On the Time Course of Naming Multidimensional Objects in a Referential Communication Task

Analyses of Eye Movements and Processing Times in the Production of Complex Object Specifications

Doctoral Dissertation

submitted by

Eva Belke

Department of Linguistics
(Fakultät für Linguistik und Literaturwissenschaft)
University of Bielefeld

Advisors and Reviewers:

Dr. Hans-Jürgen Eikmeyer, University of Bielefeld

Dr. Antje Meyer, University of Birmingham

PD Dr. Ulrich Schade, University of Bielefeld

December 2001

Gedruckt auf alterungsbeständigem Papier °° ISO 9706

“Among the most important functions of language is the communication of perceptual experience. Language affords each of us the ability to have a private perceptual experience and then tell other people what we have seen or heard. […] While this is obviously an important linguistic capacity, we know very little about the process by which people ‘transform’ their perceptions into language, nor about the processes by which people ‘transform’ someone else’s description into an ‘understanding’ of a perceptual experience.”

(Clark, Carpenter, & Just, 1973: 311)

Acknowledgement

I am indebted to the following persons (in alphabetical order) for supervision, technical support, helpful comments, proofreading, or for simply encouraging me during the last three years:

Stefan Barattelli

Herbert Baumann

Gerlind und Horst Belke

Kathryn Bock

Gary Dell

Ger Desserjer

Hans-Jürgen Eikmeyer

Gijs van Elswijk

Kathy Gabriel

Frouke Hermens

Willem Levelt

Dieter Metzing

Antje Meyer

Rolf Peter Otto

Kay Peterson

Gert Rickheit

Katharina Rohlfing

Ulrich Schade

Kirsten Schindler

Martin Schmitz

Max Sichelschmidt

Birgit Sommer

Johan Weustink

Danksagung

Im Frühjahr 1999 habe ich während eines Praktikums am Max-Planck-Institut für Psycholinguistik erstmals mit Augenbewegungsmessungen gearbeitet. Dabei konnte ich Kenntnisse erwerben, die für die Experimente im Rahmen der vorliegenden Arbeit unerlässlich waren. Frau Dr. Antje Meyer, die dieses Praktikum begleitet und mich hervorragend betreut hat, gilt mein besonderer Dank. Aufgrund des großen Vertrauens, das sie mir entgegenbrachte, durfte ich schon früh an aktuellen Projekten zu Augenbewegungsmessungen bei der Produktion gesprochener Sprache mitarbeiten. Gijs van Elswijk und Frouke Hermens haben mich in die professionelle Bedienung der Experimentalsteuerung und des Eye-Trackers eingewiesen und mich dadurch in den Stand versetzt, schon bald selbst Augenbewegungsexperimente durchführen zu können – eine unabdingbare Voraussetzung für die spätere Arbeit an der Dissertation.

Danken möchte ich ferner all jenen, die mich bei der Planung und Durchführung der Experimente unterstützt haben und ohne deren freundliche Hilfe, nicht zuletzt bei so manchem technischen Problem, ich die empirischen Daten für die vorliegende Arbeit nicht so zügig und reibungslos hätte erheben können (in alphabetischer Reihenfolge): Stefan Barattelli, Herbert Baumann, Ger Desserjer, Hans-Jürgen Eikmeyer, Gijs van Elswijk, Frouke Hermens, Hendrik Koesling, Antje Meyer, Ulrich Schade, Martin Schmitz, Max Sichelschmidt und Johan Weustink. Nicht vergessen seien all die Versuchspersonen, die an den Experimenten teilgenommen haben.

Dank schulde ich vor allem meinen Betreuern, Hans-Jürgen Eikmeyer, Antje Meyer und Ulrich Schade. Hans-Jürgen Eikmeyer und Ulrich Schade waren im Studium meine wichtigsten Lehrer – sie haben mich in vielerlei Hinsicht gefördert und unterstützt und mein Interesse an der Sprachproduktionsforschung zunächst geweckt und später voller Wohlwollen gefördert. Sie waren vor Ort stets für mich da, wenn ich inhaltliche bzw. organisatorische Fragen oder Schwierigkeiten hatte. Mein größter Dank gebührt der externen Betreuerin meiner Arbeit, Frau Dr. Antje Meyer, die mich – trotz räumlicher Entfernung – zunächst aus Nijmegen und später aus Birmingham kontinuierlich beraten hat. Sie hat ungeachtet ihrer vielfältigen Verpflichtungen regelmässige persönliche Treffen ermöglicht und sich dabei sehr viel Zeit für mich genommen. Diese Treffen haben mir immer wichtige Impulse für das weitere Vorgehen und neuen Schwung zum Arbeiten gegeben.

Die Stipendiaten des Graduiertenkollegs „Aufgabenorientierte Kommunikation“ waren als Kollegen und Diskussionspartner wichtig und gaben mir viel Rückhalt. Gemeinsam haben wir interessante Projekte bewerkstelligt und dabei wichtige Erfahrungen gemacht. Das Graduiertenkolleg bot exzellente Rahmenbedingungen, um meine Dissertation zügig voranzubringen und dabei gleichzeitig aktuelle Ergebnisse auf verschiedenen Konferenzen vor internationalem Publikum zu präsentieren und zu diskutieren. Dabei konnte ich wichtige Erfahrungen sammeln und viele interessante Kontakte zu anderen Promovenden und Forschern im Bereich der Psycholinguistik knüpfen. Im Sommer 2001 hatte ich die Möglichkeit an einer Sommerschule zur konnektionistischen Modellierung kognitiver Prozesse in Oxford teilzunehmen, bei der ich wichtige neue Inhalte und Kenntnisse erwerben konnte. Mein Dank gilt in diesem Zusammenhang besonders Prof. Dr. Dieter Metzing, der sich nachdrücklich dafür eingesetzt hat, dass ich zusammen mit Katharina Rohlfing an dieser Sommerschule teilnehmen konnte.

Ich bin meinen Eltern sehr dankbar dafür, dass sie mein Studium und meine späteren Aktivitäten stets unterstützt und mit Interesse begleitet haben. Meine Familie und meine Freunde, insbesondere Peter und Kathy, wissen, dass sie sehr wichtige Gesprächspartner für mich waren, die mir über manche Durststrecke hinweggeholfen haben!

Contents

1 Introduction

Referring to objects in the outside world is one of the fundamental functions of language. Although there are simple ways of referring to an object by non-verbal means, such as pointing at it, speakers usually specify the objects they are referring to verbally, in order to talk about them and to make the object name available for further discourse (cf. Pechmann, 1984; Pechmann & Deutsch, 1982). Depending on the complexity of the situation, referring expressions may differ with regard to their degree of elaboration: If there is only one exemplar of a given object type, the object's name should be sufficient to identify it; however, when speakers refer to an object in the context of several other objects, they often have to specify it by means of a set of features that clearly distinguishes it from the other objects (Olson, 1970). In order to assess the act of referring to an object, the referential communication task has been developed (Piaget, 1926). In this task, subjects are asked to name one of several multidimensional objects in such a way that a listener will be able to uniquely identify the intended object. Multidimensionality, in experiments on referential communication, generally means variation between objects in terms of dimensions such as color, size, and object class[1] (Danks & Schwenk, 1972; Eikmeyer & Ahlsén, 1998; Ford & Olson, 1975; Herrmann & Deutsch, 1976; Olson, 1970; Pechmann, 1989; Whitehurst, 1976).

Until now, little effort has been spent on investigating the procedural aspects of the cognitive processes involved in the production of complex noun phrases in a referential communication task. Pechmann (1989, 1994; see also Pechmann & Zerbst, 1994) was the first to conduct experiments on referential communication using reaction times as a dependent variable. His research yielded important insights into the nature of the relevant cognitive processes and opened up new directions in the empirical research on referential communication.

However, neither qualitative analyses of utterance structures nor reaction time measurements can fully capture the exact time course of the processes involved in generating complex noun phrases in a referential communication task, as they are

[1] The term object class is used in the logical sense of the term *class*. Thus, the object class of cap, for example, contains elements like a large red cap, a small red cap, a blue, and a green cap, but no other object except for a cap.

solely based on the analysis of overt reactions. Eye tracking techniques, in contrast, permit more detailed analyses of the time-course of the conceptual preparation and the formulation processes preceding the overt reaction (Meyer, Sleiderink, & Levelt 1998; Meyer & van der Meulen, 2000; Pechmann, 1989). Since they provide information about when and for how long objects in a display are viewed, they allow on-line measurements of the processes underlying referential communication, such as the evaluation of the referent object (*target object*) and its relation to the surrounding objects (*context objects*), the selection of properties of the target object to be verbalized and the formulation of the object's specification.

In the following, I will first give an introduction to the theoretical background of the investigation presented in this work and discuss open issues in empirical research on referential communication. Some of these issues were addressed in a series of experiments: Experiments 1 and 2 were designed to investigate the processes underlying stimulus discrimination and the relation between perceptual and linguistic encoding processes during the generation of complex object specifications. In Experiment 3, the results from the first two experiments were extrapolated to more complex object configurations, in order to explore in more detail the perceptual and procedural determinants of the *form* of complex object specifications. The results will be discussed in a general framework in view of procedural aspects of language production, shedding a new light on previous findings and open issues in the research on referential communication.

2 Empirical Framework and Theoretical Background

As indicated above, the procedural aspects of the cognitive processes involved in referential communication have not yet been explored in detail. The present work is embedded in two broad lines of empirical research:

- The work is tightly linked to present research on speech production processes: In the 1990ies, most of the empirical work on the production of referring expressions was focused on naming simple objects. Lexical access was one of the major research fields that were investigated by various experimental paradigms, most of them using one-word utterances (see Levelt, Meyer, & Roelofs, 1999, for a review). Although the theoretical views on lexical access processes are still controversial, there seems to be a shift of interest towards more complex utterances that provide insight to coordinative lexical processes during the production of syntagms and phrases (Levelt & Meyer, 2000; Meyer, 1996, 1997; Schriefers, 1992, 1993; Schriefers, de Ruiter, & Steigerwald, 1999; see also Caramazza & Miozzo, 1997; Miozzo & Caramazza, 1999; Schriefers & Teruel, 2000). The results obtained on lexical processes in complex utterance generation will have to be aligned with the wide range of empirical investigations on syntactic processes in sentence production (e.g., Bock, 1982; Bock, Loebell, & Morey, 1992; Griffin & Bock, 2000; Hartsuiker, Kolk, & Huiskamp, 1999).
- The experiments tie in with previous investigations on the linguistic form of object specifications in referential communication. The referential communication task has been widely used in language acquisition research, and many investigations therefore focused mainly on the information that is conveyed by the observed specifications (Deutsch & Pechmann, 1982; Ford & Olson, 1975; Sonnenschein, 1982, 1985; Whitehurst, 1976; Whitehurst & Merkur, 1977). Similarly, studies on adults' performance were primarily aimed at determining the impact of pragmatic variables, such as common ground, on the speaker's object specifications. There are only a few studies that focused mainly on the *form* of the utterances; most of them are dated back to the 1970ies and 1980ies (Byrne, 1979; Deutsch & Pechmann, 1982; Herrmann & Deutsch, 1976; Martin, 1969a; Olson, 1970; Pechmann, 1984; Danks & Schwenk, 1972). Extensive empirical research has been done in the framework of research projects on object

naming and on the variability of referential object specifications at the Universities of Marburg and Mannheim, Germany (cf. Deutsch, 1994; Mangold-Allwinn, 1994 for reviews and Herrmann & Deutsch, 1976; Mangold-Allwinn, Barattelli, Kiefer & Koelbing, 1995 for details).

In this chapter, after a short survey of object reference and naming in referential communication, I will point out which processing stages can be assumed to underlie complex object descriptions, and how information is processed within and between these stages. After that, the empirical framework of the present work will be introduced, and methodological issues of eye tracking as experimental technique will be considered in detail.

2.1 Object Reference and Naming in Referential Communication

In the present work, a distinction is made between the terms "naming" and "object reference" in the following sense: Following Stachowiak (1978: 208) *naming* mainly occurs in "fairly restricted situations", such as introducing each other or teaching and testing. Naming is typically verbal and is commonly used in experiments on the processes underlying speech production (cf. Bock, 1996, for an overview). *Object reference*, in contrast, is much less restricted and is ubiquitous in everyday interactions. Its general pragmatic aim is to identify a referent. Therefore, referring to objects is not necessarily verbal, and naming is but one special case of object reference. Pechmann and Deutsch (1982: 331) argue that, from a more general point of view, "object reference may be defined as an action by which one person tries to focus the attention of another person on a certain part of the environment." As indicated at the beginning, there are multiple ways to refer to objects – the most obvious being pointing at something that both speaker and listener can see (see also Terrace, 1985, for an evolutionary perspective on (non-)verbal reference). Using a referential communication task that allowed both verbal and gestural reference to an intended object, Pechmann and Deutsch (1982) investigated how frequently children (aged two to nine years) and adults use pointing gestures as opposed to verbal descriptions. When pointing gestures served a referential function, children and adults used them about equally often. At the same time, they tended to reduce the contextual detail of their verbal descriptions. However, when – due to spatial conditions – pointing was not sufficient for a successful identification of the intended object, the frequency of pointing decreased with age, while at the same time the frequency of adequate descriptions increased. The high frequency of ineffective pointing by younger children is not due to their difficulties in assessing the functionality of the pointing gesture in relation to the given spatial configuration. It rather supports the view that children use non-verbal means of reference before the appropriate linguistic means have been acquired (see Clark, 1978).

