *MIT Press*

Wyss R, König P, Verschure PF. (2006) A model of the ventral visual system based on temporal stability and local memory. *PLoS Biol.* ;4(5):e120.

Pseudoneglect as a Function of Hand Use and Line Length

EVA BUECHEL, ANNAMARIE VOGT, KLAUS OPWIS

Department of Cognitive Psychology and Methodology
University of Basel, Switzerland[1]

When bisecting horizontal lines at the midpoint as accurately as possible, neurologically normal individuals reliably show a leftward error, called Pseudoneglect. The aim of the study was to systematically examine the influence of line length for lines bisected with the right and the left hand. Healthy participants were asked to bisect lines of six different lengths. Our data suggests that there is a significant leftward error for all line lengths. For absolute deviations, this error tends to be greater for lines bisected with the left hand than for lines bisected with the right hand. Furthermore, although the absolute deviations clearly increase with longer lines, no main effect was found for relative deviations. A reliability analysis suggests that Pseudoneglect is a relatively constant phenomenon within each participant over all line lengths.

Introduction

When asked to bisect horizontal lines at the midpoint as accurately as possible, neurologically healthy individuals show a systematic leftward error. This phenomenon is referred to as pseudoneglect (Bowers & Heilman, 1980). The term pseudoneglect stems from the name of the neurological condition “neglect”. Neglect patients fail to attend to the contralesional side of space, which mostly occurs after right hemispheric lesion (Buxbaum et al., 2004). Ignoring their left visual field, neglect patients bisect lines to the right of their actual middle in line bisection tasks.

Pseudoneglect, in contrast, is a phenomenon that occurs in healthy individuals whose error to the left is not as prominent as the right error of neglect patients (Jewell & McCourt, 2000). One explanation for this phenomenon that has received much support is the activation-orientation hypothesis (Reuter-Lorenz, Kinsbourne & Moscovic, 1990). This hypothesis states that visuospatial attention is distributed in the opposite direction to the more activated hemisphere. Since line bisection is a visuospatial task involving activation of the right hemisphere, the left half of the horizontal line in a line bisection task receives

[1] Anschrift: Annamarie Vogt, Institut für Psychologie, Missionsstrasse 62a, 4055 Basel, Switzerland, email: annamarie.vogt@unibas.ch

more attention and is thus perceived longer than the less attended right half (Bultitude & Davies, 2006). However, Porac, Searleman and Karagiannakis (2006) have demonstrated that also perceptual factors contribute to the phenomenon of pseudoneglect. They showed in their study that performance on line bisection tasks could be altered toward the right or the left of the actual middle using dots placed on or near the line. The dots were placed according to the theory of visual illusions of length, called *centroid extraction* (Morgan, Hole, & Glennerster, 1990), which argues that the position of a visual target is calculated as the mean position of all stimuli in close proximity to the target stimulus.

Extant studies found that the extent and direction of line bisection errors vary as a function of various individual factors such as sex, age and handedness and can be modulated by methodological manipulations such as scanning direction, used hand, positioning of line, and line length (Porac et al. 2006). For example, Fujii, Fukatsu, Kimura, Saso and Kogure (1991) found a slight non-significant tendency for participants to bisect lines more towards the right of the actual middle with increasing age. These findings were supported by Stam and Bakker (1990), who found a rightward error for older participants.

Hausmann, Ergun, Yazgan and Güntürkün (2002) found an interaction for sex and hand used to bisect the lines. In their study, females showed a similar leftward error with both hands used to bisect the lines, while males showed the error predominantly when using the left hand. Earlier studies on sex differences summarized in the meta-analysis by Jewell and McCourt (2000) found no gender differences in line bisection. Two studies (Roig & Cincero, 2004; Wolfe, 1923) found gender effects that do not, however, support each other's findings.

The differences in hand use, as mentioned above, are of interest to understand pseudoneglect, as the use of one hand will activate the motor cortex, resulting in a more activated left or right hemisphere. Left hand use, which is associated with the right motor cortex, would activate the right hemisphere and lead to a stronger pseudoneglect. The majority of studies conducted on pseudoneglect as a function of hand use have found pseudoneglect patterns that support this reasoning. When the left hand is used, participants tend to err farther to the left than when the right hand is used (Jewell & McCourt, 2000; Luh, 1995). Considering this observation, it seems to be appropriate to investigate pseudoneglect with a non-motor task. In the traditional line bisection task commonly used to test neglect patients (Porac et al. 2006), subjects are asked to manually bisect a horizontal line in the actual middle. To rule out that pseudoneglect is a phenomenon occurring due to motor components, different tasks, such as the landmark task, have been developed. On the landmark task, participants are asked to make a forced choice judgment about the location of the mark, stating if the mark is closer to the right or the left end of the line (Bultitude & Davies, 2006). Even if motor components are ruled out or at least minimized to a key-press response, healthy participants and neglect patients still show a biased performance consistent with their performance on manual line bisection tasks.

Another critical factor in line bisection tasks is the length of the lines. This is assumed to be especially important due to fact that line length has an influence on line bisection performance for neglect patients (Jewell & McCourt, 2000). Neglect patients seem to show a crossover effect, bisecting long lines to the right of the objective center and short lines to the left of the objective center (Halligan & Marshall, 1988). Some studies also found a crossover effect for normal individuals, whereby normal participants bisect long lines (> 2cm) too far to the left and short lines (< 2cm) too far to the right (Mennemeier et al., 2005; Rueckert, Deravanesian, Baboorian, Lacalamita, & Repplinger, 2002). Rueckert et al. (2002) found this crossover effect only on purely perceptual tasks (e.g. Landmark task), but not on traditional manual line bisection tasks.

Other studies on the influence of the length of lines on pseudoneglect have yielded inconsistent results. Chokron and Imbert (1993) found that subjects reading from the right bisected the lines to the left of the actual middle, whereas subjects reading from the left bisected the lines to the right from the actual middle. For both groups, the absolute error in millimeter was reported to increase as a function of line length. Luh (1995) found that the error of line bisection in millimeters also increased for longer lines, with a greater effect if the lines were displaced to the left. Manning, Halligan and Marshall (1990) found that the within-subject variability of the deviations in millimeters increases with line length and that the mean displacement from the middle in millimeters (be it left or right, depending on subjects) is linearly related to line length. In a case study comparing two neglect patients with normal subjects, Halligan, Manning and Marshall (1991) also found a trend of pseudoneglect in healthy subjects increasing with longer lines when considering the absolute deviations in millimeters.

While the above three studies found pseudoneglect to increase with length of line, three other studies found no effect of line length on pseudoneglect. Using absolute deviations in millimeters, Mattingley, Pierson, Bradshaw, Phillips and Bradshaw (1993) found a main effect for line length with neglect patients. However, they found no main effect for the normal participants. Interestingly, Butter, Mark and Heilman (1981) also found a main effect for line length with a neglect patient and no main effect for their five healthy controls, using relative deviations from the middle of the line in percent. In Halligan and Marshall's data on healthy individuals, pattern of the deviations in millimeters again does not suggest that pseudoneglect increases with line length (Halligan &Marshall, 1988). Furthermore, while one study found that subjects bisected on the left of the middle for short lines and right of the middle for long lines, with no error for middle lines, a study by McCourt and Jewell (1999) found rightward errors with short lines and leftward errors with medium and long lines.

The results of the extant studies are very inconsistent. One reason for this inconsistency could be that the line length categories used were sometimes crude (e.g. long vs. short lines) (McCourt & Jewell, 2000) and others were more continuous. The role of line length in pseudoneglect is therefore not clear from the literature thus far. Although line length was one independent variable in many studies conducted on pseudoneglect, it appears that it was never considered to be critical by the researchers and thus not thoroughly investigated.

Due to these inconsistencies in the findings, the aim of this study was to systematically examine the role of line length on pseudoneglect using multiple line lengths. We were especially interested in the relative deviations from the actual middle, which measure the deviations from the middle in percent. Surprisingly, most extant studies on pseudoneglect as a function of line length did not differentiate between the absolute deviation and the relative deviation from the actual middle, reporting absolute values only. For example Jewell and McCourt (2000) did not distinguish between studies using relative or absolute deviations in the review on factors influencing pseudoneglect. However, considering that the deviation from the middle cannot be more than half of the line length and therefore cannot exceed a certain magnitude, we could suppose that the increased error with longer lines is due to greater error potential, whereby the subjects' chance to err to one side or the other increases with the increase of line length. This is why it is essential to differentiate between the two deviations, as they measure distinctive bisection errors. The relative deviations enable us to compare deviations across line lengths, disregarding the length of the line. Hence, it is also meaningful to investigate whether pseudoneglect increases with line length when relative deviations, which measure the deviations from the middle in percent, are used. We hypothesized that pseudoneglect increases with longer lines for the absolute deviations, not, however, for the relative deviations. Consistent with extant literature, we also expected participants to err more to the left when the lines were bisected with the left hand.

Methods

Participants

A total of 54 undergraduate students (36 female and 18 male) from the University of Basel, Switzerland volunteered to participate in the experiment. They were given a choice of chocolate bars upon completion of the task. The mean age of the participants was 25.35 (*SD*=7.25; range: 19-49 years). The participants' handedness was determined using the Edinburgh Handedness Inventory (Oldfield, 1971). The majority of the participants were right-handed, one participant was left- handed, and one participant was ambidextrous.

Materials

The stimuli for the line bisection task consisted of a total of 24 black lines centered on a DIN A4 paper in landscape format. The length of the lines was varied in six different lengths (4cm, 8cm, 10cm, 12cm, 16cm, 20cm). Each length was presented four times, twice for the left hand and twice for the right hand used to bisect the lines. The line lengths were approximately one millimeter in width. Every sheet bearing a line to bisect was followed by a blank sheet to ensure that the following line could not be seen and therefore could not influence the line bisection task. The bundle of 12 lines for each hand used was stapled on the upper corner of the sheets, on the left side for the lines that were to be bisected with the right hand and on the right side for the lines that were to be bisected with the left hand. This enabled the participants to turn the pages without hindrance.

The sequence of the line length was presented in five randomly assembled orders and participants were randomly assigned to one of the five sequences. The order of the lines was identical for the right hand and for the left hand. There was an instruction sheet on the first page of the first bundle and at the beginning of the second bundle, which was viewed when changing the hand used to bisect the lines.

At the end of the two bundles, there was a questionnaire where information on age, gender, brain or head injuries, previous knowledge of the purpose of the study and handedness (Oldfield, 1971) was collected.

Procedure

The data was collected after an introductory class. The participants were seated in an auditorium with a fold down table that was large enough to allow for a DIN 4 landscape-formatted paper to be centered in front of the participants.

The participants were instructed to bisect the presented lines in the middle as accurately as possible, without measuring and deliberating. They were provided with sharpened pencils to ensure an exact measurement. They were instructed to use their right hand for the first set of lines and the left hand for the second set of lines, regardless of their handedness. After completing the task, the sets of lines were collected by the experimenter, and the participants were given a choice of chocolate bars.

Design

The design of the study was a 2 x 6 design, with the repeated measure factors hand use and length of lines. We carefully measured the deviations from the objective center of the line to a 0.5 millimeter accuracy. Our dependent variables were the absolute and relative deviations from the actual middle. The absolute deviation was calculated by taking the mean of the same two length lines that were bisected with the same hand. In addition, we computed the relative deviations as follows: ((measured right half – true half)/ true half) x 100). Again, the mean relative deviation of the two same-length lines bisected with the same hand was calculated. These relative deviations allow comparison of the deviations over all line lengths, disregarding line length.

Results

Two of the 54 participants were not right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971) and were thus excluded from our analysis. It is assumed that left-handed participants have a different hemispheric dominance and could thus show different patterns on line bisection tasks. As none of our subjects reported having any history of brain injuries, no further subjects were excluded.

The dependent measures were the absolute deviations (mean of equal length bisected with same hand) and the relative deviations (mean percentage of deviation of equal line length bisected with the same hand) from the objective middle.

We hypothesized that participants would bisect the lines left of the actual middle, regardless of the hand used. To test this hypothesis, we conducted one-sample *t*- tests over the mean deviations per line length bisected with the right/left hand respectively to test if the deviations deviated significantly from zero. All mean deviation values (*M*s > .47, *SD*s >.91 for absolute values, *M*s > 1.92, *SD*s > 3.08 for relative values) deviated significantly to the left of the actual middle (*t*s > 3.42, *p*s ≥ .001). For an overview of the results, see Table 1.

Furthermore, we hypothesized that the leftward error would increase with line length for the absolute values, not, however, for the relative values. For the absolute values, a two factorial ANOVA with the repeated measure factors used hand (left and right) and line length (4cm, 8cm, 10cm, 12cm, 16cm, 20cm) revealed a main effect for both used hand, F (1, 51) = 3.96, p = .026 and line length, F (5, 255) = 14.42, p < .001. The main effects were moderated by a significant interaction of the two repeated measure factors, F (5, 255) = 3.63, p = .002. The pattern of the data is displayed in Figure 1.

To further investigate this interaction, we conducted a linear contrast analysis for the two separate hands used as a function of line lengths. The linear contrast analysis revealed significant linear positive trends for both hands, $F(1, 51)_{\text{right hand}} = 9.69$, p = .003, $F(1, 51)_{\text{left hand}} = 38.71$, p = < .001.

As expected, a two factorial ANOVA for the relative deviations, with the repeated measure factors used hand and line length, did not yield a main effect of the line length, F (5, 255) = .45, p = .811. For the relative deviations, there was also no main effect for hand use, F (1, 51), p = .186. For an illustration of the findings, see Figure 2.

Table 1

Mean deviations towards the left of the actual middle for the absolute and the relative deviations and their standard deviation. The p-values for the means are shown, which are the same for the absolute and the relative deviations.

	M (SD) in mm	*M (SD)* in %	*t*	df	*sig.* (one-tailed)
Right hand					
4 cm	.74 (.91)	3.61 (4.63)	5.87	51	<.001**
8 cm	.93 (1.85)	2.32 (4.63)	3.61	51	.001*
10 cm	1.04 (1.89)	2.09 (3.79)	3.97	51	<.001**
12 cm	1.45 (2.29)	2.42 (3.82)	4.57	51	<.001**
16 cm	1.54 (2.85)	1.93 (3.56)	3.90	51	<.001**
20 cm	1.92 (3.08)	1.92 (3.08)	4.49	51	<.001**
Left hand					
4 cm	.47 (.98)	2.33 (4.91)	3.42	51	.001*
8 cm	1.40 (1.80)	3.50 (4.50)	5.60	51	<.001**
10 cm	1.33 (2.12)	2.66 (4.24)	4.53	51	<.001**
12 cm	1.84 (3.15)	3.06 (5.25)	4.20	51	<.001**
16 cm	2.59 (3.33)	3.24 (4.16)	5.61	51	<.001**
20 cm	3.36 (3.44)	3.36 (3.44)	7.04	51	<.001**

* $p < .01$, ** $p < .001$.

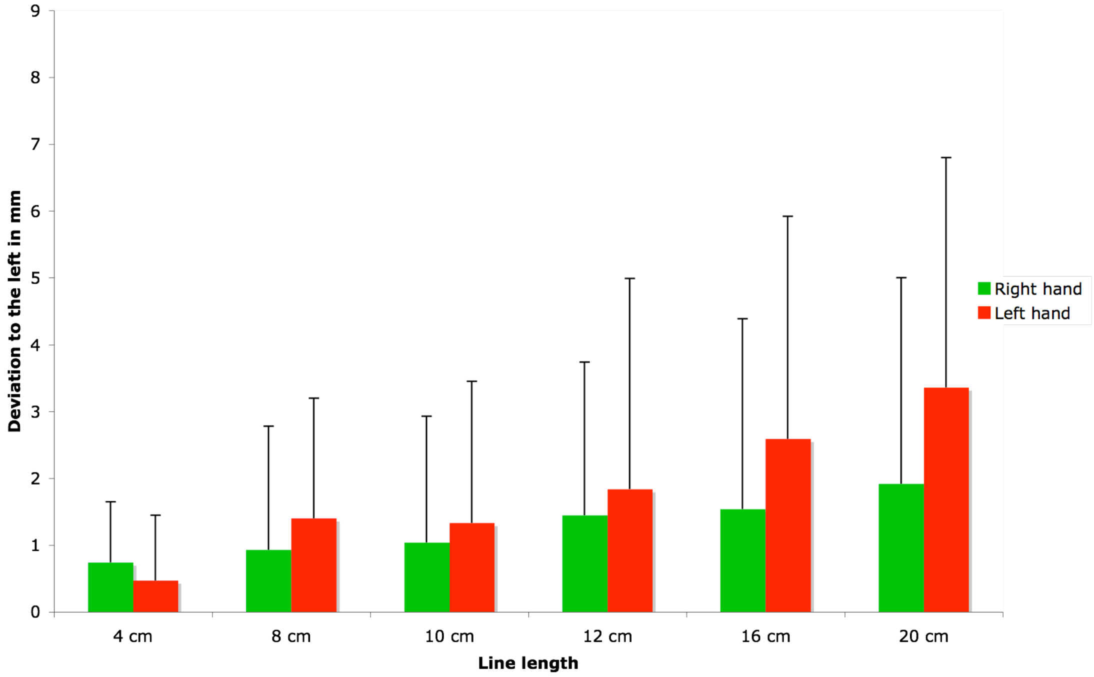


Fig. 1 Mean and standard deviations of the absolute deviations to the left in millimeters as a function of line length and hand used to bisect the lines.

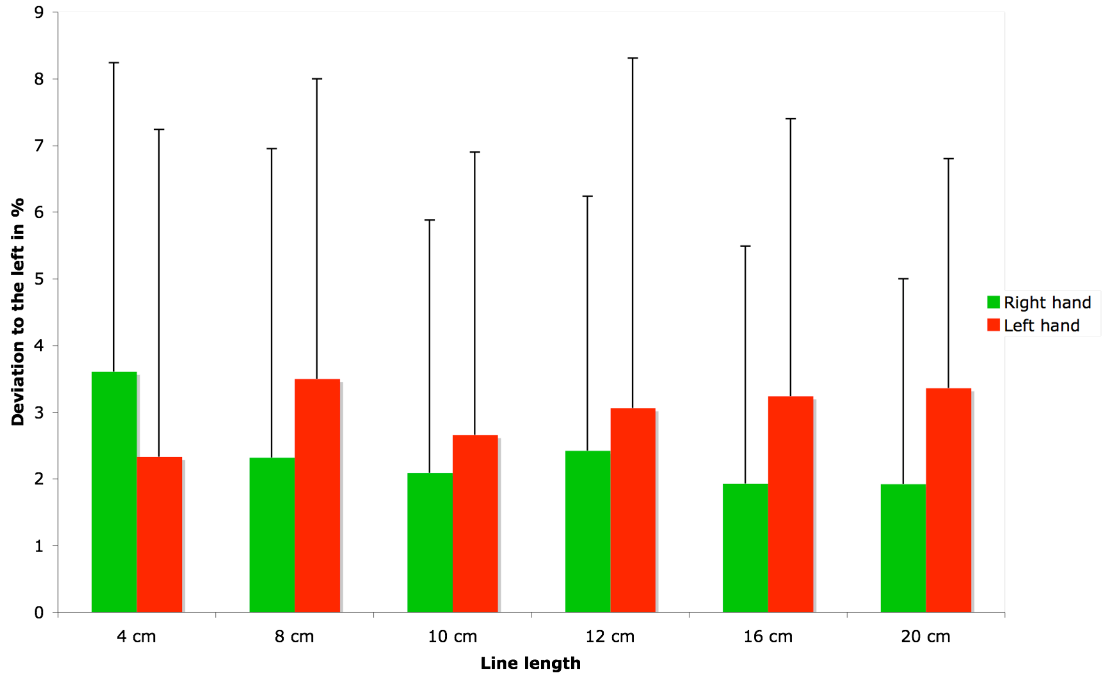


Fig. 2 Mean and standard deviation for the relative deviations to the left in % as a function of line length and hand used to bisect the lines.

Unexpectedly, however, we found an interaction of the line length and hand use, $F\ (5, 255) = 3.60$, $p = .004$. To gain further insight into the role of line length on pseudoneglect, we again conducted a linear contrast analysis for separate hands used. Linear contrasts revealed that there is a significant negative linear relation to line length for the right hand, $F\ (1,51) = 5.37$, $p = .025$. For the left hand, however, linear contrasts did not reach significance, $F\ (1,51) = 1.28$, $p = .264$.

We additionally did a reliability analysis to find out if individuals' performance on line bisection tasks remains constant over the different line lengths. The relative deviations were used to investigate reliability. With respect to the relative deviations, we found a Cronbach's $\alpha = .84$ for the right hand and a Cronbach's $\alpha = .89$ for the left hand.

Discussion

The aim of this study was to replicate the leftward error of pseudoneglect and examine the influence of line length on pseudoneglect. As predicted, we found a significant leftward error in line bisection tasks for all line lengths bisected with both the right and the left hand. Furthermore, we systematically examined the influence of line length on pseudoneglect in line bisection performances. We explored this by looking particularly at the relative deviations from the middle.

Our main prediction was that the extent of pseudoneglect increases when the length of the lines increases for the absolute deviation values, not, however, for the relative deviation values. This is exactly what we found. There was a significant main effect for line length for the absolute deviations, and linear contrast analysis revealed that there is a linear trend for the leftward errors to increase with longer lines, regardless of hand use. For the relative deviations, there was no main effect. Our results are supported by a study of Butter et al. (1993). In an experiment with one neglect patient and five controls, they found that the relative deviations from the middle increased with line length for neglect patients, not, however, for the controls.

Interestingly, we found a positive linear trend for the right hand for the relative deviations not, however, for the left hand. One possible explanation for this is that the right hand covers up a great part of the line while bisecting, making the right half appear smaller. In contrast, when bisecting with the left hand, the full line is visible.

Furthermore, we assumed that the leftward error would be stronger when the left and non-dominant hand was used to bisect the lines. We found this main effect for hand use for the absolute deviation values. We did not, however, find a main effect for the hand use for the relative deviations.

We additionally did a reliability analysis to find out if the extent of pseudoneglect varies within subjects. Our relatively high Cronbach's alpha coefficients suggest that the intra-individual variation of pseudoneglect is rather small. This means that the line bisection task reliably measures the extent of each individual's deviations, resulting in pseudoneglect, and that each individual's extent of pseudoneglect remains relatively constant across all line lengths.

Our results therefore provide clarification of the role of the length in line bisection tasks with healthy individuals. Our data also highlights the importance of differentiating between relative and absolute values.

As mentioned above, for the absolute values, the participants erred more to the left when they bisected the lines with their non-dominant left hand. Because we only had one ambidextrous and one left-handed participant, it remains unclear if this effect is due to the dominance of the hand or to the hand used. Furthermore, all participants bisected all the lines with the right hand first. The greater error manifested in the absolute deviations to the left could therefore also be caused by a tiring effect. As our stimuli did not contain lines shorter than 4cm, we were unable to replicate the crossover effect found by Mennemeier et al. (2005).

For further studies, it would therefore be advantageous to add a number of shorter lines and to randomize the order of the hand used to bisect the lines. Further investigations are also required to explain the discrepancies in linear contrast for the right and the left hand.

References

Bowers, D., & Heilmann, K. M. (1980). Pseudoneglect: Effects of hemispace on a tactile line bisection task. *Neuropsychologia*, *18*, 491-498.

Bultitude, J. H., & Aimola Davies, A. M. (2006). Putting attention on the line: Investigating the activation-orientation hypothesis of pseudoneglect. *Neuropsychologia, 44*, 1849-1858.

Butter, C. M., Mark, V. W., & Heilman, K. M. (1988). An experimental analysis of factors underlying neglect in line bisection. *Journal of Neurology, Neurosurgery, and Psychiatry, 51*, 1581-1583.

Buxbaum, L. J., Veramonti, T., Frassinetti, F. Ladavas, E., Coslett, H. B., Ferraro, M. K., et al. (2004). Hemispatial neglect: Subtypes, neuroanatomy, and disability. *Neurology, 62*, 749-756.

Chokron, S., & Imbert, M. (1993). Influence of reading habits on line bisection. *Cognitive Brain Research, 1*, 219-222.

Fujii, T., Fukatsu, R., Kimura, I., Saso, S., Kogure, K. (1991). Unilateral spatial neglect in visual and tactile modalities. *Cortex, 27,* 339-343.

Halligan, P. W., Manning, L., & Marshall, J. C. (1991). Hemispheric activation vs. Spatio-motor cueing in visual neglect: A case study. *Neuropsychologia, 29,* 165-176.

Halligan, P. W., & Marshall, J. C. (1988). How long is a piece of string? A study of line bisection in a case of visual neglect. *Cortex, 24*, 321-328.

Hausmann, M., Ergun, G., Yazgan, Y., & Güntürkün, O. (2002). Sex differences in line bisection as a function of hand. *Neuropsychologia, 40*, 235-240.

Jewell, G., & McCourt, M. E. (2000). Pseudoneglect: A review and meta-analysis of performance factors in line bisection tasks. *Neuropsychologia, 38*, 93-110.

Luh, K. E. (1995). Line bisection and perceptual asymmetries in normal individuals: What you see is not what you get. *Neuropsychology, 9*, 435-448.

Manning, L., Hallingan, P. W., & Marshall, J. C. (1990). Individual variation in line bisection: A study of normal subjects with application to the interpretation of visual neglect. *Neuropsychologia, 20*, 647-655.

Mattingley, J. B., Pierson, J. M., Bradshaw, J. L., Phillips, J. G., & Bradshaw, J. A. (1993). To see or not to see: The effects of visible and invisible cues on line bisection judgments in unilateral neglect. *Neuropsychologia, 31,* 1201-1215.

McCourt, M. E., & Jewell, G. (1999). Visuospatial attention in line bisection: Stimulus modulation of pseudoneglect. *Neuropsychologia, 37,* 843-855.

Mennemeier, M., Pierce, C. P., Chatterjee, A., Anderson, B., Jewell, G., Dowler, R., et al. (2005). Biases in attentional orientation and magnitude estimation explain crossover: Neglect is a disorder of both. *Journal of Cognitive Neuroscience, 17*, 1194-1211.

Morgan, M. J., Hole, G. J., & Glennerster, A. (1990). Biases and sensitivities in geometrical illusions. *Vision Research, 30,* 1793-1810.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia, 9*, 97-113.

Porac, C., Searleman, A., & Karagiannakis, K. (2006). Pseudoneglect: Evidence for both perceptual and attentional factors. *Brain and Cognition, 61,* 305-311.

Reuter-Lorenz, P.A., Kinsbourne, M. & Moscovitch, M. (1990). Hemispheric control of spatial attention. *Brain and Cognition, 12*, 240-266.

Roig, M., & Cincero, F. (1994). Hemisphericity style, sex and performance on a line-bisection task: An exploratory study. *Perceptual and Motor Skills, 78*, 115-120.

Rueckert, L., Deravanesian, A., Baboorian, D., Lacalamita, A., & Repplinger, M. (2002). Pseudoneglect and the crossover effect. *Neuropsychologia, 40*, 62-173.