In the following, I will focus on genuinely verbal object reference only, i.e. reference to objects without the aid of pointing or other gestures.

2.1.1 Cognitive Determinants of Object Reference

The act of referring presupposes the ability to categorize perceptual data and to decide which features of the intended referent are relevant to identify it. Perceptual categorization or classification integrates two basic mechanisms, namely differentiation and generalization of features. Correctly assessing which features of a referent are relevant to identify it, minimally requires that a speaker is able to differentiate the referent from potential alternatives. Note that, for a unique object specification, a speaker needs not necessarily limit his description to the minimally distinctive features (*minimal distinctiveness*), but may mention non-distinctive features as well (*referential overspecifications*). Beyond that, there may be multiple ways to (minimally) refer to an object on different levels of specification (*degree of elaboration*): In the context of a *large red* ball, a *small green* ball might be referred to as “the green ball”, “the small ball”, or “the small green ball”. Similarly, a male person might be named *Daddy, Father, Heinz, Uncle, Love* and by many other terms (Herrmann & Deutsch, 1976).

The development of non-verbal cognitive abilities goes hand in hand with linguistic development. During language acquisition, children tend to overextend the semantic content of words or the use of syntactic rules (*overgeneralizations*; cf. E.V. Clark, 1973), which can be regarded as an instance of evolving differentiation and generalization processes. Haviland and Lempers (1984) found that children’s classification skills contribute to their performance on referential communication tasks. The impact of their performance on classification tasks requiring an understanding of common properties, such as a block-sorting task, a class-inclusion task, or a class-intension task, was highly dependent on age and vocabulary size. Similarly, Camaioni and Ercolani (1988) showed that comparison performance was significantly related to performance in referential communication tasks. In particular, young children with high comparison abilities performed similarly to older children with lower comparison abilities. Stachowiak (1978) reports studies that show that performance on the Token Test (de Renzi & Vignolo, 1962), which has proved to be a highly sensitive diagnostic tool for the detection of aphasic deficits, correlates not only with performance on language comprehension tests, as originally intended, but also with production tests (e.g. object naming). He argues that the Token Test, “although it concerns not the encoding but the decoding of the naming function, measures exactly these faculties” (Stachowiak, 1978: 214), i.e. the faculties to categorize and select features for the exclusion of referential alternatives.

2.1.2 Pragmatic Determinants of Object Reference

In everyday situations there are multiple pragmatic factors that determine object specifications during communication (Clark & Wilkes-Gibbs, 1986; Hupet & Chantraine, 1992). The common ground of the interlocutors plays an important role in deciding whether a message is informative or not (Fussell & Krauss, 1992; Horton & Keysar, 1996, 1998; Sodian, 1988; Sonnenschein, 1986, 1988). Young children’s failure to communicate an adequate and unambiguous manner is often attributed to their general difficulties in giving up their own perceptual or conceptual perspective (*egocentrism*; cf. Piaget, 1954; see also Sonnenschein, 1988, for a broader account on the basis of general communicative skills). In contrast, adult speakers tend to give redundant information in their object descriptions (*referential overspecifications,* see Pechmann, 1984).

Pechmann (1984) investigated the origins of referential overspecifications taking into account the speakers’ discourse model. He identified two types of referential overspecifications, marked by means of prosodic stress: *Endophoric* overspecifications are related to the preceding discourse. Speakers may specify redundant features of an object in order to contrast it with an object that has been described immediately before. *Exophoric* overspecifications, in contrast, are related to the set of contextual alternatives that are perceivable for both speaker and listener.[2] In his analyses, Pechmann found that extra stress is given on endophoric, but not on exophoric overspecifications. Exophoric overspecifications have proved to help the listener identify the target object (Mangold, 1986; Sonnenschein, 1982, 1984). Some authors even proposed that speakers use referential overspecifications strategically, to help the listeners find the referent (Mangold & Pobel, 1988).[3] Note, though, that particularly for the case of referential overspecifications, such pragmatic factors are closely intertwined with perceptual determinants of object reference: An important determinant of exophoric overspecifications seems to be the detectability of features, which is determined by the types of dimension involved (relative vs. absolute) and the distribution of features in the field of contextual alternatives (Herrmann & Deutsch, 1976; Herrmann & Grabowski, 1994). The latter aspect has been proved to affect feature selection in minimal specifications, too: When two dimensions are equally adequate to minimally specify the referent, speakers tend to specify the dimension that is easier to detect (Herrmann & Deutsch, 1976).

Speakers usually act according to conversational rules and conventions (Herrmann & Grabowski, 1994). Grice (1975) assumes that speakers, being co-operative, conform to a set of conversational maxims. The maxims of *quantity* and *manner* state that a speaker should be brief but informative and avoid obscurity of expression and ambiguity. Adult speakers normally perform well in terms of informativeness, since they are able to monitor their utterances with regard to ambiguities (for examples, see Eikmeyer & Ahlsén, 1998; Levelt, 1983). The maxim

---

[2] For an illustration of the terms *exophoric* and *endophoric* overspecification, take, for instance, the following fictive dialogue on a set of three objects, a RED PLASTIC CUP (Object 1), a RED PORCELAIN CUP (Object 2) and a YELLOW PLASTIC CUP (Object 3). Imagine speakers A and B having the task to pack these objects into a box for removal.

Speaker A: "Hand me the yellow plastic cup, please."

Speaker B: "No, we should take the red porcelain one first."

In the context of the two red cups, speaker A had to name the color only to refer unambiguously to object 3, i.e. he produces an *exophoric* overspecification of the material. Speaker B produces an overspecified utterance, too: Although it would be sufficient to name the material of object 2, being the only porcelain cup of the three objects, he names its color explicitly to contrast it with speaker A's description of object 3 (discourse related *endophoric* overspecification).

[3] My own findings support the assumption that referential overspecifications are rather perceptually determined than conceptually or linguistically planned. I will take up this issue in the discussion of the findings presented in this work.

of quantity, in contrast, is often neglected: Speakers tend to specify more than the minimally distinctive features of a referent (Dale & Reiter, 1995; Zhu 1995). Levelt (1989: 133) argues that Grice's maxims are formulated too generally to be tested experimentally, but that "there is more than one way to be cooperative in referring to objects. What violates the maxims from the exophoric point [object context] of view doesn't do so from the endophoric [discourse context]" (cf. Pechmann, 1984; see also Footnote 2).

Mangold-Allwinn et al. (1995) developed a model of the process of generating complex object specifications, taking into account several pragmatic determinants of object specifications, such as the communicative aim of speaker and hearer and the discourse context. They sketch the processing stages in object reference and make predictions as to the time course of the ongoing processes. However, although the model is grounded on rich empirical evidence for the existence of the assumed processing stages, there is little evidence regarding the time course of the processing within and between these stages. Note though that this reflects a basic methodological problem: It is almost impossible to assess the time course of cognitive processes and to maintain natural settings at the same time. In most experiments, the influence of situational factors and contextual determinants has to be eliminated for the sake of empirical validity (see Rohlfing, Belke, Rehm, & Goecke; submitted).

In the experiments presented in this work, the processes underlying object reference were studied by combining experimental means typically used to investigate speech production processes in naming with a referential communication task. For the aforementioned methodological reasons discourse-related factors and the aspects of non-verbal means of object reference could not be considered.[4] The present work thereby resumes a line of experimental research on the linguistic form of object specifications in referential communication tasks and – by working with measurements of reaction times and eye movements – extends it to the field of behavioral research on underlying processes and representations (cf. section 4 of this chapter).

---

[4] In the terminology of Brennan and Clark (1996), this is an "ahistorical" approach to object reference, compared to "historical" accounts, which incorporate discourse related contextual factors. Note, however, that in the setting of a naming experiment with measures of procedural variables (reaction times, fixation times), establishing a 'natural' discourse context would undermine the empirical control of the experimental situation (see Rohlfing, Belke, Rehm & Goecke, submitted).

2.2 Perceptual, Conceptual, and Linguistic Encoding in Referential Communication

In the following, I am going to present an outline of the processes underlying the generation of complex object specifications. As to include the broader framework of the present work, models of naming and speech production (Dell, 1986; Humphreys, Lamote & Lloyd-Jones, 1995; Levelt, 1989; Levelt, Roelofs & Meyer, 1999; Seymour, 1973) will be combined with models of object reference (Mangold-Allwinn et al., 1995; Herrmann & Grabowski, 1994). Only few models of naming incorporate the complete process from the perceptual analysis of a stimulus to the articulation of its name (e.g., Hoffmann & Kämpf, 1985). Most of them focus either on the first stages of conceptual and lexical processing (Schade & Eikmeyer, 1998; Herrmann, 1982; Herrmann & Grabowski, 1994; Humphreys et al., 1995; Humphreys, Riddoch, & Quinlan, 1988; Mangold-Allwin et al., 1995) or on the later stages of linguistic encoding (Dell, 1986; Dell, Chang & Griffin, 1999; Levelt, Roelofs & Meyer, 1999; Roelofs, 1997b; Schade, 1999).

During the last three decades, many different models of speech production have been developed (Fromkin, 1971; Garrett, 1975, 1988; Dell, 1986; Herrmann & Grabowski, 1994; Levelt, 1989; Schade, 1992, 1999). Three basic levels of processing are distinguished more or less explicitly in all these models, namely the levels of conceptual preparation, grammatical and lexical encoding, and articulation (Levelt, 1989). As indicated above, these models do not incorporate the representations underlying the processes of perceptual analysis and conceptual preparation. In general, all models of speech production must meet two fundamental assumptions on the underlying processing mode:

- Speech production proceeds from conceptual preparation via formulation to articulation.
- Speech production is an incremental process (Kempen & Hoenkamp, 1987). All components of the model can work in parallel: While later processing stages are still working on the first elements of an utterance, earlier stages of processing can already prepare later parts of the utterance.

There is an extensive controversy as to whether the cognitive processes underlying naming and speech production work in a modular, cascaded, or interactive way (Dell, 1985; Goodglass, 1998; Harley, 1993; Humphreys et al., 1988, 1995; Levelt et al., 1999; Schade, 1999). In the framework of the *modular* or *discrete stages* view, the processing components are assumed to work autonomously, and to minimally affect each other (informational encapsulation, cf. Levelt, 1989). With this assumption, the interaction between processing components is minimized. In *cascaded* processing, this strict informational encapsulation is given up, and early processing components are assumed to potentially influence later processing stages. Information is transmitted to subsequent levels as soon as the processing at given level starts (Blanken, 1998; Coltheart, Curtis, Atkins, & Haller, 1993; Humphreys et al., 1988, 1995; Peterson & Savoy, 1998). Note, however, that in cascade models, the informational exchange between processing components is unidirectional, i.e., there is no feedback of information from one component to its predecessor (Humphreys et al., 1988). In *interactive* models, such a feedback of information is possible (Dell, 1985, 1986; Humphreys et al., 1995): Here, a processing component can influence earlier components; in several models of lexical access, for instance, it is assumed that complex selection processes on the semantic and phonological levels of representation occur in a parallel and interactive way (Dell, 1985, 1986; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Harley, 1984; Schade, 1992, 1999).

In each class of models, additional assumptions have to be made to account for the fact that while language production is rapid and incremental, allowing for simultaneous processing on different levels, it is at the same time designed to produce utterances that are sequential and unfold over time. Modular models are inherently sequential, so here additional assumptions have to made as to how to accomplish the rapidity and simultaneity of processing during production. Cascaded and interactive activation models, by contrast, are inherently parallel; nevertheless they have to be able to produce sequential output in order to model the process of language production appropriately.

In the following, I am going to present current models of object reference and speech production and refine these models with regard to the special case of object naming in a referential communication task. As articulatory processes are of minor interest for the present work, only the processes underlying the perceptual analysis of the (visual) input, conceptual preparation, and formulation will be discussed. Mind, though, that the resulting schema of representations and processing stages underlying the generation of complex referential expressions, as depicted in Figure 1, is not intended to be a 'model' of the production process in the strict sense (see Schade, 1999, for a review of the principles of modelling cognitive processes). Its main purpose is to serve as a guideline to the considerations and experiments presented below, and to introduce the basic terminology used in the remainder of this work.

2.2.1 Representations and Processing Stages

Mangold-Allwinn et al. (1995: 223ff.) subdivide the processes involved in producing referential noun phrases in three stages, namely *perceptual analysis*, *concept generation*, and *activation of lexical representations*. Based on the terminology of Levelt (1989), I use the more general terms *conceptual preparation* and *formulation* for the latter two stages.[5] Levelt's "blueprint of the speaker" represents a comprehensive account of the processes underlying speech production (see Levelt, 1989, 1999). Yet, it does not capture the details of conceptual preparation in referential communication (see also Hirst, 1999). Humphreys et al. (1988, 1995) provide a detailed account of the processes mediating between the perceptual analysis of object features, the identification of the object, and the activation of its name. I integrated the ideas proposed by Mangold-Allwinn et al. (1995) and Humphreys et al. (1988, 1995) with Levelt's (1989) model of speaking, combining them to a general outline of the processing stages that can be assumed to be involved in the production of complex object specifications (see Figure 1).

[5] Beyond terminological issues, the models by Mangold-Allwinn et al. (1995) and Levelt (1989) differ in varios aspects, such as the underlying notion of concepts in the two models. For the present work, Levelt's model, being a procedural model of language production, will represent the narrower framework of the experimental investigation.