Stam, C. J., & Bakker, M. (1990). The prevalence of neglect: Superiority of neuropsychological over clinical methods of estimation. *Clinical Neurology and Neurosurgery, 92*,229-235.

Wolfe, H. K. (1923). On the estimation of the middle of lines. *The American Journal of Psychology, 34*, 313-358.

Contributions of modality-specific brain regions to working memory and long-term memory: An ERP study.

THORSTEN A. BRINKMANN & HUBERT D. ZIMMER

Brain and Cognition Unit, Department of Psychology, Saarland University, Saarbruecken, Germany

In the present ERP study, we investigated whether modality-specific brain regions were activated in a working memory (WM) task and re-activated in a subsequent long-term memory (LTM) task. Furthermore, slow-wave potentials evoked during both tasks should have similar topographies, indicating partially overlapping neuronal networks. Therefore, subjects participated in an S1 – S2 matching task (WM) followed by a pair association task (LTM). Using a cueing paradigm we were able to separate the processes taken place during both tasks. Event-related potentials (ERPs) were recorded for working memory encoding, maintenance and target presentation as well as for long-term memory retrieval and long-term memory target representation. We looked for different distributions of slow-wave potentials between the modalities during the maintenance and the retrieval of modality-specific information. In addition, the topographies for each modality were expected to be similar for the WM and the LTM task. For the working memory maintenance we found no modality-specific modulations of slow-wave potentials. Instead, slow-wave potentials occurred already during the working memory cue presentation, indicating general working memory processes. ERPs for the WM target presentation provided some evidences for modality-specific information processing. While for the LTM retrieval slow-wave potentials did not vary as a function of modality, modality-specific effects were obtained during the LTM target presentation. Furthermore, as revealed by topographic analyses, the distributions of the observed effects were similar for LTM and WM suggesting that both mental representations are supported by similar neuronal networks, including early sensory brain regions.

Introduction

Psychological memory research is traditionally divided into long-term and short-term memory approaches. This division of labor is caused by differences in the used research paradigms (matching tasks and serial recall tasks with delays of a few seconds versus item and source recognition, as well as free recall tasks after longer delays) and by dissociations observed in patients. Patients may show nearly perfect short-term memory while they are clearly impaired in long-term memory and vice versa (Shallice & Warrington, 1970; Warrington & Weiskrantz, 1968). Recently, however, this distinction has been increasingly challenged (Ranganath & Blumenfeld, 2005). Two lines of results are put forward in support of this revised position. First, in several experiments it was shown that patients with impaired long-term and seemingly intact short-term memory were also impaired in short-term memory tasks when previously unknown material is used or when binding information is relevant (Olson, Page, Moore, Chatterjee, & Verfaellie, 2006). Second, similar memory structures were found active in both memory tasks (Nyberg, Forkstam, Petersson, Cabeza, & Ingvar, 2002) which suggest that the same neural structures may provide both types of memory. The latter finding, however, was only shown for a small set of features and it was not shown for ERP components under conditions where a short-term paradigm is used as study task for the long-term memory test within the same subjects. This was actually the aim of the present study. Demonstrating by such a procedure that the same neural structures are involved in processes contributing to working and long-term memory would be a strong argument in favor of common neuronal networks for these two tasks. Therefore our participants were required to solve an S1 – S2 matching task with auditory and visual items and afterwards their long-term memory was tested for these items. During both tasks EEG was recorded in order to investigate the topography of neural activity in dependence of the item's modality.

Findings from brain imaging studies suggest that long-term memory (LTM) representations are stored in distributed cortical structures which were originally involved in perceptual processing and encoding. According to the reinstatement hypothesis, the formerly generated representation of an item is re-activated within material-specific brain structures during retrieval. The most compelling evidence for this hypothesis arises from brain imaging studies showing distinct patterns of re-activations for different types of material-specific information (Damasio, 1989; Haxby et al., 2001; Vaidya, Zhao, Desmond, & Gabrieli, 2002). For example, in an event-related functional MRI study, Wheeler et al. (Wheeler, Petersen, & Buckner, 2000) have demonstrated that regions within the visual and auditory cortex were re-activated during the retrieval of sensory-specific information. In this study, a label was presented to the subjects followed by either the according picture or the according sound and subjects were required to memorize either the pictures or the sounds with their according labels. In the recall task, only the labels were presented and subjects had to retrieve the respective items from LTM. Retrieval of sounds was accompanied by bilateral activations near superior temporal gyrus – an auditory processing region – whereas the fusiform gyrus – a visual object processing region – was selectively active during retrieval of pictures. Furthermore, those re-activated regions overlapped with regions that were found active during a perception task that subjects were required to perform prior to the actual experimental session.

More specifically, remembering different types of visual information (e.g. objects, faces, and locations) caused distinct patterns of activation in occipital-temporal cortex and fusiform cortex (Henson et al., 2003; Khader, Burke, Bien, Ranganath, & Rösler, 2005; Ranganath, Cohen, Dam, & D'Esposito, 2004). Using a delayed-paired associate task (DPA) in an event-related fMRI study, Ranganath (Ranganath et al., 2004) demonstrated that retrieval of faces was associated with greater signal changes in the fusiform face area (FFA) and retrieval of houses with changes in the parahippocampal place area (PPA). In a training phase, subjects learned associations between a face and a house. During DPA trials either a face or house was repeated as cue and the paired object had to be retrieved. During the cue presentation, face cues were associated with greater activations in FFA while activations for house cues yielded an increase of activation in PPA. However, in the delay period, activations in both regions switched according to the target feature, i.e. the retrieval of face targets was associated with higher changes in the BOLD-signal in FFA, although a house cue was processed just before. The opposite pattern was obtained for the parahippocampal place area whenever houses had to be retrieved from long-term memory. Additional evidence for a re-activation of the visual processing stream in LTM tasks was provided by Slotnick (Slotnick & Schacter, 2006). During study, they presented shapes either on the left or the right side of a fixation cross so that they were initially processed only within one – the contralateral – hemisphere. During testing, shapes were presented in the center of the screen and subjects were instructed to remember each shape and its location (source memory). Neuronal activity within the right fusiform gyrus (BA 18) was associated with shapes previously presented on the left side of the fixation cross. In contrast, shapes presented on the right side during encoding, caused increased activations in left lingual gyrus (BA 18) during recognition. In a PET study, Ueno (Ueno et al., 2007) found overlapping activation in the inferior occipital gyrus during encoding and retrieval of color information.

However, not only brain imaging data suggest modality-specific memory entries. Also ERP data provide evidence that material-specific brain structures are active during LTM target processing and the retrieval of information from long-term memory (Cycowicz, Friedman, & Snodgrass, 2001; Johansson, Stenberg, Lindgren, & Rosén, 2002; Khader, Heil, & Rösler, 2005). For example, Khader et al. (Khader et al., 2007) reported more negative going slow-wave potentials for the recall of positions at parietal and occipital electrode sites compared to the recall of object information. Retrieval of objects instead evoked more negative potentials at frontal and anterior temporal electrodes compared to positions. In addition to the EEG experiment, the same subjects participated in an fMRI experiment (using the identical design) several days later. Dipole source localization of the measured ERP data revealed an overlap between the EEG dipole model and the activations observed in fMRI. Additionally, dipole strength corresponded to the relative amount of fMRI activation. Taken together these results provide evidence for material-specific LTM representations and their re-activation during retrieval. Further evidence was put forward in source memory tasks asking for specific knowledge about previously learned episodes. Johansson (Johansson et al., 2002) investigated recognition memory for previously perceived and imagined pictures. In general, ERPs for correct source memory judgments were more negative going than those for correct rejections and the negative slow-wave potentials had a posterior scalp distribution. Although the functional basis of this negative slow-wave is still a subject of discussion (Johansson & Mecklinger, 2003), is it likely that these deflections are linked to mnemonic rather than response-related processes (as suggested by Wilding & Rugg, 1997). Similarly, ERP studies using the visual half-field paradigm revealed memory effects reflecting the retinotopic cortical organization during perception (Fabiani, Stadler, & Wessels, 2000). Gratton et al. (Gratton, Corballis, & Jain, 1997) presented line patterns either on the right or the left side of a fixation cross. During the test, stimuli were presented centrally and subjects had to make old/new judgments, i.e. whether the stimulus was presented during encoding or not. Correctly remembered old items previously shown in the left visual field (right hemisphere) were associated with a right lateralized negativity whereas this pattern was reversed for correct old items presented in the right visual field (left hemisphere).

Taken together, findings from brain imaging as well as ERP studies strongly support the view that content-specific brain structures are re-activated during the retrieval of stored long-term memory information. Retrieval of those information is indicated by increased activation in material-specific brain structures (fMRI) and by slow-wave potentials (EEG) with their corresponding topographies.

The question is whether the same structures are also active in working memory. A first hint is the logical inference that successful retrieval implicates previous successful encoding and consolidation. If working memory provides stimulus encoding one should expect neural activity in corresponding brain structures. Also at a conceptual level, many similarities between encoding and working memory can be found, for example, updating and manipulating information. Therefore, it is likely that mental representations maintained in working memory and those retrieved from long-term memory are structurally similar and recent neuropsychological findings support this suggestion. In an event-related fMRI study, subjects performed a visual working memory task prior to a visual associative memory task (Ranganath, Johnson, & D'Esposito, 2003). Conjunction analysis revealed an overlap in activations for both encoding and retrieval phase of the WM and the LTM task. Encoding was associated with overlapping activations in right and left inferior frontal gyri (BA 44, 45, and 47) and left posterior middle frontal gyrus (BA 9). During the retrieval phase, overlapping activations were observed in right and left inferior frontal gyri (BA 44, 45, and 47), right and left posterior middle frontal gyri (BA 9), left anterior middle frontal gyri (BA 10/46), and right superior frontal gyrus (BA

10). Consequently, neural activity during processing of information in WM can predict LTM performance (Blumenfeld & Ranganath, 2006; Davachi, Maril, & Wagner, 2001) and interference during WM delay can cause significant decrease also in subsequent long-term memory performance (Ranganath, Cohen, & Brozinsky, 2005). Therefore, it can be argued that maintenance of information in WM contributes to successful long-term memory which suggests common neuronal structures.

Given this strong relationship between WM and LTM and the evidence for content-specific storage sites in long-term memory one should expect that content- or sensory-specific brain regions show specific activation patterns during WM maintenance. In fact, a broad range of fMRI experiments brought evidence for this assumption. For example, in an S1 – S2 paradigm, Courtney and colleagues (Courtney, Ungerleider, Keil, & Haxby, 1997) instructed subjects to keep images of faces in mind over a short delay period. During maintenance of face information, they found selective activations in mid-to-anterior fusiform gyrus (BA 37). Similarily, Druzgal and D'Esposito (Druzgal & D'Esposito, 2001) reported activation in the fusiform face area during an n-back task for faces, and this activation increased with working memory demands. Ishai and colleagues (Ishai, Haxby, & Ungerleider, 2002) reported activity in face specific areas during mental imagery of faces which were generated either from short-term or from long-term memory. Song and Jiang (Song & Jiang, 2006) investigated WM for colors and shapes (polygons). They found feature-unspecific but load dependent activation in the prefrontal cortex and feature-specific activation in occipito-temporal regions. These and many other studies suggest that WM for modality specific contents is provided by the corresponding sensory-specific brain structures.

In a similar vein, negative slow potentials in ERP studies were observed over sensory-specific brain regions during maintenance. In a phonological delayed match-to-sample task negative slow-wave potentials were found over left frontal brain regions (Ruchkin, Johnson Jr, Grafman, Canoune, & Ritter, 1992) and the amplitudes of these potentials were directly related to working memory demands. Therefore, it can be assumed that the distribution of the slow-wave potentials reflect specific phonological processing. Content-specific topographies of slow potentials were also reported for visual WM (see Zimmer (in press) for an extended discussion). Using geometric shapes as stimuli, slow potentials occurred over parietal areas for spatial and over frontal areas for object tasks and they increased with the number of objects (Mecklinger & Pfeifer, 1996). For 3D nonsense shapes, spatial tasks showed again a parietal negative potential and object tasks showed an occipital positive potential (Bosch, Mecklinger, & Friederici, 2001). Similarly, topographies for face and object processing could be dissociated while the topography for each feature representation was the same for an acquisition phase as in a subsequent long-term memory task (Khader, Heil et al., 2005).

In summary, fMRI as well as ERP data suggest that brain regions engaged in perceptual processing are activated during maintenance in WM tasks and during retrieval from LTM. Considering the operations taking place during LTM encoding, we argue that processing information in WM provides memory representations of items which are also available in subsequent long-term memory tasks. These entries are re-activated during retrieval (reinstatement). As a consequence, the same sensory-specific brain regions should be active during maintenance of information in WM and during retrieval from LTM. An implication hereof is that negative slow-wave potentials should have similar scalp distributions during WM tasks and LTM retrieval. Khader (Khader, Heil et al., 2005) reported empirical data in support of this assumption. They recorded slow potentials during encoding and a subsequent LTM task and found comparable topographies. However, during the acquisition phase they presented the items three times and the intervals between the repetitions were long (up to 53 items). It therefore cannot be excluded that retrieval operations occurred also during encoding. Hence encoding

(WM processing) might have been contaminated by LTM retrieval. In the present study, we wanted to avoid this by using a classical WM task (S1 – S2 matching task) as encoding task for a later long-term pair-associate test. At the same time, we wanted to extend the empirical basis for common working and long-term memory structures by demonstrating reinstatement in further sensory modalities. For that purpose, we presented mixed pairs of an auditory and a visual stimulus during encoding and EEGs were recorded during the S1 – S2 matching and the subsequent LTM task. At study we instructed participants to encode the two pair elements that were simultaneously presented and to associate the two objects. In order to introduce maintenance processes, we then required participants to mentally "refresh" one of the pair elements by presenting the other element as cue in order to match this memorized stimulus with a target item. By that either a visual or an auditory stimulus was maintained. In the LTM task, we presented either the auditory or the visual element of the pair as cue and instructed participants to retrieve the originally paired element. After that the target was presented and participants had to decide whether it was the correct pair element or not (a recombined target or a new target).