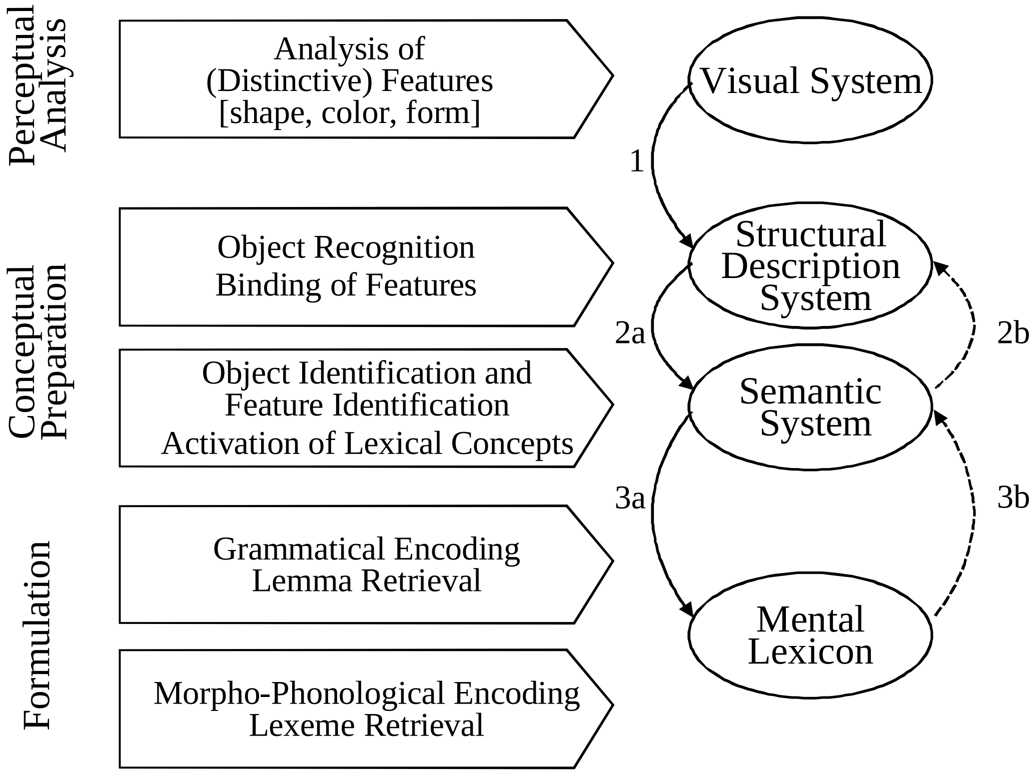


Figure 1. Processes and representational systems underlying the production of complex object specifications in a referential communication task (adapted from Humphreys et al. 1988, 1995; Mangold-Allwinn et al. 1995; Levelt, 1989, 1999).

2.2.2 Perceptual Analysis

As indicated in Figure 1, the perceptual analysis of the input is primarily aimed at identifying the distinctive features of the object to be named. Following Herrmann and Deutsch (1976), I consider the object class – here: the shape of the target object – as one of these features. There is a rich diversity of features that can be employed in referential communication tasks (Herrmann & Deutsch, 1976); in the framework of the present work, however, I will focus on color, size and object class, as these are the features most commonly used in previous experiments (see section 3 of this chapter for a review). No detailed analysis of the visual processes underlying the perception of color, size and shape features of objects (see, e.g., Marr, 1982) is given at this stage, as for the present purposes it is sufficient to know how *differences* in these features are detected, and to what extent procedural aspects of the detection process are relevant for later processing stages. There is extensive literature on the detection of multidimensional differences, most of which is based on the experimental paradigm of “same”-“different” judgments (see Farell, 1985, for a review). However, in most of the experiments of this paradigm artificial stimuli, such as geometrical shapes or non-objects, were used, which are probably processed differently than outline drawings of objects (Boucart & Humphreys, 1992; Grill, 1971; see also chapter 3). Therefore, I explicitly assessed the processes underlying multidimensional stimulus discrimination in a “same”-“different” decision experiment, using the same objects as in a later experiment on referential communication (Experiments 1 and 2, cf. chapters 3 and 4). For reviews of previous theoretical and empirical research on the processes underlying two-dimensional shape recognition and the visual perception of color and size see, e.g., Logothetis & Sheinberg (1996) or Quinlan (1991).

2.2.3 Conceptual Preparation

In Levelt’s (1989, 1999) model of the speech production process, the stage of conceptual preparation comprises several planning processes that are broadly subdivided into macro- and microplanning. I will not present those aspects of the macroplanning processes that are related to situational and discourse-related factors, as they are not in the focus of the present investigation. Yet, in the framework of the present work, an important aspect of macroplanning is the selection of information for making reference to objects (cf. Levelt, 1989: 129ff.).[6] During microplanning a propositional form of the selected information is generated and lexical concepts are retrieved (Levelt, 1989).[7,8]

---

[6] Another aspect subsumed under the macroplanning processes is the ordering of the information selected for expression. These ordering processes are particularly important for the generation of complex syntactic combinations of several referential expressions (Ferreira & Henderson, 1998; Levelt, 1981, 1982). For the case of referential noun phrases, however, only one referential expression is encoded and the ordering of words in the phrase is accomplished on the level of grammatical encoding (Schriefers et al., 1999; Schriefers & Teruel, 2000), rather than on the level of conceptual preparation.

[7] Note that the assumption of lexical concepts as terminal elements of conceptual preparation is highly controversial: Other than decomposed representations on the basis of combinations of primitive concepts, lexical concepts are non-decomposed representations of semantically complex words that are linked to the primitive conceptual features they are composed of (functionally decomposed representations; see Roelofs, 1992, 1997a for a theoretical and empirical review of this controversy).

As indicated above, Levelt's model does not capture the details of the conceptual preparation processes for the case of referential communication. Humphreys et al. (1988, 1995) provide a detailed account of the encoding processes and the representational formats mediating between the perceptual analysis of object features and the microplanning processes associated with the activation of lexical concepts. There is extensive evidence for the distinction between a *Structural Description System* and a *Semantic System* from neuropsychological, developmental, and behavioral studies (Coltheart, Inglis, Cupples, Michie, Bates, & Budd, 1998; Flores d'Arcais & Schreuder, 1987; Humphreys et al., 1988; Humphreys & Riddoch, 1999; Kelter, Grötzbach, Freiheit, Höhle, Wutzig, & Diesch, 1984; Sartori & Job, 1988). In picture naming, access to object names seems to proceed mainly via semantic representations. There is little evidence that phonological encoding processes in picture naming can be initiated without semantic mediation (Brennen, David, Fluchaire, & Pellat, 1996; Goodglass, 1998; Warrington, 1975), as it has been shown for word reading (Coltheart et al., 1993; Jacobs & Grainger, 1994; Plaut, McClelland, Seidenberg, & Patterson, 1996).

For the purpose of the present work, I assume that not only the Semantic System but also the Structural Description System represent conceptual knowledge and are therefore involved in the conceptual preparation processes (Coltheart et al., 1998; Flores d'Arcais & Schreuder, 1987; Humphreys & Riddoch, 1988; Klix & Metzler, 1982; Shallice, 1988).[9]

In particular, it can be argued that the assumption of both lexical concepts and lemmas is redundant (see, e.g., Harley, 1999; Zorzi & Vigliocco, 1999).

[8] There may be language specific differences as to how and which information can be encoded in lexical concepts. Stachowiak (1978), for instance, distinguishes between referring via naming or "labeling" vs. describing a referent. He argues that in spite of this formal distinction both types of reference can be represented in terms of logical predicates. This is obvious for the case of descriptions; for labels he suggests a predication like APPLIES.

[9] Humphreys et al. (1988) leave open whether both structural descriptions and semantic knowledge are represented in a prepositional format as part of a common conceptual system. Coltheart et al. (1998) assume that the semantic system consists of several subsystems, namely a "non-perceptual" subsystem, that is independent of sensory modalities, and modality specific subsystems representing "perceptual-attribute" knowledge (see also McCarthy & Warrington, 1994). Among these, the subsystem for the visual modality may correspond to the structural description system proposed by Humphreys et al. (1988, 1995).

2.2.3.1 Structural Descriptions

In the Structural Description System (see Figure 1), representations of object forms and structures are stored. A common diagnostic tool to tap the access to structural descriptions are object decision tasks, where participants are asked to judge whether objects are real or not (cf. Kroll & Potter, 1984). Non-objects may be created from real objects in such a way that object decisions require access to stored perceptual knowledge in any case (e.g., Humphreys et al., 1988).

Structural similarity is generally defined on the basis of perceptual features.[10] In the framework of the present work, I assume that color and relative size of a target object in a referential communication task are coded as part of its structural description. This process of "binding" features is often assumed to be accomplished by means of a master map of spatial locations (Treisman, 1988). Separately extracted features, such as color or orientation, are integrated via their common position at a particular location within this master map. However, more recent findings challenge this idea of spatial location as unique determinant of feature binding. They suggest an "account of object perception as the process of setting up and utilizing temporary 'episodic' representations of real world objects" (*object files* or *object tokens*; cf. Kahnemann, Treisman, & Gibbs, 1992: 177; see also Gordon & Irwin, 1996, 2000; Prinzmetal, 1981; Yantis, 1992). Kanwisher and Driver (1992) give a comprehensive overview of the evidence on the role of object tokens in directing visual attention and feature binding (see also Driver, 2001, for a review of space-based vs. object-based accounts of selective visual attention). I will not discuss this issue in more detail here, as – for the present purpose – the main claim I want to make is that there are ways to link an object and its features to a complex structural description, and that this linking or feature binding takes place during the perceptual encoding of the target object in a referential communication task.[11] The issue whether context objects

---

[10] In the interactive activation connectionist (IAC) model of Humphreys et al. (1995), structural similarity is modeled by means of excitatory connections between similar, i.e. consistent, representations and inhibitory connections between inconsistent representations.

[11] As described in section 1 of this chapter, highly salient features of an object are often over-specified, maybe to make it easier for the listener to identify the intended object. Weiß and Mangold (1997) present findings from a referential communication task that show that the color of an object is often not (over)specified when it is characteristic of the target object (e.g. *yellow* – banana). This result suggests that in some objects the color is an integral part of long-term memory structural

are processed similarly has to be left open. Schriefers (1990) provides evidence for the encoding of both a target and a context object in naming features of a target object. Participants were asked to judge the relative size or length of a target object in relation to a context object. The objects were presented simultaneously, and the target object was marked by means of a cross. Schriefers found that the overall size/length of the presented objects influenced the conceptual processing of the relative size/length of the target object. For the case of size (Experiment 5), the judgment "smaller" was given faster when both objects were small, than when both were large. Similarly, the answer "larger" was given faster when the overall size of both objects was large. Schriefers interprets this result in terms of the congruency of absolute information from the overall size of both objects and relative information from the individual size of the target object. The facilitating effect of the congruency of absolute and relative information suggests that both objects are processed conceptually at least on the level of structural descriptions.[12] However, considering referential communication tasks with complex object displays including more than two objects (see, e.g., Pechmann & Zerbst, 1990), it might be more plausible to assume that the context objects only serve as a foil for the analysis of distinctive features in the target object. They would thus play a role in the perceptual analysis of the display, but not in the later processing stages (as in Schade & Eikmeyer, 1998; see also Eikmeyer, Schade, Kupietz, & Laubenstein, 1999).

2.2.3.2 Semantic Representations

Semantic representations specify conceptual, associative, and/or functional knowledge about objects. They are stored in the Semantic System. Evidence on the internal structure of the Semantic System comes from neuropsychological studies on

---

descriptions of the object and is probably stored in a bundle of features associated to that object. Objects that do not have a characteristic color per se will probably be stored without any color information. When such an object is presented as colored drawing, the representations of the object features and its color are presumably activated and represented as separate units of processing that have to be bound to form a whole (see above).

[12] In another experimental condition, the cross that marked the target object was presented 1500 ms after the objects had appeared on the screen and the main effect of consistency vanished for reaction times. Schriefers (1990) argues that in this POST-condition subjects had had enough time to reject "absolute size as inadequate information" so that it did not interact with the naming of the relative size any more (ibd.: 130).

disorders of semantic memory in patients with brain lesions or Alzheimer's disease (Chertkow, Bub, & Caplan, 1992; Coltheart et al., 1998; Garrard, Perry, & Hodges, 1997; McCarthy, & Warrington, 1994)[13]. Such disorders may indicate deficits to access or retrieve information; in Alzheimer's disease, however, the disorder may also be due to a breakdown within semantic memory itself (cf. Garrard et al., 1997; Harley, 1998). Disorders of semantic memory after focal cerebral lesions often have been reported to selectively affect specific categories, such as animate vs. inanimate things (see Caramazza, 1998, and Saffran & Schwartz, 1994, for reviews). Consequently, the diagnostic tools to assess semantic disorders are designed to test categorical knowledge. The tests usually cover the whole range of input and output channels to assess potential modality specific impairments, too (see Garrard et al., 1997, for a list of examples).