The goal of the present study was two-fold. First, if WM can be conceived as re-activated LTM then maintenance and retrieval processes should operate on partially identical representations. The second aim was to demonstrate that these brain activities during retrieval and maintenance are modality-specific. With regard to the above mentioned neuropsychological results we argue that both types of processes evoke modality-specific negative slow-wave potentials over sensory-specific brain areas. Hence, the topographies of the slow-wave potentials should be comparable for the two types of memory tasks (WM and LTM), but still be different according to the used target modality. We expected slow-wave potentials with a frontal scalp distribution during the maintenance and the retrieval of sound information and posterior scalp distributions for visual information. Finally, in order to separate target processing from maintenance and retrieval processes, our focus during the WM task was on the maintenance phase (after cue presentation) and on the retrieval phase during the LTM task (also after the cue presentation). If WM maintenance and LTM retrieval recruit similar networks then the topographies of slow potentials should be comparable in both tasks.

Methods

Subjects

Subjects were 23 students from Saarland University who provided informed consent and were paid for their participation. None of them had prior experience with the task. All subjects had normal or corrected-to-normal vision. Due to incomplete datasets, 3 subjects were excluded from further analyses, resulting in 20 subjects (14 females, mean age 24.4 years, range 19 – 40 years).

Material and Design

Stimuli were 200 realistic pictures and 200 sounds of common objects. Pictures as well as sounds were chosen from four different categories (instruments, animals, domestic appliances and humans). From these items the experimental stimulus material was generated. Each item pair consisted of a picture and a sound. Pictures and sounds were pseudo-randomly combined so that the two elements were semantically unrelated. All sounds were edited to a duration of 4500 ms and then presented on one of four different positions (front-left, front-right, back-right and back-left) via four loudspeakers. These sounds were recorded with a spatially realistic head-microphone (soundman, okm-II professional). After digitalization, sounds were normalized for loudness and a short fade-in was chosen to ensure that the beginnings of the sound presentation were approximately identical in volume. By this procedure, positional information was still available when sounds were presented with stereo headphones so that each auditory stimulus was heard from one of the four virtual locations around the participant. Pictures were resized to 160 pixels along their principal axis preserving proportions. This resulted in an angle of vision of 3°, approximately. During the experiment, pictures were presented on four different positions on the screen (top-left, top-right, bottom-right and bottom-left) comparable to the four sound locations. To minimize eye movements during the experiment positions for pictures were arranged adjacent to the center of the screen. Non-matching stimuli for the working memory task were generated by manipulating perceptual features in each modality. Non-matching pictures resulted from changing color saturation and orientation (mirror-reversals). Non-matching sounds were generated by applying different kinds of echoes, mechanical distortions and delay effects to the original sounds. All changes were small and difficult to be verbalized in order to force subjects to encode and to concentrate on perceptual features. In pretests, we checked that these manipulations could be detected. Locations of stimuli were varied in order to enhance the sensory variability of input but locations always matched between study and test.

In the S1 – S2 matching task, half of the targets were sounds and half were pictures, half of the stimuli were presented as matches and the other half as non-matches. For non-matching items, perceptual features changed between the S1-presentation and the S2-target-presentation. Assignment of item status (match/non-match) was counterbalanced across subjects. The same stimuli were used in the subsequent LTM task. Now, items were presented as intact or recombined pairs, and additionally, new items were presented. For the long-term memory task, the correct pair information was of interest and therefore sensory information was not manipulated. For the new item condition, meaningless visual and auditory stimuli (geometric structures and sine-waves) were additionally constructed. These items served as cues in the "new" target condition to avoid that participants accidentally attempted to retrieve any target information.

During study, we presented these picture-sound pairs in an S1 – S2 matching task. After study, a paired associate task followed. One of the old items or a new item was presented as cue, and after a short retrieval phase the target was presented to be judged as same pair, recombined, or new. On third were old items with a matching pair element, one third were old items with a recombined element, and one third were new items. Again, half of the targets were pictures and half sounds. Assignment of items to target status was counterbalanced across subjects.

Procedure

Prior to the experimental sessions, subjects worked through three practice lists to get familiar with the experimental procedure. After each list an overall feedback was presented on the screen. Only subjects who reached a criterion of at least 70 % accuracy in both tasks were allowed to participate in the actual experiment. One subject did not reach this criterion and did therefore not participated in the experiment.

Each experiment consisted of ten runs of a WM block (S1 – S2 matching task) followed by a LTM block (pair-associate task). The WM block served as encoding phase for the LTM task. Sixteen pairs of objects were presented during one WM block. Each WM trial (see Fig. 1) started with a fixation cross shown in the center of the screen (1000 ms) followed by a stimulus pair (4500 ms). The picture and the sound were simultaneously presented at one of the four positions[1]. Subjects were instructed to a) encode all the specific perceptual details of both objects and b) to encode the pair association. Then the fixation cross was briefly presented (1000 ms) as warning cue, followed by a frame at the pictures' location. This frame was visible until the end of the trial to minimize eye movements. Together with the frame one pair element (a sound or a picture)[2] was presented as cue for 3000 ms and an empty maintenance interval of 5000 ms followed. In eight trials a picture was presented as cue and a sound had to be maintained in working memory, and in eight trials a sound was presented and a picture had to be maintained. Subjects were instructed to rehearse the percept of the non present pair element (S1) during the empty interval in order to match the rehearsed item with the following target (S2). To announce the target stimulus, the frame turned into green after 4500 ms and 500 ms later the target stimulus appeared. Subjects had to decide whether S2 was identical to S1 or whether any perceptual feature had changed. Participants indicated their answers by pressing the "c" or the "m" button of the keyboard with their index fingers. Assignment of Item status (match, non-match) and response buttons were counterbalanced across subjects.

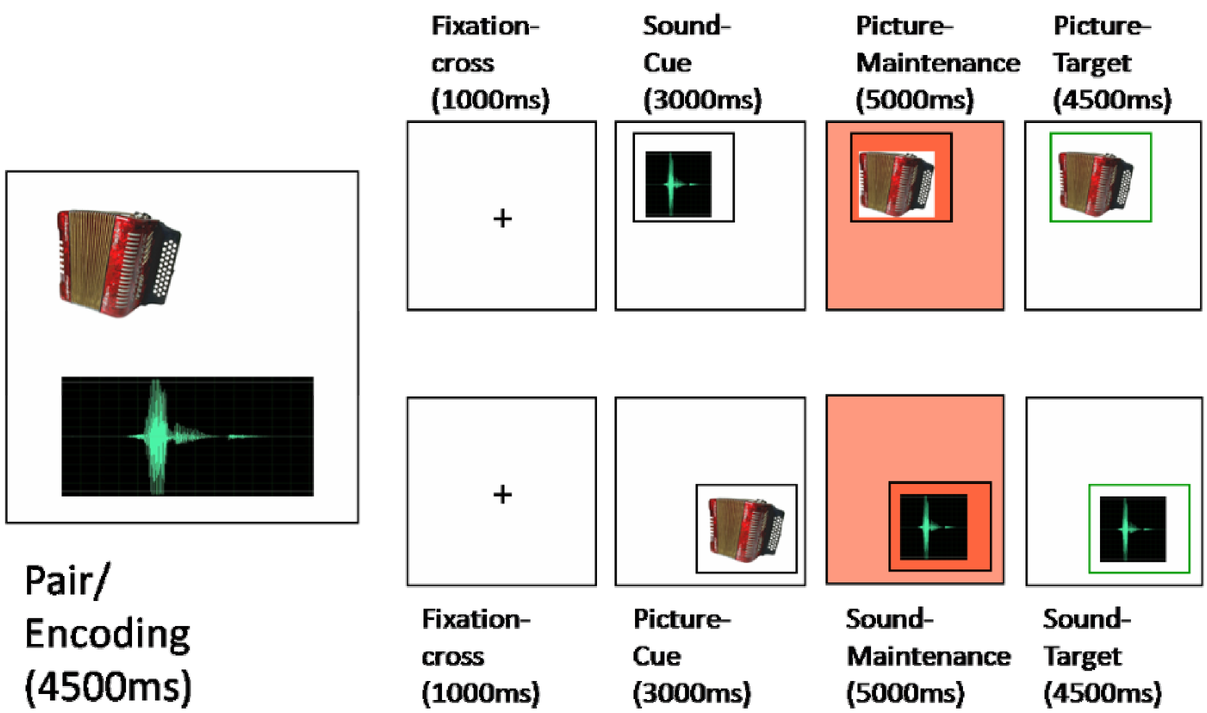


Figure 1. An illustration of the trial structure of the WM task (upper row: picture target condition; lower row: sound target condition). Note, during the maintenance interval (colored in red) subjects had to refresh the information and no stimulus was presented.

In the LTM block (Fig. 2), eight pictures and eight sounds from the prior WM block were repeated as cue (cue/old condition). LTM trials started with the presentation of a fixation

[1] Positions of pictures and sounds were identical in 25% of the cases. In the remaining 75% of the cases, positions were pseudo-randomized.

[2] In the picture cue condition the picture appeared within the frame, in the sound condition it remained empty.

cross in the center of the screen for 1000 ms and then either a picture or a sound from the prior WM task was presented as retrieval cue. The cue had a duration of 4500 ms followed by an empty interval of 3000 ms. Subjects were instructed to encode the cue and to retrieve the associated object (the target) when the cue presentation ended. After this interval the target stimulus was presented (for 3000 ms), which either was the correct pair element of the encoding phase in the WM trial or an old recombined object. In case of the four new items (two pictures and two sounds) a meaningless stimulus was presented as cue, and a new real object as target. Subjects had to decide whether the target was "correct", or incorrect – "recombined" or "new". Answers were given by pressing the "c" and "m" buttons. The assignment of item status (same pair, recombined, new) and response buttons were counterbalanced across subjects. Additionally, after each response, subjects indicated whether they had retrieved the target right after the cue presentation or whether they had checked the relation only after target presentation.

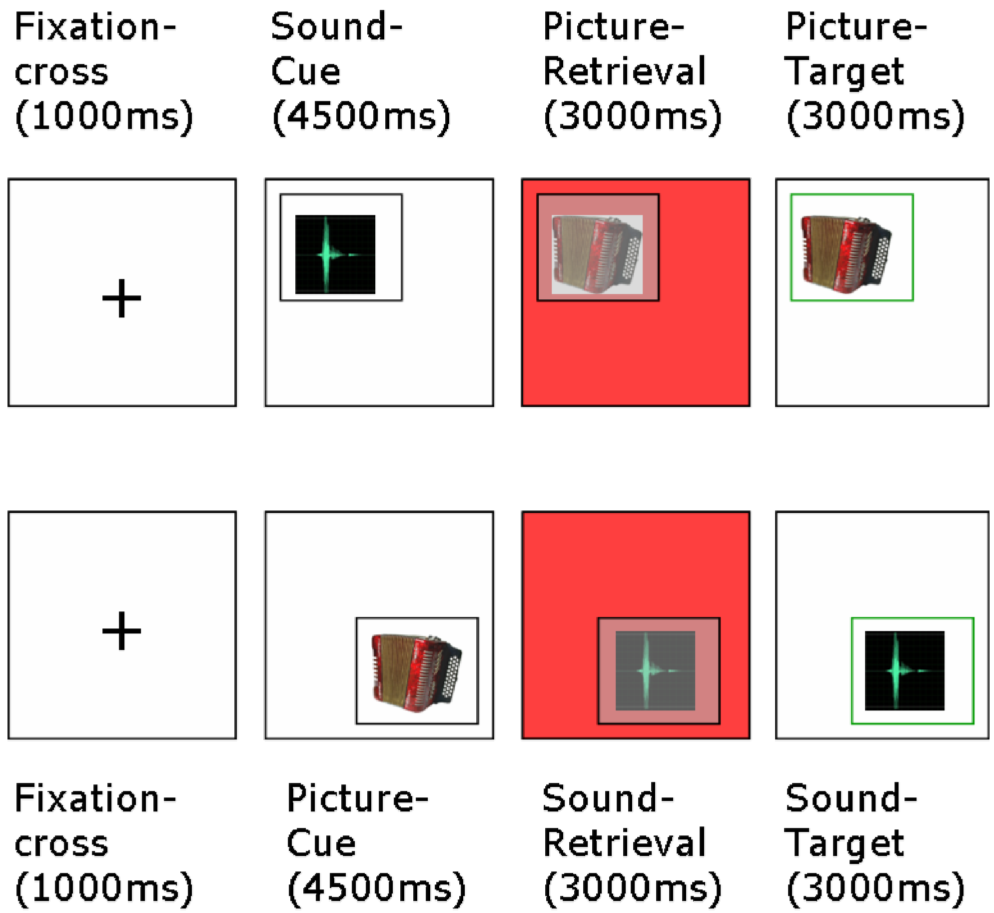


Figure 2. An illustration of the trial structure of the LTM task (upper row: picture target condition; lower row: sound target condition). Again, during the red marked interval (retrieval), no stimulus was presented to the subjects.

ERP recordings and analyses

The experimental session was run in a sound- and electromagnetically shielded chamber. EEG activity was recorded from 64 Ag/AgCl electrodes mounted in a preconfigured elastic cap

(System Falk Minow, Munich, Germany). They were arranged according to the extended international 10-20-system. All scalp electrodes were referenced on-line to the left mastoid and re-referenced off-line to linked-mastoids. Vertical and horizontal EOG were monitored and eye blinks were corrected semi-automatically off-line according to Gratton et al. (Gratton, Coles, & Donchin, 1983). Impedances of all electrodes were kept below 10 kΩ. Signals were digitized by a DC coupled amplifier (Brain Amp MR, Brain Products, Munich) with a sampling rate of 500 Hz. DC drift was corrected according to Hennighausen et al. (Hennighausen, Heil, & Rösler, 1993). EEG recordings were first segmented according to the phases in the WM and LTM trials due to different durations. For WM trials three ERPs were extracted, beginning 200 ms prior to the stimulus onset (durations given in parentheses): encoding (4700 ms), cue presentation and maintenance (8200 ms), and target (4700 ms). LTM trials were segmented into two parts: cue plus retrieval (7700 ms), and target (3200 ms). All data were baseline-corrected with respect to the 200 ms pre-stimulus interval and a low-pass filter was set to 20 Hz (slope 48 dB). Trials containing artifacts (maximum difference of two values in the interval 250 µV) were excluded from further analyses. ERPs were then computed for two conditions in WM maintenance trials (picture, sound) and for four conditions in the LTM retrieval trials (picture/old, picture/new, sound/old, sound/new). For statistical analyses mean voltages were computed for three (WM) and two (LTM) time bins, each 1000 ms, respectively, starting after the cue offset signal: WM maintenance (3500 – 4500 ms, 4500 – 5500 ms, 5500 – 6500 ms) and LTM retrieval (5000 – 6000 ms, 6000 – 7000 ms). Additionally, early and late components known from recognition memory tasks were investigated during the target processing. Time windows were defined by visual inspection of the ERPs for pictures and were adopted for the working memory task. Therefore, mean voltages for four WM target conditions (picture/match, picture/non-match, sound/match and sound/non-match) and for six LTM target conditions (picture/same pair, picture/recombined, picture/new, sound/same pair, sound/recombined, and sound/new) were also computed for two time windows: WM target presentation (250 ms – 450 ms, 450 ms – 700 ms), LTM target presentation (250 – 450 ms, 450 ms – 700 ms).

Results

Behavioral data

Response times and accuracies were analyzed for the test conditions in the WM and the LTM tasks. In both tests the data provide evidence for sufficiently high memory performances. Reaction time and accuracy data are depicted in Fig. 3 and Fig. 4. Only same pair and recombined conditions were presented for the LTM task because new items were already indicated as new ones during the cue presentation.

Working Memory

A repeated measures ANOVA on reaction times in the working memory task with the factors Modality (visual, auditory) and Item status (match, non-match), revealed a main effect of Modality, $F(1,19) = 114.32$, $p < .001$, as well as an interaction between Modality and Item status, $F(1,19) = 6.82$, $p < .05$.

The comparable ANOVA on accuracy data yielded a different pattern of result. For the proportion correct, a significant main effect of Modality, $F(1,19) = 68.18$, $p < .001$, a main effect of Item status, $F(1,19) = 44.17$, $p < .001$, and an interaction effect between Modality and Item status, $F(1,19) = 56.40$, $p < .001$, were observed. Tukey HSD indicated a significant difference between match and non-match items for both modalities (all p < .001) which was clearly higher for sounds than for pictures. Sensory mismatches were sometimes overlooked which caused a reduced rate of correct rejections for non-matching items compared to hits for matching ones, especially for sound targets.

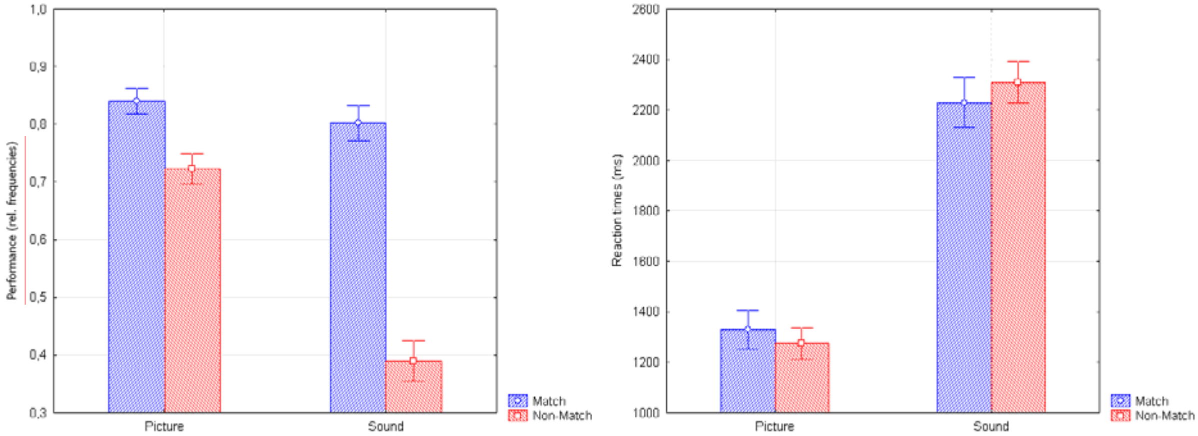


Figure 3. Mean response times for correct trials and accuracies (proportion correct) in the WM task (match: hit, non-match: correct rejection). Bars indicate standard errors.

Long-Term Memory

A 2 × 3 repeated measures ANOVA with the within-subject factors Modality (visual, auditory) and Item status (same pair, recombined, new) on hit reaction times indicated a significant main effect of Modality, $F(1,19) = 133.51$, $p < .001$, and a main effect of Item status, $F(2,38) = 132.96$, $p < .001$. Post-hoc testing (Tukey HSD) revealed that response times for new pictures and new sounds were significantly faster than for same pair and recombined pictures as well as for same pair and recombined sounds (all ps < .001). Differences concerning reaction times for same pair vs. recombined items were neither significant for pictures nor for sounds.

A comparable analysis of accuracy data yielded a main effect of Item status, $F(2,38) = 19.07$, $p < .001$, and an interaction of Modality by Item status, $F(2,38) = 5.27$, $p < .01$. As expected, new items were judged nearly perfectly as new although for sounds a few errors occurred. However, also item binding was sufficiently remembered. About eighty percent of the pairings were recognized with a small tendency to falsely reject a correct sound target as recombined ($p < .05$). Obviously, the item association was well remembered.

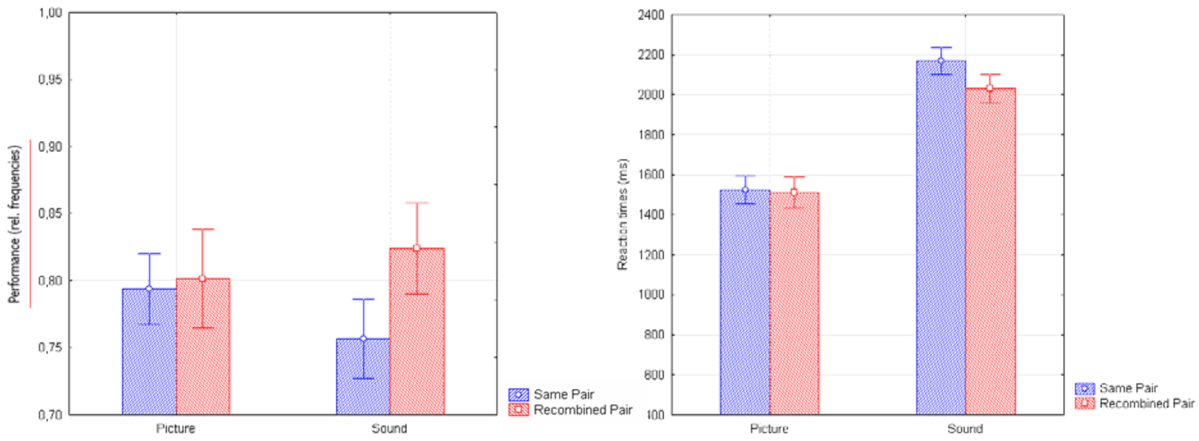


Figure 4. Mean response times for correct trials and accuracies (proportion correct) in the LTM task (same pair: hit, recombined and new: correct rejection). Bars indicate standard errors.

Event-related potentials

For the statistical analyses, one electrode from each ROI was selected, resulting in a 3 by 3 schema defined by Location (anterior, central, posterior) and Laterality (left, midline, right). Therefore, the following electrodes were included in the statistical analyses: F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4. The same analyses were computed for the nine ROIs (anterior-left: F3, FC3, FC5, anterior-midline: FC1, FC2, Fz, central-left: C3, CP3, CP5, central-midline: CP1, CP2, Cz, parietal-left: P3, P5, PO3, parietal-midline: O1, O2, Pz, and the right counterparts). Although the same pattern of result was observed for the ROI analyses, the effects for single electrodes reached higher levels of significance and therefore are reported.

WM maintenance and LTM retrieval

ERP data for WM maintenance were analyzed by means of analyses of variance (ANOVA) with the factors Location (3), Laterality (3) and Modality (visual, auditory). ANOVAs were computed separately for each time window. For all three time windows, only a main effect of Modality was observed (3500 – 4500 ms: $F(1,19) = 9.73$, $p < .05$; 4500 – 5500 ms: $F(1,19) = 9.84$, $p < .05$; 5500 – 6500 ms: $F(1,19) = 9.23$, $p < .05$). ERPs for picture maintenance were in general more negative than ERPs elicited during the maintenance of auditory information, and this effect was observed at all electrode locations (see Fig. 5.a, left).

For LTM retrieval (Fig. 5.a, right), ERP differences were examined by computing an ANOVA, including Location (3), Laterality (3), Modality (2), and additionally Item status (old, new) as factors. For the first time interval, a significant main effect was observed for Modality, $F(1,19) = 6.54$, $p < .05$. In addition, the four-way interaction for Location x Laterality x Modality x Item status was also significant, $F(4,76) = 2.63$, $p < .05$. Retrieval of visual information was associated with a more negative deflection at all electrodes in contrast to the retrieval of sound information which elicited more positive waveforms. In the 6000 – 7000 ms interval, the ANOVA yielded a main effect of Item status, $F(2,38) = 5.12$, $p < .05$. As can be seen in Fig. 5.b, although differences in amplitudes were obtained, topographies were quite similar for both modalities and therefore not further analyzed.

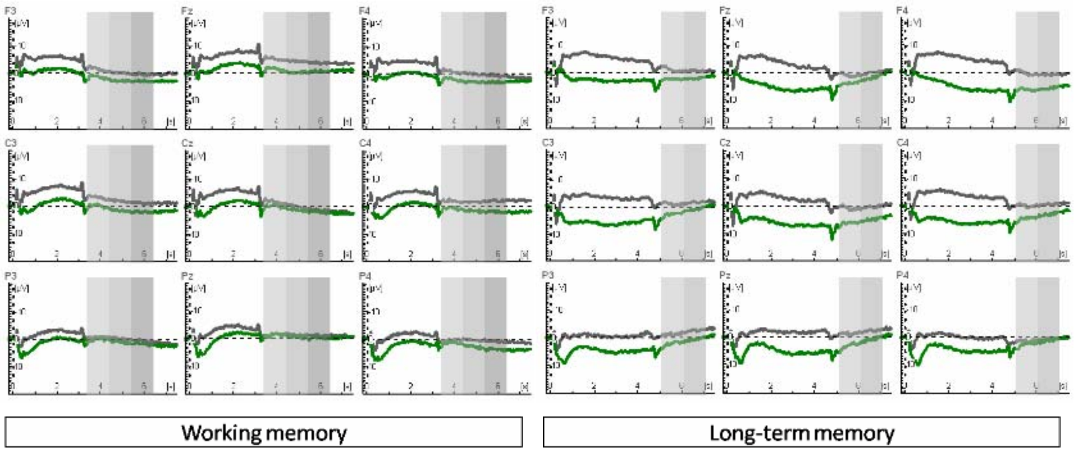


Figure 5. a) Grand average waveforms for working memory maintenance (left) and LTM retrieval (right) at the 9 electrode locations (grey: picture, green: sound), positive values are displayed downwards.

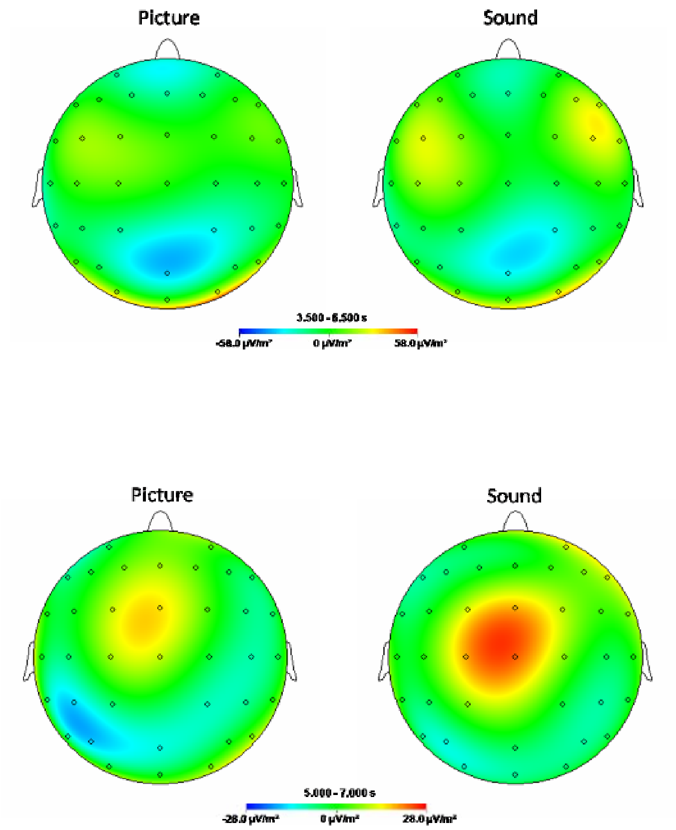


Figure 5.b) Current source density maps for working memory maintenance (upper row) and LTM retrieval (lower row).