In the framework of category specific impairments, selectively preserved color naming has been reported before (Mummery, Patterson, Hodges, & Price, 1998; Robinson & Cipolotti, 2001), indicating that concepts for colors and natural kinds are represented in dissociated cortical areas (see also Damasio, McKee, & Damasio, 1979 on cases of selectively *impaired* color naming). Several positron emission tomography studies provide additional evidence for distinct neural correlates for processing colors and color names (Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995; Price, Moore, Humphreys, Frackowiak, & Friston, 1996).

Additional specializations can be found within the semantic network for objects (e.g., in terms of animacy; see also Rosch, 1973, 1975a; Loftus, 1975). In the framework of the present investigation, however, this will be of minor importance, as the colors and objects used in the experiments presented below are only prototypical instances of objects and their attributes (cf. Rosch, 1973, 1975a, 1975b; Rosch & Mervis, 1975). Therefore, I assume that in these experiments the differences between the conceptual and semantic representations of objects and attributes will not affect the retrieval of lexical concepts during conceptualization (Levelt, 1989; Roelofs, 1992).

[13] As indicated above, I assume that the modality specific visual subsystem, that Coltheart et al. propose to represent perceptual attribute knowledge, corresponds to the structural description system proposed by Humphreys et al. (1988), while the "non-perceptual system" is comparable to the "Semantic System" proposed here.

Lexical concepts constitute the input to the formulation stage that will be depicted in the following.

2.2.4 Formulation

During formulation, the selected lexical concepts are given syntactic and morpho-phonological shapes. It is commonly assumed that these encoding processes are linked to and have access to distinct lexical representations, referred to as *lemmas* and *lexemes* (Bock & Levelt, 1994; Levelt et al., 1999). While lemmas comprise (semanto-)syntactic lexical information, lexemes are lexical representations of phonological forms.[14] In the following, I will give a brief outline of the formulation processes in complex noun phrase production.

2.2.4.1 Grammatical Encoding

Complex noun phrases, as well as other syntactic structures, are commonly viewed as syntactic frames with slots or rules for the ordered insertion of lexical items with certain syntactic specifications (Dell, 1986; Schade, 1999; Schade & Eikmeyer, 1998). Bock and Levelt (1994) assume four main processes in grammatical encoding (see also Garrett, 1988):

- During lexical selection, different classes of lemmas (e.g., adjectives vs. nouns) are accessed depending on the respective slots in the noun phrase. As indicated above, lemmas carry the grammatical information associated with individual lexical concepts.
- During function assignment, each lemma is assigned a syntactic role and the respective grammatical information is accessed. In a complex noun phrase like *das rote Auto* (the red car), an adjective (*rot*) is assigned the syntactic role of an attribute of the head noun *Auto*. Accordingly, the relevant grammatical information is retrieved: In the present example, the lemmata of the correct inflectional ending -*e* for nominative, singular, neuter adjectives and the lemma

[14] Note that in the original terminology as introduced by Kempen and Huijbers (1983) the lemma comprises not only syntactic, but also semantic information. As outlined above, Levelt et al. (1999) represent this semantic information in the form of lexical concepts, defining a “new” notion of a lemma as lexical representation of syntactic information (see Levelt et al., 1999; Zorzi & Vigliocco, 1999).

of the correct definite determiner *das* for nominative, singular, neuter nouns would have to be retrieved for the definite noun phrase.

- During constituent assembly, a “control hierarchy for phrasal constituents” is set up (ibd.: 947). In the present example of a German definite noun phrase, this hierarchy reflects the syntactic dependencies between head noun and determiner and adjective(s) in the noun phrase and controls the correct sequentialization of the elements in the noun phrase (see Schade, 1992, 1999, and Eikmeyer & Schade, 1991 for a connectionist account of such sequentialization processes).
- During inflection the lemmas of the determiner of the head noun and the inflectional endings of the adjectives are inserted (see Bock & Levelt, 1994; Lapointe & Dell, 1989, for detailed accounts).

The assumption of these four stages and their functional and positional purposes is primarily motivated by analyses of speech errors. The actual process of grammatical encoding has to be conceived of as being highly incremental, allowing for simultaneous processing of different pieces of information on different processing levels (see Levelt et al., 1999; Schade, 1999; Schade & Eikmeyer, 1998). Individual stages of processing are difficult to isolate since they are closely time-locked. Along with encoding processes in one domain, activation is built up in other domains: During lemma access, for instance, the retrieval of the head noun of a noun phrase will activate corresponding function words going along with the word class NOUN, such as lemmas for the definite and indefinite article.

Following the procedural localist connectionist model of noun phrase production by Schade and Eikmeyer (1998), the processes underlying grammatical encoding of a complex object specification, such as *das rote Auto* (see above) can be outlined as follows: Based on representations built up on the conceptual level (→ object space) and the level of structural descriptions (→ feature space), the lemma node of *Auto* (car) and the feature nodes *rot* (red) and *definite* will be activated in the target space. In addition, a network of control nodes will be activated in the control space, corresponding to the structural frame of a complex noun phrase (DET – ADJ(COL) – NOUN). The lemma of *Auto* (car) passes its activation on to the morphological stem of *Auto* and the corresponding category nodes for its gender (neuter), number (singular) and case (nominative). Likewise, the lemma *rot* activates the morphological stem of the adjective *rot*. *Definite* activates all forms of the definite article (*der*, *die*, *das*). Similarly, each of the grammatical category nodes passes its activation on to grammatical morphemes that fit its description. In the end, the article *der* and the inflectional ending *–e* should be most activated and will thus be selected. The network of control nodes accomplishes the correct sequential selection of word and morpheme forms (lexemes) according to the structural frame, namely the definite article, the adjective and its inflectional suffix and the noun, yielding the nun phrase *das rot-e Auto*.

Except for lexical access processes, the time-course of grammatical encoding processes in complex noun phrases has not been well explored yet. Schriefers et al. (1999) assessed the relative durations of lemma retrieval processes for nouns vs. color adjectives in complex noun phrase production. Depending on the number of nouns in the response set, they obtained prolonged lemma retrieval times for nouns, but not for adjectives. They conclude that their participants operated with “two separate ‘subvocabularies’ or response sets, one containing the adjectives and one containing the nouns” (ibd.: 717; see also Eikmeyer et al., 1999; Schade, 1999 for a corresponding connectionist implementation).

Beyond these syntactic aspects, complex inflectional rules have to be met for the production of a definite noun phrase (see above): In German noun phrases the definite determiner and adjective attributes are inflected according to the gender of the head noun. This implies that although the head noun is the last element in a complex referential noun phrase, its gender has to be accessed early during the formulation process to select the correct syntactic gender for the determiner (see Miozzo & Caramazza, 1999; Schriefers & Teruel, 2000; but see Schade & Eikmeyer, 1998 for an alternative account)[15].

---

[15] Caramazza and colleagues assume that languages can be classified into early- vs. late-selection languages according to the timing of the selection of the determiner. The timing of these selection processes was tapped by using the picture-word interference paradigm, using gender-incongruent distracters. For German noun phrases, Schriefers & Teruel (2000) found an early gender interference effect suggesting that German is an early-selection language. Using plural noun phrases that are inflected identically across genders, however, Schriefers (2000) did not obtain a gender interference effect, which suggests that there may exist an intralanguage variability (see also LaHeij, Mak, Sander, & Willeboordse, 1998).

2.2.4.2 Phonological Encoding

During the phonological encoding of a German complex noun phrase, the lexemes of the noun and the adjective stems are retrieved. In addition, the phonological form of the definite determiner (*der*, *die* or *das*) and of the inflectional endings of the adjectives are retrieved (-*e* for all singular, nominative, definite determiner NPs). I will not go into detail on the selection and sequential ordering of phonemes, as this is of minor importance for the present study (cf. Dell, 1986; Levelt et al. 1999; Schade, 1999 for comprehensive accounts).

2.2.5 Processing Modes in Conceptual Preparation and Formulation

The timing of perceptual, semantic, grammatical and phonological encoding is still a matter of substantial controversy on theoretical and empirical grounds. In the following, I will briefly outline the modes of informational exchange between processing stages (cf. Figure 1).

As outlined above, speech production is commonly agreed to be an incremental process (Kempen & Hoenkamp, 1987), but there is considerable disagreement about whether a given part of an utterance is processed in *discrete stages*, in a *cascaded* or an *interactive* processing mode. In the following, I want to present an argument for cascaded and partially interactive processing in the model presented in Figure 1.

Speakers often begin to speak before they have completed the visual exploration of the whole set of context objects (Pechmann, 1989). The occurrence of overt repairs and postnominal adjectival attributes (such as “the yellow shirt, the big one”; cf. Eikmeyer & Ahlsén, 1998; Eikmeyer et al., 1999; Schade & Eikmeyer, 1998) suggests that during the linguistic encoding of the target object, the visual exploration continues and the perceived information is continuously transmitted (“cascaded”) to the conceptual and linguistic encoding stages (see Arrow [1] in Figure 1).

With regard to the processes on the level of structural descriptions and semantic representations, Humphreys et al. (1988) argued for a cascaded processing mode (see Arrow [2a] in Figure 1; cf. Humphreys et al., 1995, for a short review of the main results in support of this view). In a picture naming experiment, Humphreys et al. (1988) found stronger frequency effects for structurally dissimilar objects, than for structurally similar objects. They concluded that the mode of informational exchange between semantic and phonological encoding processes is cascaded, too (see Arrow [3a]). From a representational point of view, recurrent connections from the semantic system to the structural description system are indispensible for tasks like answering questions about viusal attributes of objects or drawing to dictation (Arrow [2b] in Figure 1; see Coltheart et al., 1998). From a procedural viewpoint, however, it is not clear yet, whether the informational exchange between the two representation systems is bi-directional (see Chertkow et al., 1992, Sartori & Job, 1998, for discussions). In the local connectionist implementation of their model, Humphreys et al. (1995) therefore used bi-directional links between the three levels of representation (ibd.: 557; see also Humphreys, Riddoch & Price, 1997; Humphreys, Price & Riddoch, 1999).

The assumption of interactive processing between the levels of semantic and lexical encoding would be in line with the general view that lexical concepts and lemmas interact bi-directionally (Schriefers, 1990; Bock & Levelt, 1994; Levelt et al., 1999)[16]. In view of the time course of semantic and phonological encoding processes in lexical access, however, the cascade processing view as opposed to the assumption of discrete stages or the interactive processing view is highly controversial (Arrow [3b] in Figure 1; see Damian & Martin, 1999, Dell et al., 1997; Dell, Chang, & Griffin, 1999; Peterson & Savoy, 1998; Levelt et al., 1999; see also Schriefers et al., 1999).

As the empirical status of recurrent connections in the later processing stages is still open, I largely adopted the cascade processing view for the model depicted in Figure 1 and included recurrent connections (dashed arrows) only tentatively,.

[16] In particular, this allows for the possibility that speech production and comprehension share the same lexicon.

2.3 Open Issues in Referential Communication

As indicated at the beginning, previous research on referential communication has mainly focused on the information on the object, that is conveyed in a noun phrase. In several Indoeuropean languages with prenominal adjectives (German, English, Dutch, Swedish), analyses of the *form* of the noun phrases and the order of mention of the properties included in the object specifications revealed a high frequency of *overspecifications of color* and a *canonical order effect* for color and size adjectives: A substantial proportion of utterances included color specifications, although color was not a minimally distinctive feature (Eikmeyer & Ahlsén, 1998; Pechmann, 1984, 1989, 1994), and in about 85 % of all complex noun phrases including a size and a color specification, size was named before color (Danks & Schwenk, 1972; Eikmeyer & Ahlsén, 1998; Ford & Olson, 1975; Herrmann & Deutsch, 1976; Martin, 1969a, 1969b; Olson, 1970; Pechmann, 1984, 1989). In the following sections, these phenomena and their theoretical implications will be reviewed in more detail.

2.3.1 Color Overspecifications

Referential overspecifications include more than the minimal set of distinctive features (Ford & Olson, 1975; Garmiza & Anisfeld, 1976; Herrmann & Deutsch, 1976; Deutsch & Pechmann, 1982; Whitehurst, 1976). In experiments incorporating color and size as differential dimensions, most of the redundant utterances include overspecifications of color (Eikmeyer & Ahlsén, 1998; Herrmann & Deutsch, 1976; Pechmann, 1989; Schriefers & Pechmann, 1988). This phenomenon has generally been viewed as a result of the high perceptual salience of the color dimension. Two more specific accounts of referential overspecifications have been developed that will be presented in the following.

2.3.1.1 Incrementality

Pechmann (1984, 1989, 1994) and Schriefers and Pechmann (1988) interpret over-specifications of color as evidence for an incremental processing mode between the stages of conceptualization and grammatical encoding: “The speaker could start his description even before the visual scanning and the conceptual planning for his description is completed, i.e. before he has identified the features that discriminate the target object from the context objects” (cf. Schriefers & Pechmann, 1988: 174). In a study on the viewing behavior of participants during a referential communication task, Pechmann (1989) observed that, indeed, participants started to articulate their utterances before they had seen all objects. In a referential communication study in German and Swedish, Eikmeyer and Áhlsen (1998) registered a considerable proportion of utterances where speakers specified the color before, but the size after the noun, saying, for instance, *das gelbe Hemd, das große* (‘the yellow shirt, the big one’; see above). Similarly, they observed overt repairs, such as *das gelbe, das GROSSE gelbe Hemd* (‘the yellow, the BIG yellow shirt’). Their findings nicely demonstrate that the linguistic planning processes seem to be continuously monitored, such that incoming perceptual information can effectuate modifications of the utterance in preparation. The incremental character of the production of object specifications thus seems to be an elegant account of referential overspecifications. Note, however, that this account cannot explain the strong syntactic preference to name the size before the color. It would rather predict the non-canonical order to occur more often. This inconsistency will be addressed in more detail in section 2.3.2.2.