WM target

An ANOVA with the factors Location (anterior, central, posterior), Laterality (left, midline, right), Modality (visual/auditory), and Item status (match/non-match) was computed for each time window, respectively. For the first interval (250 – 450 ms) main effects of Modality, $F(1,19) = 98.54$, $p < .001$, and Item status, $F(1,19) = 15.16$, $p < .001$, were observed as well as the interaction effects Location x Item status, $F(2,38) = 4.53$, $p < .05$, $\varepsilon = .65$, and Laterality x Item status, $F(2,38) = 3.84$, $p < .05$, $\varepsilon = .89$. For the second time interval (450 – 700 ms), ANOVA again yielded main effects of Modality, $F(1,19) = 243.69$, $p < .001$, and Item status, $F(1,19) = 54.03$, $p < .001$, and additional interaction effects of Laterality x Modality, $F(2,38) = 3.33$, $p < .05$, and Modality x Item status, $F(1,19) = 11.47$, $p < .05$. The interaction of Location with Item status was marginally significant, $F(2,38) = 3.30$, $p = .07$, $\varepsilon = .71$. For further analyses of the main effect of Item status, separate ANOVAs were conducted for pictures and sounds for each time window, respectively.

During the early time window an ANOVA for **pictures** (Fig. 6.a, left side) revealed a main effect of Item status, $F(1,19) = 18.95$, $p < .001$, and an interaction effect of Location by Item status, $F(2,38) = 6.90$, $p < .05$, $\varepsilon = .77$. Post-hoc testing of the interaction effect revealed that matching pictures were, in general, more negative than non-matching pictures, but the difference was reduced (still significant) at posterior electrode locations (Fig. 6.b). For the second time interval (450 – 700 ms) only the main effect of Item status, $F(1,19) = 58.48$, $p < .001$, reached significance, i.e. non-matching had a greater positive deflection as compared to matching pictures.

For the **sound** target condition (Fig. 6.a), right), an ANOVA did not yielded significant results for the early time window. In the second time interval, the main effect of Item status, $F(1,19) = 4.49$, $p < .05$, was reliable. As for the picture target condition, non-matching sounds had more positive going waveforms than did matching sounds.

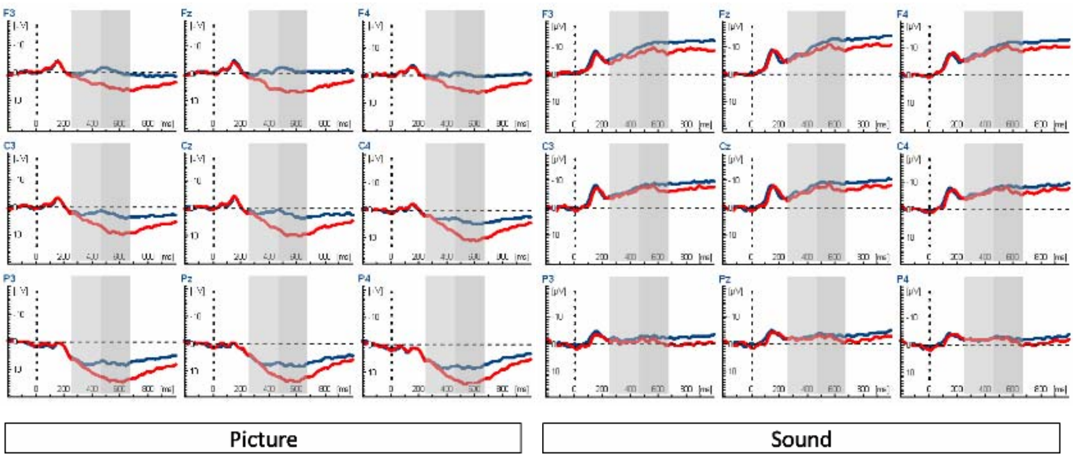


Figure 6.a) Grand average waveforms for working memory target presentation (blue: match, red: non-match) at the 9 electrode locations (left: picture, right: sound), positive values are displayed downwards.

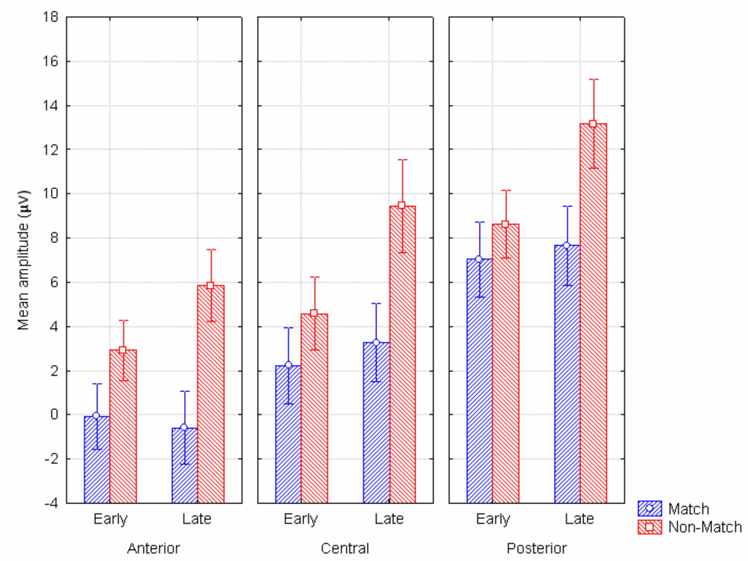


Figure 6.b) Mean amplitudes for the picture target condition in the working memory task. Bars indicate standard errors.

LTM target

To investigate the ERP results during target presentation an ANOVA with the factors Location (anterior, central, posterior), Laterality (left, midline, right), Modality (visual, auditory), and Item status (same pair, recombined, new) was performed for each time window, respectively (see Fig. 7.a and 7.b). Results for the first time window (250 – 450 ms) yielded main effects of Modality, $F(1,19) = 25.33$, $p < .001$, of Item status, $F(2,38) = 7.06$, $p < .05$, $\varepsilon = .69$, and significant interactions of Laterality × Modality, $F(2,38) = 10.33$, $p < .001$, $\varepsilon = .84$, Location × Item status, $F(4,76) = 4.63$, $p < .05$, $\varepsilon = .57$, Modality × Item status, $F(2,38) = 34.89$, $p < .001$, $\varepsilon = .69$, and Location × Laterality × Modality, $F(4,76) = 6.02$, $p < .05$, $\varepsilon = .67$. Post-hoc testing (Tukey HSD) of the interaction Modality with Item status revealed that same pair and recombined pictures had more positive going waveforms than ERPs for same pair and recombined sounds. New items from both modalities did not differ significantly. In the second time window (450 – 700 ms), the main effect of Modality, $F(1,19) = 114{,}00$, $p < .001$, Item status, $F(2,38) = 4.70$, $p < .05$, $\varepsilon = .80$, as well as the following interactions became significant: Location × Item status, $F(4,76) = 7.89$, $p < .001$, $\varepsilon = .58$, Modality × Item status, $F(2,38) = 8.83$, $p < .05$, $\varepsilon = .85$, Location × Laterality × Modality, $F(4,76) = 4.85$, $p < .05$, $\varepsilon = .80$. For further investigations of the manipulation of item status, i.e. changes of the pair associations, separate analyses for pictures and sounds were computed.

For **pictures**, a 3 × 3 × 3 ANOVA with the factors Location (anterior, central, posterior), Laterality (left, midline, right), and Item status (same pair, recombined, new) revealed a significant main effect of Items status, $F(2,38) = 29.99$, $p < .001$, $\varepsilon = .65$, as well as a significant interaction of Location with Item status, $F(4,76) = 3.13$, $p < .05$, $\varepsilon = .61$ in the early time window. Post-hoc testing of the interaction effect indicated a graded effect at centro-parietal electrode sides in which same pair and recombined pictures differed from new ones and in addition same pair pictures had a greater positive shift than recombined pictures. At frontal electrodes the difference between same pair and recombined pictures proved not to be significant. For the second time interval, only the main effect of Item status, $F(2,38) = 12.30$, $p < .001$, $\varepsilon = .87$, was reliable. Post-hoc testing indicated a significant same pair/new- and recombined/new-effect, i.e. ERPs for both old conditions were more positive going than ERPs evoked by new pictures. Waveforms for same pair pictures did not differ significantly from those for recombined pictures (see Fig. 7.a and 7.b, left).

In the **sound** target condition effects of manipulating pair associations were investigated by computing an ANOVA with the factors Location (3), Laterality (3), and Item Status (3). During the first time window, only the interaction Location by Item status, $F(4,76) = 3.65$, $p < .05$, $\varepsilon = .53$, reached significance. Tukey HSD post-hoc test revealed that effects of Item status were maximal over fronto-central electrode locations. While for frontal electrodes same pair and recombined sounds differed from new ones, only the difference between recombined and new sounds was reliable at central electrode locations. At posterior electrode sites neither same pair nor recombined sounds deviated significantly from new sounds. Furthermore, no significant differences between same pair and recombined sounds were observed. In the time interval between 450 – 700 ms, again only the interaction of Location with Item status, $F(4,76) = 10.08$, $p < .001$, $\varepsilon = .49$, reached significance. Post-hoc testing indicated an overall old/new-effect (same pair and recombined) at frontal electrodes in which old items elicited waveforms with greater positive shifts than did ERPs for new sounds. Additionally, recombined sounds differed from new ones at posterior electrode locations (Fig. 7.a and 7.b, right).

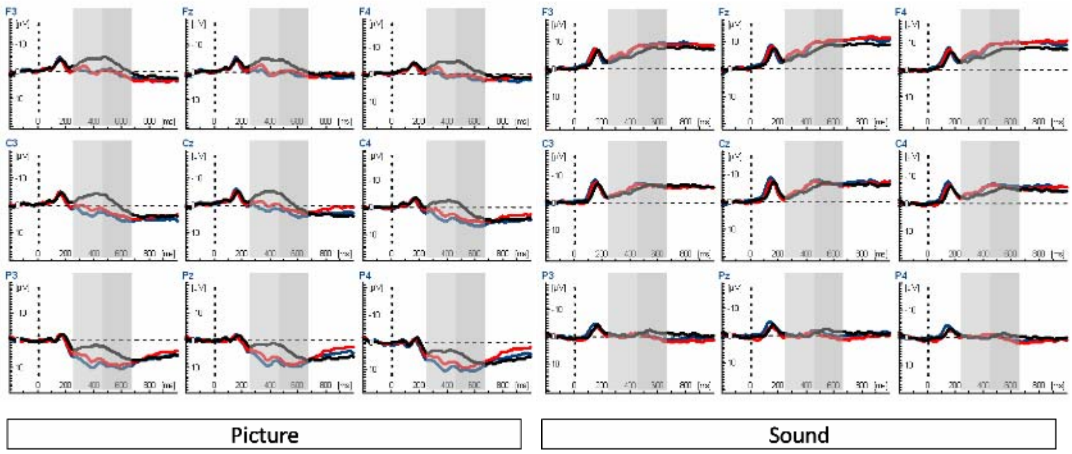


Figure 7.a) Grand average waveforms for long-term memory target presentation (black: new, blue: same pair, red: recombined). ERPs for pictures are displayed on the left and ERPs for sounds on the right side.

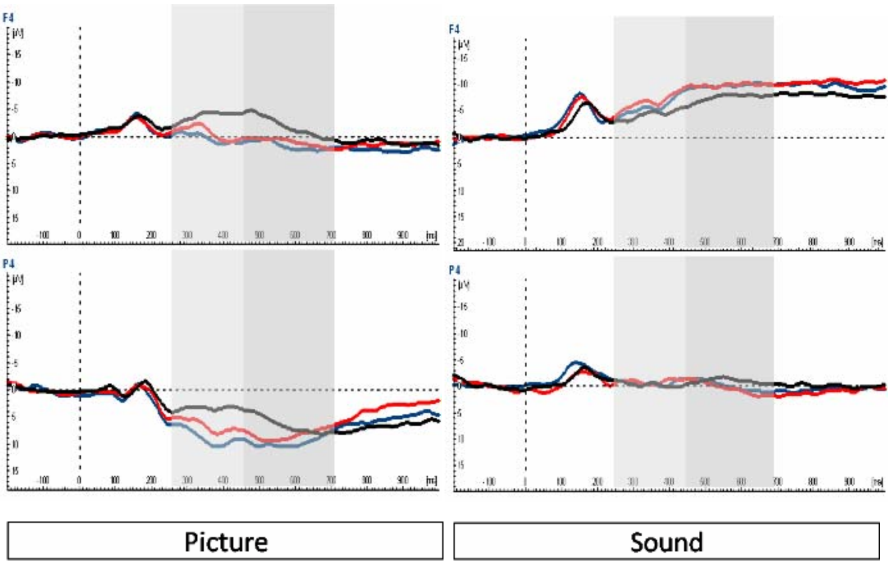


Figure 7.b) Grand average waveforms for long-term memory target presentation (left: picture, right: sound) at the F4 (upper row) and the P4 (lower row) electrodes (black: new, blue: same pair, red: recombined).

Topographic analyses

In order to investigate, whether the observed effects were due to differences in topographies or whether they were caused by differences in mean amplitudes, ANOVAs were performed for vector-scaled data, including the factors Location (anterior, central, posterior), Laterality (left, midline, right), and Item status (same pair, recombined, new) for the LTM task and Location (anterior, central, posterior), Laterality (left, midline, right) and Item Status (match, non-match) for the working memory task.

In a first step, we tested whether pictures and sounds differed according to their scalp distribution in the working and long-term memory task, respectively. For both time intervals of the long-term memory task, a significant main effect of Modality, $F(1,19) = 28.09$, $p < .001$ and $F(1,19) = 120.48$, $p < .001$, respectively, were obtained. In addition, the ANOVA revealed significant interactions of Location by Modality, $F(2,38) = 5.63$, $p < .05$, $\varepsilon = .64$, for the first time window and $F(2,38) = 5.00$, $p < .05$, $\varepsilon = .67$ for the second. An ANOVA for ERP topographies in the working memory task indicated a significant main effect for Modality, $F(1,19) = 98.54$, $p < .001$, for the time window between 250 and 450 ms and a main effect of Modality, $F(1,19) = 243.69$, $p < .001$, for the late interval. Results for the WM as well as for the LTM task indicated modality-specific scalp distributions for the observed effects.

In a second step, we investigated differences in the topography between the two types of tasks by computing an ANOVA with the factors Type of task (WM, LTM), Location (anterior, central, posterior), Laterality (left, midline, right), and Modality (visual, auditory). Because non-match items and recombined items differ according the perceptual features[3], only ERPs for identically presented test stimuli were analyzed. For the first time interval, main effects for Type of task, $F(1,19) = 5.24$, $p < .05$, and Modality, $F(1,19) = 73.91$, $p < .001$, were significant. Additionally, the three-way interaction Type of task x Laterality x Modality, $F(2,38) = 3.67$, $p < .05$, $\varepsilon = .92$, became significant. The same main effects were also obtained for the second time window (Type of task: $F(1,19) = 8.90$, $p < .01$; Modality: $F(1,19) = 185.56$, $p <$

[3] For non-match items, perceptual features were changed whereas perceptual features for recombined items were identical for encoding and test.

.001). No other interaction with Type of task reached significance for the early interval or for the late one, respectively (see Fig. 8.a – 8.c).

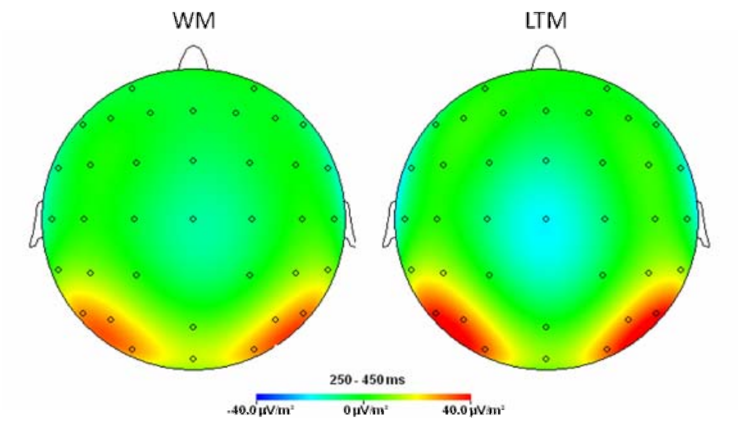


Figure 8.a) Current source density maps for picture target presentation in the WM task (left) and the LTM task (right).

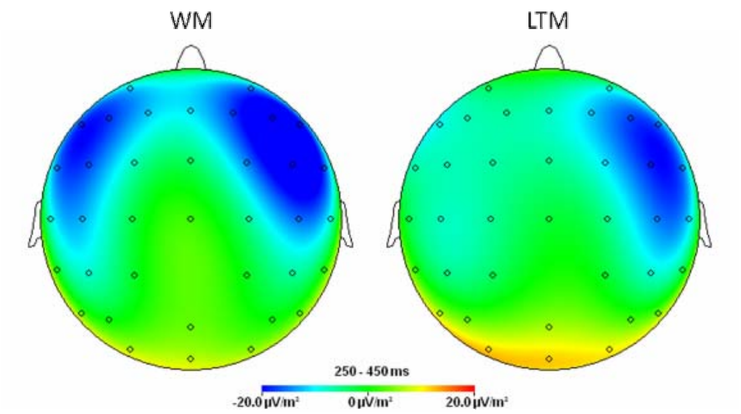


Figure 8.b) Current source density maps for sound target presentation in the WM task (left) and the LTM task (right).

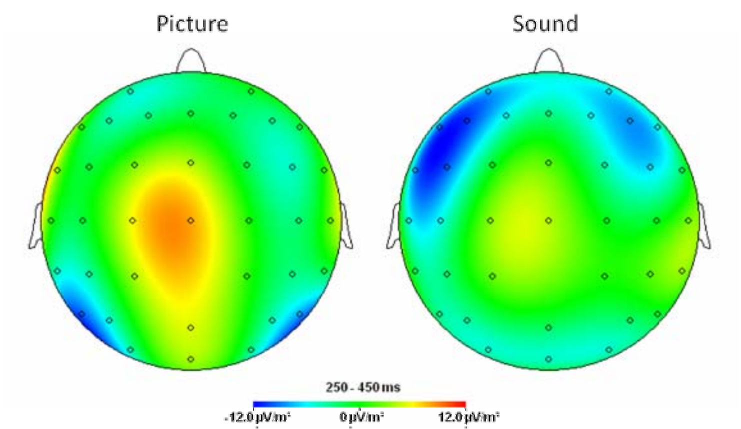


Figure 8.c) Current source density maps of different wave (WM – LTM) of the target presentation during the early time window (left: picture, right: sound).

Discussion

In the present study we focused on modality-specific representations in long-term and working memory. Neuropsychological studies suggest that, according to the reinstatement hypothesis, retrieval of stored mental representations lead to re-activations of neuronal structures originally involved in perceptual processing. Event-related fMRI studies further revealed a partial overlap of activity during WM tasks and LTM tasks. In this context, WM is often discussed as re-activated long-term memory, implicating that representations in WM and LTM, at least, overlap with each other. In addition, activation patterns obtained in fMRI studies and scalp distributions of slow-wave potentials observed in ERP studies suggest that re-activated networks are specific to memory content. We examined modality-specific modulations (visual vs. auditory) in ERPs recorded during an S1 – S2 matching task and a pair-association task as well as topographic differences or similarities between tasks and/or modalities.

We were specifically interested in activations and re-activations of early sensory brain regions because both item types should activate distinct cortical brain regions, namely primary visual or auditory cortex. We predicted that those regions are activated and re-activated, during maintenance or retrieval, respectively, and this activation is reflected in different topographies of recorded slow-wave potentials.

WM maintenance and LTM retrieval

As opposed to our expectation, ERP results did not indicate modality-specific processes during maintenance or retrieval. Also the absence of a significant interaction effect of Location by Modality seems to speak against the assumption that sensory information is actively rehearsed during WM maintenance and actively retrieved from LTM at the time of the cue presentation. Contrary, more negative potentials were observed for visual than for auditory information during maintenance, even at frontal electrode locations, and this polarity is opposite to the one expected. However, this overall negativity for picture maintenance should be interpreted with caution because it could be a direct consequence of cue processing. It may result from differences between early components of perceptual processing of the cue, which was a sound for picture targets and vice versa. While the picture cue presentation elicited a pronounced N1 - P2 complex, sound cues were only associated with an N1 component. Statistical analyses of the cue presentation revealed that these early components of the ERPs differed significantly for pictures and sounds (see Fig. 5)[4]. Those differences continued over the whole epoch of cue presentation and might even have extended into the maintenance interval. However, besides this main effect no evidence for specific slow-wave potentials during the maintenance and the retrieval interval were apparent, although such components were reported by other authors (Bosch et al., 2001; Khader, Heil et al., 2005).

One explanation for this divergent result would be that participants were rather passive during the maintenance and retrieval interval and waited for the target presentation. Especially in the WM task, they might have directly used the target as retrieval cue relying on the bottom-up

[4] During the cue presentation early ERPs effects due to sensory processing were observed. While sounds elicited an early frontal negative component, ERPs for pictures were associated with a positive posterior deflection.

triggered information[5]. Because our paradigm asks for a judgment only after target presentation this strategy might have been adopted by subjects. Furthermore, statistical analyses, including time window as an additional factor, yielded neither a significant main effect nor a significant interaction with the factor Time for working memory maintenance. Therefore, slow-wave potentials associated with working memory processes were not observed during the maintenance interval. We assume that in this study the intended maintenance interval did not require additional working memory processes. Instead, those processes were postponed to the moment of the target presentation. This result is in line with other behavioral studies (Quinn & McConnell, 2006; Zimmer, Speiser, & Seidler, 2003) reporting that interference during the maintenance did not affect performance in working memory tasks.

However, negative slow-wave potentials were observed during the cue presentation after the N1 – P2 complex and this negativity was distributed over the whole scalp. As can be seen in Fig. 5 the time course of this slow potential was identical for both modalities, thus speaking rather for modality-unspecific processes. The onset of the slow-wave potentials indicated that participants may already have used the cue presentation for WM operations. This strategy did not allow for any modality-specific modulations of slow-wave potentials during the intended maintenance, because processing of the cue as well as the imagined target lead to activations of sensory-specific brain regions, respectively. In this case, modality-specific activations would be overlapping. This assumption is supported by fMRI results (Ranganath, 2006) showing co-activations of the fusiform face area already during the presentation of house cues and vice versa. During the delay interval the activation in the FFA increased for face targets and became higher than the activation for house targets. In a similar vein, co-activations during the cue presentation in our study might have lead to general negative slow-wave potentials, instead of showing modality-specific modulations.

ERP components during target presentations

Recent studies investigating feature-specific manipulations have shown that this kind of information directly influences several ERP components. For example, episodic recognition studies have shown that target processing can be accompanied by a negative slow-wave potential (Cycowicz et al., 2001). The authors suggest that this negative potential might index the search and/or the retrieval of sensory-specific information. In addition, several studies from our lab (Groh-Bordin, Zimmer, & Ecker, 2006; Groh-Bordin, Zimmer, & Mecklinger, 2005) indicated that changing sensory information between encoding and test affected the FN400 (familiarity) and the late parietal component (recollection). According to the dual-process account, two functional distinct processes, familiarity and recollection, can be dissociated in recognition tasks. In this account, familiarity refers to a general feeling of having encountered an object before, without any conscious access to details of the study episode while recollection refers to conscious retrieval of details. In the context of electrophysiological measures, familiarity is associated with an early midfrontal old-new effect, beginning at approximately 300 – 500 ms post-stimulus onset. In contrast to old repeated items, new items show a more negative deflection (FN400). On the other hand, the electrophysiological correlate of recollection (the late positive component, LPC) occurs at (left-) parietal electrode sites in which old items are more positive going than new items. In conclusion, in our study effects due to the specific manipulation of the target's modality should be obtained in the ERP recorded during the target processing.

[5] Instead of actively refreshing the associated item as they were instructed to do.

Long-term Memory. Long-term memory target processing did not require the retrieval of any sensory-specific information. During the long-term memory task, information of a pair association (picture-sound or sound-picture) had to be retrieved subsequent to the cue presentation. However, if the item is re-instantiated in a network comprising early sensory brain regions, specific modulations of the memory content (visual or auditory information) should be observable in the respective scalp distributions during target processing.

The results for the **picture** target condition yielded a significant old/new-effect during the early and late time interval at all electrode locations. Same pair and recombined pictures showed a larger positive deflection than did new pictures. While for central and parietal electrode sites same pair and recombined pictures differed reliably, this effect was far from reaching significance at frontal electrode locations. This graded effect diminished in the time interval from 450 – 700 ms, resulting in a general same pair/new- and recombined/new-effect, i.e. whenever the target item is the same as the retrieved one or the target item is an old one but not belonging to the cued position, waveforms were more positive than waveforms for new pictures. Similar findings were observed in studies manipulating object-context associations (Ecker, Zimmer, Groh-Bordin, & Mecklinger, 2007; Tsivilis, Otten, & Rugg, 2001) and it is thought that the early frontal old/new effect is an index of familiarity while the parietal component indexes recollection.

Despite the fact that general old/new effects were observed in the present study, manipulation of pair information (same pair vs. recombined) did not result in significant differences for the FN400 and the LPC. In other words, recombined associations neither influenced familiarity nor recollection. However, and most critical in the context of the reactivation theory, an additional effect was observed at parietal electrode locations during the early time window. Here the correct target picture elicited a greater positive deflection than did recombined pictures. Interestingly, this effect was maximal over those electrode locations were visual-specific processing was expected. As indicated by topographical analyses, scalp distributions were identical for same pair and recombined pictures resulting in the fact that differences were only due to amplitude differences and therefore varied as a function of neuronal activity.

However, the nature of this observed effect is not clear. One possible explanation is that differences between same pair and recombined items during the early time window reflect the beginning of the late parietal component, because subjects were able to retrieve the information prior to the target presentation. But two findings speak against this assumption. First, reaction times for the picture target decision were quite long. If subjects already knew which item should appear as target item we would expect that the target item is primed. This should lead to faster reaction times because representations of the target item should be activated and this should fasten the identification of the object, and therefore the decision. Second, statistical analyses of vector-scaled data (including Time window as factor) revealed significant interactions with the factor Time window, indicating different topographies for the parietal effect in the early and that in the late time window. According to those differences, indicating different processes, it is more likely that the observed effects resemble modality-specific processing.