2.3.1.2 Economy

Whitehurst (1976) argued that color overspecifications occur because speakers apply a principle of least effort on the evaluation of the detected differences (see also Pechmann, 1994; Pechmann & Zerbst, 1994). It may cost less effort to specify any detected feature of the referent than to explicitly assess the distinctiveness of each of these features, or – in other words: “While contrastive descriptions are efficient in terms of words they may be inefficient in terms of the effort to appropriately analyze the stimulus array.” (Whitehurst, 1976: 478). In the framework of the present investigations, I assumed that the analysis of distinctive features in a referential communication task can be reduced to multiple “same”-“different” decisions. In Experiment 1 (see chapter 3), I used the “same”-“different” paradigm to assess the processes underlying the detection of differences in color, size, or object class in multidimensional stimulus discrimination. Combining the stimuli used in Experiment 1 with a referential communication task, I then assessed perceptual determinants of referential overspecifications (Experiment 2). The results will be discussed in chapter 4.

2.3.2 Canonical Order of Prenominal Adjectives

The canonical order of color and size adjectives is surprising in view of incremental theories of speech production: Although color is the more salient feature and is thus available for the verbalization process much earlier than size, size is often named before color in complex noun phrases. This holds for several languages with prenominal adjectives, such as German, Dutch, English or Swedish (Eikmeyer & Ahlsén, 1998; Martin & Molfese, 1972; Deutsch & Pechmann, 1982). It is found not only in naming tasks, but also in acceptability ratings (Danks & Glucksberg, 1971; Martin, 1969b). Until now, two pathways have been followed to account for this canonical order effect, which will be presented in more detail in the following.

2.3.2.1 Ordering Rules

In the 1970ies, a rather descriptive *visuo-semantic approach* to the internal structure of noun phrases was developed on the basis of qualitative analyses of utterances and of the attributes verbalized in the object specifications (Byrne, 1979; Ertel, 1971; Hetzron, 1978; Martin, 1969b; see also Sichelschmidt, 1989, for a review). The dimensions were classified according to their perceptual and semantic properties, such as intrinsicality, absoluteness, or definiteness of denotation. The definitions of these terms are often rather vague and their denotation can be mediated at best by means of examples:

- Intrinsicality (Byrne, 1979: 73): “For example, […] of the ‘simple’ modifiers […], the ones that refer to kinds or species refer to the most intrinsic properties of objects, more so than adjectives which qualify in terms of colors or shapes”. In this sense, attributes, such as “wooden”, that depict the material an object is made of are more intrinsic than color or size attributes.
- Definiteness of denotation (Martin, 1969b: 472): “Adjectives which denote the same property regardless of the meaning of the modified noun are […] more definite in denotation than adjectives which denote different properties in the context of different nouns.”

- Absoluteness (Martin, 1969a: 700) is “a logical correlate of the definiteness of denotation”. Martin (ibd.: 702) gives the following operational definition: “Adjective absoluteness was explicitly defined in terms of the relative number of comparisons required for the choice of a given adjective”. Note, though, that absoluteness, in the strict sense of the term, refers to dimensions that are inherent to an object and do not require any comparison to other objects for a correct identification.

Generally speaking, the more absolute, intrinsic, and definite a dimension is, the closer it will be placed to the noun.[17] This holds for both, prenominal and post-nominal ordering, e.g. in Spanish (see also Greenberg, 1963).[18]

Based on such descriptive analyses, many grammars of the respective languages incorporate these findings in the form of ordering rules in idiomatic style (cf. Bache, 2000; Eichinger, 1991; Hetzron, 1978; Seiler, 1978; Zifonun, Hoffmann & Strecker, 1997). Note, however, that inverted adjective order phrases are not ungrammatical and – as indicated above – do occur to a certain extent in referential object descriptions (Greenberg, 1963; Hetzron, 1978; Pechmann, 1994; Pechmann & Zerbst, 1995; Teyssier, 1968): “[…] the speaker will not hesitate to ‘violate’ a convention which, in many cases, is proclaimed as a grammatical rule only because it reflects preponderant usage (Hörmann, 1981: 266 as quoted in Sichelschmidt, 1986: 146).

---

[17] Danks & Schwenk (1972; see also Danks & Glucksberg, 1971) formulated a “pragmatic communication rule” to account for ordering preferences in prenominal adjective order. They claimed that ordering effects should follow a principle of relative relevance of the adjectival predications and thus predict high frequencies of inverted adjective orders in specific situational contexts. Although they present empirical findings that seem to verify these predictions, Richards (1975) re-analyzed these findings and showed that there is no conclusive evidence of the pragmatic communication rule. In addition, data from her own experiments clearly contradict the pragmatic communication rule (Richards, 1975, 1977).

[18] Comparative investigations even suggest that “the adjective ordering principle based on the semantic classification of the qualifications is a universal, because it is attested in such a distribution in several languages that no mutual influence is to be suspected. Yet it is not a strong universal that must manifest itself everywhere” (Hetzron, 1978: 175; see also Foorman, 1983; Foorman & Kinoshita, 1983; Sobin, 1984).

Developmental studies show that children acquire knowledge on the meaning of adjectives separately from knowledge about ordering rules (Foorman, 1983; Richards, 1979). Kemmerer (2000) showed that, accordingly, semantic knowledge of adjective meanings and knowledge of adjective ordering rules can be selectively impaired in aphasic speakers. On the whole, these results suggest that speakers have knowledge of the internal semantic structure of complex noun phrases. It still remains unresolved, though, how this knowledge is represented (see below) and when and how it enters the production process.

2.3.2.2 Procedural Principles

Pechmann (1994) argues that assuming a system of learned rules for the sequence of all classes of adjectives would imply that we were to store a vast number of possible combinations. This would be incompatible with the idea of grammar as a set of abstract rules. Pechmann investigated the canonical order effect via a *procedural approach* to noun phrase production (see also Pechmann, 1989; Pechmann & Zerbst, 1995; Schriefers & Pechmann, 1988). The occurrence of referential overspecifications suggests that the transfer of information from conceptualization to formulation is incremental. Early color information is transferred immediately to the formulator. Pechmann investigated whether whole noun phrases are the units of incremental production on the level of grammatical encoding, or whether their internal grammatical and phonological structure is built up incrementally. The canonical order of color and size adjectives suggests that the grammatical structure of the noun phrase as a whole is planned before phonological encoding processes are initiated.[19]

In one of his experiments, Pechmann (1994) registered more inverted adjective orders when more context objects were present in the display. He interpreted this finding as an effect of time pressure that was more prominent in more complex situations and assumed that these inverted adjective orders might be a result of incremental processing. Accordingly, he hypothesized that the reaction times

[19] As indicated in section 2.2.4.1 (see also footnote 15), German is an early-selection language, and for the correct selection of the determiner and the inflectional suffixes for the prenominal adjectives, the gender of the head noun has to be accessed. Thus, the phonological encoding of the first elements in the noun phrase can only start when the last element in the phrase has been retreived on the level of lexical access.

associated with non-canonical adjective order phrases, where color is named first, should be shorter than those associated with canonical adjective order phrases. However, the data revealed the opposite: Although size was named first in the canonical order phrases, the reaction times for these phrases were significantly shorter than for the non-canonical phrases. Pechmann concluded that the processing mode between the conceptual stage and later formulation stages was partly incremental or even strictly serial and that noun phrases were produced as units of incremental production on the level of grammatical encoding (cf. Pechmann, 1994; Pechmann & Zerbst, 1994). Pechmann and Zerbst (1995) report similar findings on the effects of canonical syntactic structure on the duration of grammatical encoding during the production of SVO- v. OVS-sentences. Both structures are acceptable and occur in spontaneous speech. OVS-sentences, however, are used in few pragmatically and contextually constrained situations only, i.e., they are *marked* and occur substantially less often than the *unmarked* SVO-sentences. Pechmann and Zerbst propose that processing times might be affected by the markedness of the syn-tactic structure of the utterance. They leave open, however, how the knowledge about the markedness of specific structures is represented in the speakers' minds.

2.3.2.3 Effects of Task Difficulty

As indicated above, Pechmann (1994) found that the number of context objects influenced the occurrence of non-canonical adjective orders and assumed that these inverted adjective orders might be a result of the incremental noun phrase production under conditions of increased time pressure. Pechmann and Zerbst (1990) reported parallel results in a referential communication task with variations of the number of context objects and of the detectability of color differences: A larger number of context objects and less detectable color differences led to a significant increase of inverted adjective orders. This seemed to contradict the hypothesis of the incrementality of noun phrase generation: Although color was less detectable, it was specified more often in initial position.

To conclude, a comprehensive account of the occurrence of inverted adjective orders and the discrepancy in processing times between canonical and inverted adjective order phrases should incorporate effects of task difficulty. The origin of inverted adjective orders might lie in a combination of the canonical order constraints and the incremental nature of the production process. In case of increased task difficulty, participants might not take into account any semanto-syntactic constraints but articulate the object's features in the order of detection.[20] In Experiment 3 (see chapter 5), this hypothesis was addressed by including a variation of task difficulty.

[20] In order to account for the finding that the number of inverted adjective orders increases when the color difference is less detectable, one would have to assume that speakers detect or attend to color differences first. Evidence on perceptual grouping and search strategies in visual search might support this view (cf. Carter, 1982; Cave, 1999; Cohen & Shoup, 1997; Feldman, 1999; Treisman, 1982; Treisman & Gelade, 1980).

2.4 Methodical Considerations

Previous studies on referential communication, which mainly focused on the degree of elaboration and the form of object specifications, were based on qualitative analyses of utterance structures. A first attempt to approach the processes underlying referential communication by means of procedural measures was made by Pechmann (1989, 1994). Reaction times, as they were measured by Pechmann, have proved as a useful tool to track the time course of generating one-word utterances (Levelt, Schriefers, Vorberg, Meyer, Pechmann, & Havinga, 1991; O'Seaghdha & Marin, 1997; Schriefers, Meyer, & Levelt, 1990). In the framework of the referential communication task, however, reaction times fail to capture the exact time-course of the processes underlying the generation of complex object specifications. Unlike simple naming tasks, the referential communication task is hard to control for the duration of perceptual processes preceding the linguistic processing of the stimulus; in fact, the perceptual processes are an integral part of the task and are thus of interest for procedural investigations, too (Sanders, 1993).[21] Tracking participants' eye movements in a referential communication task can draw a more complex picture of both perceptual and linguistic processes underlying the production of complex object specifications.

In the following, I will present some fundamentals on eye movements in information processing and attention, which are essential for understanding the rationale of using eye monitoring as an experimental technique. I will then give a brief outline of the use of eye tracking techniques in empirical research on language perception and production and discuss alternative methods.

---

[21] Pechmann (1994; see also Pechmann & Zerbst, 1994) adopted the following solution: To locate a difference in RTs that he had obtained in the naming latencies for CSO- and SCO noun phrases he conducted a series of experiments to test each processing stage separately. Note, however, that such a strategy implies considerable variation between tasks, which makes it difficult to compare the results on individual processing stages.

2.4.1 Eye Movements and Information Processing

When we process visual information during reading or scene perception we continually move our eyes. These movements are called *saccades*.[22] The stable states between two saccades are commonly referred to as *fixations*. Typical variables observed in empirical research on visual processing are fixation durations and locations and saccade lengths and locations or landing points (Liversedge & Findlay, 2000). Particularly saccade length and fixation duration have proved to be good parameters of task characteristics (cf. Rayner, 1984, 1998).

It is commonly assumed that during a saccade, visual information processing is reduced, because the eyes move that rapidly that nothing but a blur would be perceived (*saccadic suppression*; Matin, 1974). Nevertheless, we do not perceive the world in chunks of fixations, but as being stable (Carlson-Radvanski & Irwin, 1995; Irwin, 1991; McConkie & Currie, 1996). This seems to be in part due to backward and forward masking of information during saccadic movements (Brooks, Impelman, & Lum, 1981; see also Henderson, Pollatsek, & Rayner, 1987; Pollatsek, Rayner & Collins, 1984). In addition, there are short term memory representations retaining visual information across saccades (*transsacadic memory*; Carlson-Radvanski & Irwin, 1995). It is still unresolved whether – apart from saccadic suppression of *visual* processing – *cognitive* processing is suspended during saccades, too. For the time being, it seems plausible to assume that, particularly in higher order processes, some processing occurs during saccades (Henderson, Dixon, Petersen, Twilley, & Ferreira, 1995; Irwin, 1998; Matin, Shao, & Boff, 1993; Rayner, 1998).

With regard to eye movement analyses in higher order cognitive processes there is some controversy as to what might be the best measure of processing time (see Rayner, 1998, for a review). In research on object processing and naming, *gaze duration* has proved to be a useful tool to measure processing times. It includes the sum of the durations of all successive fixations on one object (Just & Carpenter, 1980; Henderson, Pollatesek & Rayner, 1987; 1989). Sums of gaze durations, e.g. when the eyes return to an object, are commonly referred to as *(total) viewing times*.

[22] Other types of movements are pursuit eye movements, which are characteristic of tracking targets in motion, and vestibular eye movements, occurring when head or body movements have to be corrected for to maintain a stable direction of vision (see Rayner, 1998).

2.4.2 Eye Movements and Visual Attention

Understanding the interrelations of eye movements and visual attention is an essential prerequisite for an appropriate interpretation of the data from eye movement experiments on cognitive processing. The crucial point is how attentional allocation and cognitive processes are associated with oculomotor behavior, such as fixations or saccades. The following principles have been proved to hold:

- Saccades are obligatorily coupled to shifts of attention (Deubel & Schneider, 1996; Remington, 1980); more specifically, attention precedes a saccade to a given location in space (Hoffmann & Subramaniam, 1995; Kowler, Anderson, Dosher, & Blaser, 1995).

The opposite, however, does not hold:

- Shifts of attention are not obligatorily coupled to saccades, i.e., we can shift the focus of attention without moving our eyes (*covert orienting*; Posner, 1980; see Liversedge & Findlay, 2000, for a recent review). From a physiological point of view, however, it should be noted that although covert allocation of attention is not necessarily coupled to explicit eye movements, there is a considerable anatomical overlap of the neural correlates of overt and covert shifts of attention (Corbetta & Shulman, 1998; Remington, 1980; see also Posner, 1992, for a concise overview).

Attentional capacity is commonly assumed to operate either in parallel and distributed over the complete visual field or selectively at a particular focus of attention (Duncan, 1980; Hoffmann, 1979; Treisman, 1977; Treisman & Gelade, 1980). Visual search processes, for instance, are often considered to proceed from a parallel (preattentive) scan to a focused (attentional) serial search for few selected likely targets (feature detection and integration; see, e.g., Cave & Wolfe, 1990; Duncan, 1980; Hoffman, 1979; Treisman & Gelade, 1980; Treisman & Souther, 1985)[23]. Binding features to objects and naming them is associated with selective allocation of attention to the referent. Nevertheless, regions in the surroundings of

---

[23] Note, however, that although these accounts, among them the feature integration theory (Treisman, 1988; Treisman & Gelade, 1980) was highly influential, the allocation of different attentional processes to distinct levels of processing has often been challenged by conflicting findings and alternative models (cf. Cave, 1999; Driver, 2001; Duncan & Humphreys, 1989; Mordkoff, Yantis, & Egeth, 1990; Wolfe, Cave, & Franzel, 1989).

the focus of attention can be previewed parafoveally (cf. Henderson, Pollatsek, & Rayner, 1987, 1989; Henderson, 1992a, 1992b; Remington, 1980). Note, however, that the allocation of attention to foveal and parafoveal areas seems to be largely dependent on the overall difficulty of the task. Selective attention is dependent on capacity limitations (Henderson & Ferreira, 1990; Rayner, 1986) and some authors argue that foveal visual processing requires huge attentional resources so that only little attention can be allocated to parafoveal visual input (Henderson & Ferreira, 1993). Meyer, van Elswijk, and Tily (2001) investigated parafoveal preview effects in a naming task. They found that when several objects were named in a sequence there could be some visual and conceptual processing of upcoming targets, but for the retrieval of lexical information the objects have to be fixated. In higher order cognitive processes, such as scene perception or picture naming, fixation or gaze durations are therefore often alleged to reflect foveal but not parafoveal processing demands (cf. Henderson & Ferreira, 1993; Liversedge & Findlay, 2000).

2.4.3 Eye Tracking in Research on Cognitive Processes

There is a long tradition of using eye-tracking techniques in cognitive science, ranging from investigations of attentional allocation in visual perception (see above) to experiments on higher-order cognitive processing such as scene perception, reading, language comprehension, and language production (cf. Rayner, 1992, 1998; Tanenhaus & Spivey-Knowlton, 1996; Meyer, Sleiderink, & Levelt, 1998; to give one example of each domain). Findings from the latter two areas will be presented and discussed in the following.

2.4.3.1 Eye Movements and Speech Perception

Along with the development of fine-grained eye tracking tools, eye movements were extensively investigated in reading research (see Rayner, 1998, for a review). In the last decade, however, eye tracking has also been increasingly used in research on spoken-language comprehension (see Tanenhaus & Spivey-Knowlton, 1996; Tanenhaus, Magnuson, Dahan, & Chambers, 2000). The paradigm has proved a use-ful tool to study syntactic ambiguity resolution (e.g. in garden path sentences; see Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995) and the influence of visual context (Eberhard et al., 1995; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995a, 1995b) and real-world knowledge (see Tanenhaus et al., 2000) on auditory comprehension. The processing of both written and auditory stimulus material has been found to work immediately, in an incremental and rapid manner (cf. Eberhard et al., 1995; Just & Carpenter, 1980; Spivey, Tyler, Eberhard, & Tanenhaus, 2001; Tanenhaus et al., 1995a, 1995b; see also Cooper, 1974, for a remarkably early work in this field).

The immediacy of syntactic and semantic processing has been extensively studied in an instruction paradigm that is similar to the referential communication task. Eberhard et al. (1995; see also Tanenhaus et al., 1995a, 1995b) investigated the time course of processing complex referential expressions in instructions such as 'Touch the starred yellow square'. Depending on the display, the referent could be uniquely identified after having heard the first, middle or last element of the noun phrase.[24] As expected, the later the disambiguating information was given, the longer it took participants to detect (here: fixate) the referent,. When the mean eye movement latency was measured from the onset of the disambiguating word instead of the onset of the utterance, the latencies in the late condition were considerably faster compared to the other two conditions, suggesting that "as the noun phrase unfolded, the information from each word was used to reduce the candidate set of blocks to just the two potential referents, which were then distinguished by the last word of the noun phrase" (ibd.: 417; see also footnote 21). Sedivy, Tanenhaus, Chambers, and Carlson (1999) provide analogous evidence on the processing of scalar adjectives, such as "tall". Convergent findings were obtained in other studies on eye movements during auditory language comprehension, showing that linguistic and visual processses are closely time-locked and allow an immediate integration and disambiguation of information from either side (Spivey et al., 2001; Tanenhaus et al., 1995a, 1995b; see also Allopenna, Magnuson & Tanenhaus, 1998 on the time-course of lexical access processes). In the next section, I will present findings on the relation between viewing and naming objects, suggesting that here, too, lexical and visual processing is closely time-locked.

[24] In the first case, there was only one starred element in the display. In the second case, all elements of the display were starred, but only one was yellow. In the third case, two of the four elements were starred and yellow but only one of them was a square.

2.4.3.2 Eye Movements and Speech Production

Meyer, Sleiderink, & Levelt (1998) were the first to systematically investigate eye movements during object naming. They asked participants to name two objects presented side by side in a noun phrase conjunction, such as ‘scooter and hat’. On the basis of previous findings on scene perception and object recognition, they hypothesized that participants would fixate the objects they wanted to name in order to identify them and to find their names. The viewing times observed for the first object turned out to be synchronized with the time needed to find the phonological form of the first object name. Similar results were obtained in an experiment using complex noun phrases for the description of the first object (‘the big red scooter and the hat’; see Levelt & Meyer, 2000). In addition, phonological priming of the object to be named first proved to not only facilitate naming but to also diminish viewing times on the first object (Meyer & van der Meulen, 2000). On the whole, these results suggest that fixating an object until the phonological form of its name is retrieved is obviously *sufficient* to name it. As indicated above, Meyer et al. (2001) showed that “when several objects are named, peripheral objects may undergo some visual and conceptual processing, but lexical access to their names only begins after fixation”. Thus, fixating an object is apparently *necessary* to access linguistic information.

Applying the eye tracking technique to sentence production, Griffin and Bock (2000) investigated the time course of sentence formulation in describing simple events. They compared the eye movements observed under different task conditions (simple inspection, detection of the “victim” in each picture, and extemporaneous and prepared description of the event). The results suggest that speakers obviously first apprehend the event structure of the scene and identify agent and patient of the action. During the linguistic formulation process, the eye movements were closely linked to the order of mention, irrespective of variations in the material, e.g. in terms of orientation (agent left / agent right), or sentence structure (active / passive).

In a variant of the referential communication task, Eberhard (2000) had participants describe movements of objects on an array of 5x5 squares. In analogy to Griffin and Bock, she, too, presents examples of trials where participants first shortly preview the objects or locations involved in the movement. When formulating the movement description they do not necessarily need to refixate these objects and locations, as the results from Meyer and colleagues would suggest: When formulating a contrastive description, for instance, they rather tend to fixate the contextual alternative.

To conclude, eye tracking can be regarded as an important tool to gain insight to the processes preceding an overt reaction – not only in spoken language comprehension and reading but also in language production. As Meyer and colleagues have shown (Meyer et al., 1998, 2001; Levelt & Meyer, 2000; Meyer & van der Meulen, 2000), eye movements and fixations apparently reflect linguistic planning processes. Note, however, that the linkage between eye movements and linguistic planning processes may be task-dependent (Eberhard, 2000). The studies by Eberhard (2000) and Bock and Griffin (2000) suggest that also non-linguistic conceptual processes can be traced by means of eye movement analyses.

2.4.4 Other Research Tools

At the beginning of this chapter, the disadvantages of reaction times for procedural investigations of referential communication were briefly outlined and the benefits of eye tracking in measuring processing times prior to an overt responsewere described. Beyond the measurement of eye movements, there are several other methods to assess processes preceding the production of a (naming) response.

During the *decade of the brain* (former US-President George Bush, 1990; Presidential Proclamation 6158) the availability of brain-imaging techniques for studying cognitive processing has substantially changed empirical research in cognitive science. Earlier research on the neuroanatomical substrates of language and cognition had been restricted to neuropsychological studies of impaired performance. Brain-imaging techniques opened up new potentials for obtaining on-line data on *where* and *when* brain activity occurs in both normal and impaired cognitive processing.

There are two broad classes of functional brain imaging techniques.

- *Haemodynamic methods* measure changes in the regional cerebral blood flow (rCBF) in the brain. Increases in metabolism are generally considered to indicate an increased neural activity.[25] The haemodynamic methods commonly used are PET (positron emission tomography) and fMRI (functional magnetic resonance imaging).

- *Electrophysiological methods* allow direct measures of neural activity via electromagnetic fields, which can be detected non-invasively with electrodes mounted on specific regions of the head. Via an EEG (electroencephalogram), the electric component of the electromagnetic field can be recorded over time. Similarly, the magnetic component can be registered with an MEG (magnetoencephalogram). A frequent method to measure electromagnetic impulses in response to a stimulus or impulse are event related potentials or magnetic fields (ERPs and ERFs in the EEG and MEG, respectively).

I will not describe the rationale and the functionality of these methods in detail (see Rugg, 1999 for an overview). Generally speaking, haemodynamic methods have proved to have a high spatial but a poor temporal resolution. Electrophysiological methods, in contrast, have a poorer spatial but a good temporal resolution.

Brown and Hagoort (1999) and Gazzaniga (2000) provide comprehensive surveys of the use of brain-imaging techniques in cognitive science. In the field of language processing, functional brain imaging has been predominantly used in research on language *comprehension* (cf. Brown & Hagoort, 1999, for reviews). Language *production* is more difficult to investigate with brain imaging studies. EEG-recordings will only work in silent naming or other related tasks (e.g., van Tourennout, Haagort, & Brown, 1997, 1998), as the muscle activity associated with overt naming massively distorts the EEG-signal. Combining overt naming with MEG-recordings, in contrast, has proved to be feasible and useful in gaining fine-grained temporal and spatial information on processing stages during lexical access (Levelt, Praamstra, Meyer, Helenius, & Salmelin, 1998; Lounasmaa, Hämäläinen, Hari, & Salmelin, 1996; Salmelin, Hari, Lounasmaa, & Sams, 1994). Beyond these electrophysiological studies, PET has been used to identify the neural substrates of object recognition and naming (see, e.g., Price et al., 1996).

---

[25] Note, however, that "a change solely in the *timing* of the activity of a set of neurons (e.g. from asynchronous to synchronous firing) will have little or no haemodynamic counterpart, despite the likely significance of such a change" (Rugg, 1999: 19).

3 Determinants of Multidimensional Stimulus Discrimination

As indicated in section 2.3, I assumed that the analysis of distinctive features a in referential communication task can be reduced to multiple “same”-“different” decisions between two objects, the target object and each context object. In Experiment 1, I used the “same”-“different” decision task to assess the time course of detecting differences in color, size, and/or object class. In the present chapter, I will first introduce the “same”-“different” paradigm and discuss previous findings and their relevance the present investigation. The results from Experiment 1 will be discussed in terms of both their relation to previous research on ”same”-“different” judgments and their significance for the present work on referential communication.

3.1 “Same”-“Different” Judgments

Since the pioneering work of Egeth (1966) on the visual perception of multi-dimensional stimuli, considerable effort has been devoted to establishing the processes underlying multidimensional stimulus discrimination. Over the last 30 years, a multitude of studies especially on the “same”-“different” paradigm has accumulated (Allport, 1971; Bamber, 1969; Besner & Coltheart, 1976; Bindra, Donderi, & Nishisato, 1968; Donderi & Case, 1970; Donderi & Zelnicker, 1969; Egeth, 1966; Hawkins, 1969; Jolicoeur & Besner, 1987; Linsday & Lindsay, 1966; Miller, 1978, Miller & Bauer, 1981; Nickerson, 1969, 1971, 1972; Sekuler & Nash, 1972; Snodgrass & Townsend, 1980). In the visual “same”-“different” decision task, participants are asked to judge two stimuli as being same or different. The stimulus pairs can be presented either together or successively (*simultaneous* vs. *sequential* presentation). Depending on the task, participants are asked to base their decision either on all dimensions of the stimuli or only on some of them while disregarding others (*conjunctive* vs. *disjunctive* judgments, see Farell, 1985). Much of the previous experimental work in this area has been directed at testing models of the processes underlying “same”-“different” judgments for multidimensional stimuli (see also Farell, 1985, for a review of the extensive evidence dated from the early 70ies). Classes of models have been defined using two basic parameters as introduced by Egeth (1966): *processing time* (exhaustive vs. self-terminating) and *processing mode* (serial vs. parallel mode vs. template matching).