Concerning the **sound** target processing memory related differences were found at frontal electrode locations during the early time window. Same pair and recombined sounds had a more negative deflection compared to new presented sounds. However, the difference between same pair and recombined sounds was not significant. While the difference between recombined and new sounds were still reliable at central electrode locations it was diminished for same pair sounds, indicating that the observed effect was limited to frontal electrodes for same pair sounds. The effect for recombined sounds, instead, had a fronto-central scalp distribution[6]. For the late time interval, we found reliable effects for same pair and recombined sounds with a frontal scalp distribution and for recombined sounds at parietal electrode locations. Although those effects likely represent the FN400 and the LPC, they were heavily reduced and therefore hard to interpret.

Taken together, the obtained ERP differences between same pair/recombined picture stimuli in LTM occurred over those electrode sites where sensory-specific processing was expected[7], supporting the view that brain regions engaged in the perceptual processing were re-activated during LTM operations.

Working Memory. During the working memory task high demands on sensory-specific processing were required because small variations of different perceptual features had to be detected. Therefore, not only the maintenance of modality-specific information but also decisions about the target stimulus should lead to activation of neuronal networks involved in perceptual processing.

For the **picture** target condition we observed an early effect due to sensory non-match items. During the first time interval, pictures started to differ at approximately 200 ms after stimulus onset in which non-matching pictures had significantly more positive-going waveforms compared to matching pictures. As indicated by the observed interaction effect, this difference between matching and non-matching pictures was maximal over fronto-central electrode locations and reduced at parietal electrode sites. In the late time interval this interaction did not reach significance, resulting in a main effect of item status. In other words, changes of visual information additionally affected ERPs observed at parietal electrode locations during the early interval of the present working memory task. ERPs for non-matching pictures elicited a greater positive deflection than ERPs for matching pictures. Again, analyses of vector-scaled data, including time window as an additional factor, indicated that the topographies of the posterior effects were different for both time windows. In sum, additionally to the early frontal effect we observed an early effect with a posterior distribution due to perceptual changes. The distribution of this effect differed significantly from the effect obtained in the second time interval. This effect is similar to the observed effect during the LTM task (see discussion above). Since this effect was due to changes either in the pair-association or in perceptual details, our results provide some support for an activation or re-activation of early sensory brain region, respectively.

An ANOVA for the sound target condition did not reveal further evidence for modality-specific processing during the first time window which could be related to memory processes. However, in the second time interval ERPs for non-matching sounds were more positive going than ERPs for matching sounds. This effect was maximal over frontal electrode locations. As can be seen in Fig. 7.b), this effect had a negative-going distribution over frontal electrode sites and a positive deflection at parieto-occipital electrode sites. Most importantly, changing features of sounds had an impact on the waveforms for matching and non-matching items and this effect was maximal over frontal electrode locations. In conclusion, the distribution of the

---

[6] Vector-scaled comparisons indicated a significant difference between frontal and central electrode locations for old sounds while the same contrast did not reach significance for recombined sounds.

[7] It should be noted that sensory processing of the target stimuli was not required during the long-term memory task.

ERP effects obtained in the second time window additionally supports our hypothesis, that sensory-specific processing in working memory activates early sensory-brain regions.

Similarities between LTM and WM

Recent fMRI studies reported an overlap of activity between working memory and long-term memory tasks, suggesting that both mental representations are supported by identical neuronal networks. In this context, it is assumed that during working memory tasks, long-term memory representations are re-activated and manipulated in order to solve the task. Additional evidence arose from ERP studies (Khader, Heil et al., 2005) which reported no topographic differences of ERP slow-wave potentials between the two types of tasks. According to these results, we expected that topographies still differ for Modality but not for the Type of task. The present results indicated a significant main effect of Type of task for both time intervals and a three-way interaction during the early time interval. While the main effects might be due to amplitude differences, the interaction effects indicated different topographies for the early effects observed in the working memory and the long-term memory task. As can be seen in Fig. 7b) retrieval of sounds was accompanied by a lateralized topography over the right hemisphere. This finding fits well in the HERA (hemispheric encoding/retrieval asymmetry; model (Nyberg et al., 1996)). In this model, encoding is associated with left lateralized processing while the retrieval of information from long-term memory causes activations within the right hemisphere. Most important for our assumption that topographies did not differ according to the Type of task is the fact that the interaction of Location by Modality failed to reach significance. Except for the lateralization effects during encoding and retrieval, which can be explained by the HERA model, results from our study support the assumption that WM and LTM operations rely on similar neuronal networks, including early modality-specific brain regions.

Conclusion

ERPs for the maintenance and the retrieval interval did not reveal any modality-specific activation, raising the question whether those processes did not require the activation of early sensory brain regions or whether maintenance and retrieval were rather passive. Slow-wave potentials already obtained during the cue presentation favor the assumption that subjects started to maintain the respective associative information during this period, resulting in an overlap of activity in both sensory-specific brain regions. Nevertheless, effects due to perceptual manipulation and manipulation of the pair information, respectively, were found for the target presentation in both types of tasks, especially for pictures. Although the nature of the obtained effect is not clear it was caused by differences between representations maintained in working memory and retrieved from long-term memory, respectively, and actual picture target presentation. Importantly, this effect occurred over those electrode locations were we expected visual processing. Therefore, the general pattern of our results from the target presentations suggest, that modality-specific processing takes places during the working memory and long-term memory task. The present data support the idea, that a neuronal network is activated or reactivated, respectively, including early sensory brain regions. Furthermore, the similarities between the topographies for the WM and the LTM tasks provided support for the idea, that representations in WM and LTM were related to similar neuronal networks.

Acknowledgements

This research was supported by the Deutsche Forschungsgemeinschaft (DFG, grant SFB 378 „Ressourcen-adapaptive Prozesse“). We wish to thank Kathrin Hilberer and Iris Wiegand for research assistance.

References

Blumenfeld, R. S., & Ranganath, C. (2006). Dorsolateral prefrontal cortex promotes long-term memory formation through its role in working memory organization. *Journal of Neuroscience, 26*(3), 916-925.

Bosch, V., Mecklinger, A., & Friederici, A. D. (2001). Slow cortical potentials during retention of object, spatial, and verbal information. *Cognitive Brain Research, 10*(3), 219-237.

Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature, 386*(6625), 608-611.

Cycowicz, Y. M., Friedman, D., & Snodgrass, J. G. (2001). Remembering the color of objects: An ERP investigation of source memory. *Cerebral Cortex, 11*(4), 322-334.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. *Cognition, 33*(1-2), 25-62.

Davachi, L., Maril, A., & Wagner, A. D. (2001). When keeping in mind supports later bringing to mind: Neural markers of phonological rehearsal predict subsequent remembering. *Journal of Cognitive Neuroscience, 13*(8), 1059-1070.

Druzgal, T. J., & D'Esposito, M. (2001). 13. Working memory for faces: An fMRI study of prefrontal and extrastriate visual cortex. *Brain and Cognition, 47*(1-2), 72-75.

Ecker, U. K. H., Zimmer, H. D., Groh-Bordin, C., & Mecklinger, A. (2007). Context effects on familiarity are familiarity effects of context - An electrophysiological study. *International Journal of Psychophysiology, 64*(2), 146-156.

Fabiani, M., Stadler, M. A., & Wessels, P. M. (2000). True but not false memories produce a sensory signature in human lateralized brain potentials. *Journal of Cognitive Neuroscience, 12*(6), 941-949.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55*(4), 468-484.

Gratton, G., Corballis, P. M., & Jain, S. (1997). Hemispheric organization of visual memories. *Journal of Cognitive Neuroscience, 9*(1), 92-104.

Groh-Bordin, C., Zimmer, H. D., & Ecker, U. K. H. (2006). Has the butcher on the bus dyed his hair? When color changes modulate ERP correlates of familiarity and recollection. *NeuroImage, 32*(4), 1879-1890.

Groh-Bordin, C., Zimmer, H. D., & Mecklinger, A. (2005). Feature binding in perceptual priming and in episodic object recognition: Evidence from event-related brain potentials. *Cognitive Brain Research, 24*(3), 556-567.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293*(5539), 2425-2430.

Hennighausen, E., Heil, M., & Rösler, F. (1993). A correction method for DC drift artifacts. *Electroencephalography and Clinical Neurophysiology, 86*(3), 199-204.

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., & Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex, 13*(7), 793-805.

Ishai, A., Haxby, J. V., & Ungerleider, L. G. (2002). Visual imagery of famous faces: Effects of memory and attention revealed by fMRI. *NeuroImage, 17*(4), 1729-1741.

Johansson, M., & Mecklinger, A. (2003). The late posterior negativity in ERP studies of episodic memory: Action monitoring and retrieval of attribute conjunctions. *Biological Psychology, 64*(1-2), 91-117.

Johansson, M., Stenberg, G., Lindgren, M., & Rosén, I. (2002). Memory for perceived and imagined pictures - An event-related potential study. *Neuropsychologia, 40*(7), 986-1002.

Khader, P., Burke, M., Bien, S., Ranganath, C., & Rösler, F. (2005). Content-specific activation during associative long-term memory retrieval. *NeuroImage, 27*(4), 805-816.

Khader, P., Heil, M., & Rösler, F. (2005). Material-specific long-term memory representations of faces and spatial positions: Evidence from slow event-related brain potentials. *Neuropsychologia, 43*(14), 2109-2124.

Khader, P., Knoth, K., Burke, M., Ranganath, C., Bien, S., & Rösler, F. (2007). Topography and dynamics of associative long-term memory retrieval in humans. *Journal of Cognitive Neuroscience, 19*(3), 493-512.

Mecklinger, A., & Pfeifer, E. (1996). Event-related potentials reveal topographical and temporal distinct neuronal activation patterns for spatial and object working memory. *Cognitive Brain Research, 4*(3), 211-224.

Nyberg, L., Forkstam, C., Petersson, K. M., Cabeza, R., & Ingvar, M. (2002). Brain imaging of human memory systems: Between-systems similarities and within-system differences. *Cognitive Brain Research, 13*(2), 281-292.

Nyberg, L., McIntosh, A. R., Cabeza, R., Habib, R., Houle, S., & Tulving, E. (1996). General and specific brain regions involved in encoding and retrieval of events: What, where, and when. *Proceedings of the National Academy of Sciences of the United States of America, 93*(20), 11280-11285.

Olson, I. R., Page, K., Moore, K. S., Chatterjee, A., & Verfaellie, M. (2006). Working memory for conjunctions relies on the medial temporal lobe. *Journal of Neuroscience, 26*(17), 4596-4601.

Quinn, J. G., & McConnell, J. (2006). The interval for interference in conscious visual imagery. *Memory, 14*(2), 241-252.

Ranganath, C. (2006). Working memory for visual objects: Complementary roles of inferior temporal, medial temporal, and prefrontal cortex. *Neuroscience, 139*(1), 277-289.

Ranganath, C., & Blumenfeld, R. S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Sciences, 9*(8), 374-380.

Ranganath, C., Cohen, M. X., & Brozinsky, C. J. (2005). Working memory maintenance contributes to long-term memory formation: Neural and behavioral evidence. *Journal of Cognitive Neuroscience, 17*(7), 994-1010.

Ranganath, C., Cohen, M. X., Dam, C., & D'Esposito, M. (2004). Inferior Temporal, Prefrontal, and Hippocampal Contributions to Visual Working Memory Maintenance and Associative Memory Retrieval. *Journal of Neuroscience, 24*(16), 3917-3925.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2003). Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia, 41*(3), 378-389.

Ruchkin, D. S., Johnson Jr, R., Grafman, J., Canoune, H., & Ritter, W. (1992). Distinctions and similarities among working memory processes: An event-related potential study. *Cognitive Brain Research, 1*(1), 53-66.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: a neuropsychological study. *The Quarterly journal of experimental psychology, 22*(2), 261-273.
Slotnick, S. D., & Schacter, D. L. (2006). The nature of memory related activity in early visual areas. *Neuropsychologia, 44*(14), 2874-2886.
Song, J. H., & Jiang, Y. (2006). Visual working memory for simple and complex features: An fMRI study. *NeuroImage, 30*(3), 963-972.
Tsivilis, D., Otten, L. J., & Rugg, M. D. (2001). Context effects on the neural correlates of recognition memory: An electrophysiological study. *Neuron, 31*(3), 497-505.
Ueno, A., Abe, N., Suzuki, M., Hirayama, K., Mori, E., Tashiro, M., et al. (2007). Reactivation of medial temporal lobe and occipital lobe during the retrieval of color information: A positron emission tomography study. *NeuroImage, 34*(3), 1292-1298.
Vaidya, C. J., Zhao, M., Desmond, J. E., & Gabrieli, J. D. E. (2002). Evidence for cortical encoding specificity in episodic memory: Memory-induced re-activation of picture processing areas. *Neuropsychologia, 40*(12), 2136-2143.
Warrington, E. K., & Weiskrantz, L. (1968). New method of testing long-term retention with special reference to amnesic patients. *Nature, 217*(5132), 972-974.
Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences of the United States of America, 97*(20), 11125-11129.
Wilding, E. L., & Rugg, M. D. (1997). Event-related potentials and the recognition memory exclusion task. *Neuropsychologia, 35*(2), 119-128.
Zimmer, H. D. (in press) Visual and spatial working memory: From boxes to networks. Neuroscience and Biobehavioral Reviews.
Zimmer, H. D., Speiser, H. R., & Seidler, B. (2003). Spatio-temporal working-memory and short-term object-location tasks use different memory mechanisms. *Acta Psychologica, 114*(1), 41-65.