In the following, *same* and *different* are used to refer to the experimental conditions, i.e. the stimulus category of a given object pair. The notation “same” and “different” will be used for response types, i.e. for both potential response alternatives and actual responses made by subjects during the experiment (see Farell, 1985).

3.1.1 Methodical Problems

There is extensive evidence on the visual discrimination of multidimensional stimuli, but the findings seem to be rather inconsistent. In order to correctly evaluate them, several factors must be considered: Many “same”-“different” decision experiments were based on artificial stimuli constructed with regard to model-based predictions (Allport, 1971; Bamber, 1969; Brunel & Ninio, 1997; Egeth, 1966). These stimuli must be carefully distinguished from non-artificial stimuli. In the case of objects, as they will be used in the present series of experiments, there are semantic and linguistic associations connected to their form that can become activated even at a very early stage of visual processing (cf. Boucart & Humphreys, 1994). Therefore, transferring experimental results from one domain to the other may lead to a serious fallacy.

Beyond that, the “same”-“different” experiments conducted before have to be carefully inspected with regard to the mode of presentation (simultaneous vs. sequential) and the task involved (disjunctive vs. conjunctive judgments; see Farell, 1985).

3.1.2 Previous Findings

Although the reaction time data obtained in previous experiments could be used to draw inferences about the timing of the decision process, the results obtained under different task conditions differ widely, especially with regard to the decision latencies for the basic stimulus conditions *same* and *different* (cf. Farell, 1985; Grill, 1971 for critical reviews). Earlier findings provide evidence for a *self-terminating* search during the decision, i.e. an answer on a *different* trial should be possible, as soon as a difference in any dimension has been detected. “Same” decisions, on the contrary, would have to be made on the basis of an *exhaustive* search, considering all dimensions to be judged. Contrary to this prediction, “same” answers were often faster than “different” answers (cf. Bamber, 1969; Downing, 1970; Entus & Bindra, 1970; Farell, 1985; Grill, 1971; Hawkins, 1969; Nickerson, 1965, 1967). It turned out, however, that this so-called *fast-“same” phenomenon* occurred only in experiments with specific methodical features in the sense of the above classification in terms of task conditions and modes of presentation (cf. Farell, 1985). Further investigations revealed that for multidimensional conjunctive judgments, “ ‘same’ judgments are faster than the slowest class of ‘different’ judgments (for which a single dimension is critical)” (cf. Farell, 1985: 423). This pattern, in turn, could be explained on the basis of the inherent properties of the dimensions involved. The degree of *codability* of the dimensions appeared to be particularly influential (Bindra et al., 1968; Farell, 1985): According to a definition offered by Bindra et al., codability “refers to the property of a stimulus that enables most Ss to categorize it in absolute terms, without reference to another (e.g., standard) stimulus. […] By this definition, stimuli such as colors […] are codable, and stimuli such as […] line length are noncodable.” (ibd.: 129). Note that being codable or noncodable is an inherent property of a dimension, whereas the discriminability of difference conditions is part of the experimenter’s manipulation (cf. also Bindra et al., 1968). Nevertheless, codability is often confounded by discriminability: Size, for example, as – by definition – noncodable dimension, can be more or less discriminable. I will therefore consider the notion of codability not as a binary distinction between codable and noncodable dimensions, but as a continuum between highly codable classes of dimensions (absolute dimensions) and less codable classes of dimensions (relative dimensions).[26]

Taking into account all methodological differences between experiments, the results can be re-evaluated and reduced to a few basic hypotheses on the structural and temporal properties of the decision process in a conjunctive “same”-“different” judgment:

1. The search for differences in *different* object pairs is self-terminating, whereas the identity check on *same* object pairs is exhaustive. In conjunctive judgments, the *overall* reaction time to *same* stimuli includes a check of all dimensions and

[26] This definition is also in accordance with the definition of absoluteness, as provided in section 2.3.2.1 (see also Martin, 1969a,b).

is usually longer than the reaction time to *different* stimuli (Downing, 1970; Downing & Gossman, 1970; Egeth, 1966; Farell, 1985; Hawkins, 1969).

2. There is a codability effect for the detection of the difference dimensions in *different* stimuli (Bindra et al., 1968; Farell, 1985): Relative stimuli are processed substantially more slowly than absolute dimensions. The notion of codability can explain both the relation among processing times for *different* conditions and the relation between the *different* conditions and the *same* condition.
3. The dimensions to be judged in conjunctive search can be processed in parallel. Detection times for differences in less codable, relative dimensions are longer than reaction times to *same* stimuli, which suggests that the exhaustive check of all dimensions for giving a "same" answer is accomplished in a parallel way (see Farell, 1985).

3.2 Experiment 1

In Experiment 1, I wanted to replicate the findings listed above for the type of stimuli used in the present investigation. Other than in previous studies I used non-artificial stimuli in the form of line drawings of real objects. *Different* stimulus pairs varied in terms of object class, and/or color, and/or size. Participants were instructed to press the "different"-button, as soon as they detected a difference in one of the three dimensions, and to press the "same"-button if the two objects were identical with regard to all three dimensions. Using these dimensions, I was able to draw a within-subjects comparison of the processing differences between highly codable absolute features (object class and color) and less codable relative features (size), and between form (size, object class) and color features (cf. Garner & Felfoldy, 1970, Santee & Egeth, 1980; Watanabe, 1988a). I chose a low ratio for the size dimension (5:4). If the ratio had been very high (say 50 : 1), the small and large stimuli can easily be identified on the basis of their absolute sizes, and size ceases to be a relative dimension. Bundesen and Larsen (1975) and Larsen and Bundesen (1978) showed that the detection times for size differences got shorter with increasing size ratios. Such correlations of codability and discriminability are ubiquitous and have to be taken into account when effects of codability are considered: Choosing more discriminable size ratios makes the size dimension more codable, if not even absolute.

3.2.1 Predictions

By means of the experimental setting outlined above, I was able to analyze effects of the basic response types “same” and “different” on the structure of eye movement patterns and the duration of processing and fixation times. Similarly, I wanted to explore the effects of the numbers and types of differences, as these were of particular interest in view of the present investigation. I predicted that structural differences between viewing patterns should be positively correlated with quantitative differences in processing times. On the basis of the general hypotheses stated above, I derived the following predictions on the decision process in a conjunctive “same”-“different” judgment:

3.2.1.1 Effects of the Basic Response Types “Same” vs. “Different”

Previous experiments on “same”-“different” judgments have provided substantial evidence that the parallel processing mode and the degree of codability have a strong impact on the relation between processing times for *same* and different types of *different* stimuli (Bindra et al., 1968; Farell, 1985). Therefore, I predicted that stimuli differing in color or object class, as absolute and highly codable dimensions, should be associated with faster processing times than *same* stimuli. Similarly, the complexity of the viewing patterns, i.e. the number of glances at the objects and the number of regressions to an object viewed before, should increase for *same* stimuli. Size, on the contrary, being a less codable, relative dimension, should be associated with slower reaction times and more complex viewing patterns than *same* stimuli.

3.2.1.2 Effects of the Number of Differences in *Different* Stimuli

On the basis of findings on self-terminating search effects in conjunctive “same”-“different” judgments (Bamber, 1969; Egeth, 1966; Farell, 1985; Hawkins, 1969; Snodgrass & Townsend, 1980), I predicted that the detection times in a two- or three-dimensional difference condition should be determined by the fastest reaction time to any of the single dimensions involved. Usually this will be the dimension yielding the shortest reaction time in a one-dimensional difference condition. Similarly, the difference dimension that is easiest to detect should determine the viewing patterns and viewing times for multidimensional differences.

3.2.1.3 Effects of the Types of Differences in *Different* Stimuli

For one-dimensional differences, I predicted an effect of codability on the eye movement patterns. Differences in relative features, such as size, should be associated with complex viewing patterns and glances to and fro both stimuli, because the size discrepancy can only be detected by means of a reference system that identifies one object as being smaller or bigger than the other. By contrast, in case of a difference in an absolute dimension, like color, the “same”-“different” decision can be drawn by first retaining the structural description of the first object and then comparing the memorized information with that extracted from the second object (cf. Carlson-Radvansky & Irwin, 1995). This strategy should be associated with rather simple viewing patterns. According to Boucart and Humphreys (1992, 1994, 1997), the processing times for form differences are longer than those for color differences. Therefore, the processing times for object class differences should be longer than those for color differences.

As outlined above, the processing times for *multidimensional* difference conditions can be predicted on the basis of the self-terminating search effect: The detection times for multidimensional differences should be determined by the fastest reaction time to any of the single dimensions involved. Thus, in a multidimensional difference condition the dimension with the highest degree of codability should determine the relative complexity of the viewing pattern and the total reaction time.

Closely connected to the analysis of reaction times to multidimensional differences is the empirical validity of models incorporating serial vs. parallel processing modes. As outlined above, experimental findings on “same” vs. “different” decision times support parallel models because “same” answers for identical stimuli can be given faster than “different” answers for difference dimensions of low codability (see Farell, 1985, for a review of the argumentation). In the present experiment, I therefore expected the processing times for identical stimulus pairs to be faster than those for stimulus pairs differing in size. However, although all dimensions are processed in parallel, identical stimuli have to be checked exhaustively with regard to all dimensions, whereas “different” answers can be based on a self-terminating search. Therefore, I predicted faster reaction times to *different* stimuli differing in absolute and highly codable dimensions such as color or object class than to *same* stimuli.

3.2.2 Method

3.2.2.1 Participants

12 female and 5 male students of the University of Bielefeld took part in the experiment. All participants were right-handed. The experiment took about 40 minutes and the subjects were paid DM 8,- for their participation.

3.2.2.2 Stimuli

108 pairs of stimuli were created from combinations of the dimensions object class, color, and size. Based on German category norms (Mannhaupt, 1983), typical representatives of three categories (animals, household furniture, clothing) were selected to form the three levels of the dimension object class (Katze (cat), Lampe (lamp), Hose (trousers)). The objects were semantically and visually dissimilar and were matched in terms of grammatical gender, number of syllables of their names, and concreteness. Line drawings for each object were taken from the Snodgrass and Vanderwart collection (1980). Nine copies of each object were prepared varying in color (red, blue, yellow) and size (small, medium, large), resulting in 27 multidimensional objects.

By combining the objects to pairs, 54 *same* and 54 *different* pairs were created. The group of *different* object pairs was composed of three subsets of 18 pairs each, containing items with one-, two-, and three-dimensional differences respectively. The groups of one- and two-dimensional differences each consisted of three subgroups with six pairs each (cf. Figure 2).

In sum, there were seven groups of difference types, namely color (C), object class (O), and size (S) in the one-dimensional group (three sets of six pairs each), color and size (CS), color and object class (CO), and size and object class (SO) in the two-dimensional group (three sets of six pairs), and the group of stimulus pairs varying in all three dimensions (CSO; one set of 18 pairs). Object pairs differing in size were created by combining medium-sized objects with large and small objects in equal shares. When two objects of a pair did not differ in size, they were both medium-sized. The ratio between the sizes of two objects differing in size was always 5 : 4, i.e., the same ratio was applied for large and medium objects (1.25 : 1) and medium and small objects (1 : 0.8). The objects were scaled to fit into a frame of 3.01° x 3.14° (medium), 3.77° x 3.94° (large), and 2.41° x 2.51° (small), respectively, with a mean distance from the screen of 60 cm.

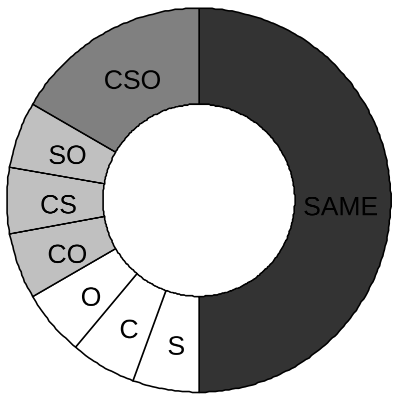


Figure 2. Construction of stimulus pairs.

For the analysis of the eye movement data, I needed to find a measurable criterion to identify the point in time when the decision had been taken and the button press was initiated. I therefore positioned the two objects used for the decision task at the top left and right corners of the display and added a small symbol at the bottom. Participants were instructed to first do the decision task on the two objects at the top of the screen, and to decide then whether the symbol at the bottom was a plus or a cross. The answers were given verbally by saying “Plus” when there was a plus, and saying nothing when there was a cross. I chose different answer modalities for the two tasks in order to minimize interference effects. Participants were asked to do the two tasks one after the other as quickly as possible. This should force them not to stick to the first task until the end of the motor reaction but to start with the second task as soon as the decision on the first task was made. Proportionally distributed among the conditions, one sixth of all object pairs was combined with a plus, the rest was combined with a cross. The plus/cross symbol was centered at the bottom of the screen. The two objects were positioned in the upper corners of the screen with a distance of 2.52° from the borders of the screen and a distance of 7.81° from each other (see Figure 3, for an example).

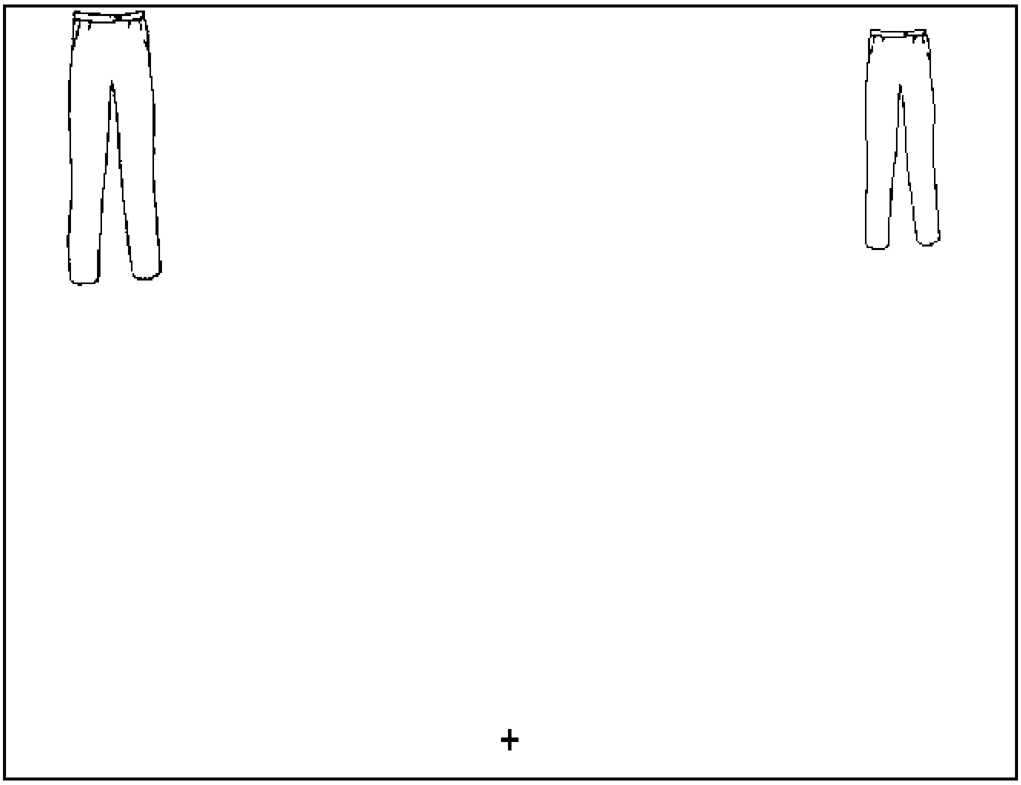


Figure 3. Example of the arrangement of the object pair and the symbol at the bottom in an item display (condition S). Note that the scales and ratios given above are not accurately sketched.

Although a differentiation between target and context object seems superfluous in case of a display with only two objects, there were several reasons to maintain this distinction for the present experiment: To control for effects of preferred scanning patterns, two items were constructed of each stimulus pair. In one item the target object was displayed in the upper left and the context object in the upper right corner, and in the other item the positions of the objects were switched. The stimulus set thus consisted of 216 items. The target object was always the object, to which the participants' first gaze was guided by means of a fixation point presented immediately before the stimulus display at the target position. In order to implement size as a relative dimension in size-discrepant object pairs, the target object was always assigned to a medium size level, so that its relative size could be varied by choosing either a smaller or a bigger context object.

3.2.2.3 Design

The experiment consisted of three nested within-subjects factors, namely *R-Type* (2) with the correct Response Types “same” and “different”, *D-Number* (3) with the number of differences involved in *different* object pairs, and *D-Type* with seven levels. The levels corresponded to the difference types described above. The items were assigned proportionally to six blocks with each block consisting of 18 *same* and 18 *different* stimulus pairs. The *different* pairs of a block included one item of each of the one- and two-dimensional conditions and three items of the three-dimensional condition. In all blocks, 1/6th of the *same* and *different* stimulus pairs was combined with a plus, the rest with a cross. In half of the *same* stimulus pairs and half of the *different* stimulus pairs of each block the fixation point and the target object were positioned on the left and the context object on the right; in the remaining stimuli the positions of target and context object were reversed. The order of experimental blocks was randomized. Within each block, the items were pseudo-randomized for each subject. Successive items never contained identical stimuli at one or both object positions of the item display.[27]

3.2.2.4 Apparatus

The experiment was controlled by a Compaq Pentium 4000 computer. The items were presented on a Sony Triniton 20’’ monitor. Reaction times were registered with pushbuttons, assigning different buttons to the *same* and the *different* condition. Via an SMI HW-EyeLink-HM eye tracking system participants’ eye movements were monitored with a sampling rate of 250 Hz. Onset and offset times and coordinates of all fixations were extracted from the data recorded by the eye tracker.

[27] This allows for the possibility that two items of the *same* or the *different* condition may follow each other. However, results of previous experiments (Krueger, 1973; Nickerson, 1973; Williams, 1972) suggest that the effects of the immediacy of individual stimulus elements are even stronger than those of the recency of response types. Therefore, I randomized the items with regard to the stimuli occurring within the items rather than the conditions they belong to.

3.2.2.5 Procedure

Prior to the experiment, participants received written instructions on the two decision tasks (“same”-“different”; “plus”-“cross”) and on the function of the fixation point. They were asked to do each task as fast and accurately as possible. Then the headband of the eye tracker was mounted and the system was calibrated. A set of nine *same* and nine *different* trials of all difference types and difference numbers was included in a practice block preceding the experimental blocks. All participants first practiced the sequential execution of the two tasks and got used to the display and the assignment of response categories to the two buttons. On the first 1000 ms of each trial, a fixation point was presented in the upper left or right corner of the screen depending on the position of the upcoming target object. Immediately following, the object pair was presented together with the symbol at the bottom of the screen for 3200 ms. Reaction times were measured from the onset of the stimulus display until a button was pressed. Reactions to the “same”-“different” decision task that took longer than 1800 ms were registered as time-outs. During each trial, participants’ eye movements were recorded. The verbal reactions to the “plus”-“cross” decision task were monitored during the experiment, but they were not analyzed any further.

3.2.2.6 Analysis

Distinct stimulus areas within the display were defined in pixels, resulting in one stimulus area each for the upper left corner, the upper right corner, and the symbol at the bottom of the screen. All fixations lying inside the contours of an object or less than 1.25° away from it were scored as object fixations. In addition, I defined a fixation area positioned between the two stimuli, as this area might be of special importance during the comparison of the objects. For each stimulus area, all fixations were extracted; onset and offset of the fixations, their durations, and their coordinates were registered. The onset times of fixations starting before and ending after stimulus onset were recoded by zero. Depending on the position of the target object, fixations on the two objects were coded as target and context object fixations respectively. When participants had turned away from an object, the viewing duration of the current fixation block was computed as difference between the offset of the last fixation and the onset of the first fixation on the object. Such blocks of consecutive fixations on an object will be termed “gazes” and their duration “gaze duration”. Note that this definition of gaze duration differs from the one given in section 2.4.1 (see Just & Carpenter, 1980; Henderson, Pollatsek and Rayner, 1987, 1989) that does not include the duration of the saccades within a fixation block. Viewing patterns were defined on the basis of the temporal order of gazes on the target object (T), the context object (C), the symbol at the bottom of the screen (x) and the intermediate region between the two objects (B). Eleven different viewing patterns were defined, namely (1) T → C → x, (2) T → C → T → x, (3) T → C → T → C → x, (4) T → x → C/T/B, (5) B → T/C/x, (6) x → T → C, (7) x → C → T, (8) C → T → x, (9) C → T → C-x, (10) C → T → C → T → x, (11) C → x → T/C/B, accounting for 96 % of the eye tracker data.

Statistical analyses were run over both subjects and items as random factors. An item was defined as one instantiation of a difference type, i.e., each of the 216 stimulus displays was regarded as an item. I will report *F1*-statistics (using subject variation) and *F2*-statistics (using item variation). It is possible to obtain significant results in separate *F1*- and *F2*-statistics, but a non-significant F-value in an ANOVA including both subject and item variance (Raaijmakers, Schrijnemakers, & Gremmen, 1999; H.H. Clark, 1973). I computed F_{min}-values based on the *F1*- and *F2*-statistics. All significant results presented below yielded significant F_{min}-statistics.

3.2.3 Results

The data from Experiment 1 were analysed with regard to error rates, viewing patterns, viewing times and reaction times.

3.2.3.1 Error Analysis

The data from 88 trials (2.4%) were discarded, because participants pressed the wrong button (21 trials, 0.6%), or did not react in time (67 time-outs, 1.8%). The analysis of the error rates revealed neither significant effects of the factor R-Type, nor of the factors D-Number and D-Type. Error rates were not systematically related to the viewing patterns.

3.2.3.2 Analysis of Viewing Patterns

The analysis of viewing patterns was restricted to trials with valid reaction times. Viewing patterns starting from the context object (266 trials = 7.7%), from the symbol at the bottom of the screen (6 trials = 0.2%), or from the intermediate area between the two objects (151 trials = 4.4%) were discarded from the analysis. The remaining valid viewing patterns starting from the target object (3026 trials = 87.7%) mainly consisted of viewing patterns with fixations on the target object and the context object (T $\rightarrow$ C $\rightarrow$ x: 1936 trials = 56.1%, T $\rightarrow$ C $\rightarrow$ T $\rightarrow$ x: 701 trials = 20.3%, T $\rightarrow$ C $\rightarrow$ T $\rightarrow$ C $\rightarrow$ x: 83 trials = 2.4%). 306 trials (= 8.9%) were associated with a direct viewing pattern (T $\rightarrow$ x), i.e. only the target object, but not the context object was fixated before the gaze was shifted to the symbol at the bottom of the screen. The intermediate region between target and context object turned out to be of minor importance, as the registered fixations were almost exclusively single fixations of short duration that occurred during the subject's change of gaze from one object to the other. They were not included in the analysis of viewing patterns, but were taken into account in the analysis of total viewing times (see below).

I classified the viewing patterns with regard to the *complexity* of the exploration into *simple* vs. *complex* patterns. All patterns with at least one gaze at each object were classified as complex viewing patterns. The complex patterns were further classified according to their *extensiveness* into patterns *without regressions* (T-C-x) and patterns *with regressions* to an object fixated before (T-C-T-x and T-C-T-C-x, respectively). Table 1 summarizes the relative frequencies of all pattern types. As Table 2 shows, the analysis of the complexity and the extensiveness of the viewing patterns revealed significant effects of the factors R-Type, D-Number, and D-Type on the viewing patterns.

3.2.3.2.1 Effects of R-Type

The proportion of complex patterns and of patterns with regressions was significantly larger under the *same* than under the *different* condition (cf. Tables 1 and 2). Participants obviously checked both objects exhaustively, i.e. with regard to all dimensions, when making a "same" decision. In contrast, a "different" decision was apparently based on any difference that was detected.

Table 1. Relative frequencies of the different viewing patterns in percent, broken down by the factors R-Type, D-Number, and D-Type.

	Viewing patterns			
	Complexity		Extensiveness	
	simple	complex	no regressions	one/two regressions
R-Type				
Same	5.3	94.7	63.0	37.0
Different	15.3	84.7	80.2	19.8
D-Number / D-Type				
1-dim. differences	9.2	90.8	68.5	31.5
C	13.3	86.7	80.0	20.0
S	2.2	97.8	46.6	53.4
O	12.9	12.9	84.4	15.6
2- dim. differences	15.2	84.8	86.5	13.5
CO	18.9	81.1	86.5	13.5
CS	11.0	89.0	84.3	15.7
SO	15.7	84.3	89.0	11.0
3- dim. differences	21.8	89.2	87.2	12.8
CSO	21.8	69.2	87.2	12.8

3.2.3.2.2 Effects of D-Number and D-Type

The observed proportions of viewing patterns under the two- and three-dimensional conditions correspond to those observed under the color and object class condition in the one-dimensional group. Participants apparently based their "different" response on the dimensions that were easiest to detect (self-terminating search; cf. Table 1). Analyses of the factor D-Type for the subgroups of one-, two-, and three-dimensional differences showed significant effects for the group of one-dimensional differences only. Size was associated with significantly more complex viewing patterns than color or object class. The latter conditions, in turn, did not differ from each other. With regard to the groups of two- and three-dimensional differences there were no significant differences (see Table 1).

Table 2. ANOVA results: Effects R-Type, D-Number and D-Type on the complexity (simple vs. complex) and extensiveness (no regressions vs. one or more regressions) of the viewing patterns.

	Subjects		Items	
	df	*F1*	*df*	*F2*
		Complexity[a]		
R-Type	1,16	15.01**	1,214	6.27***
D-Number	2,32	8.61**	2,106	7.56**
D-Type[c]	6,96	6.80***	6,102	7.15***
D-Type[d]	2,32	12.12***	2, 33	7.08**
		Extensiveness[b]		
R-Type	1,16	17.60**	1,214	17.72***
D-Number	2,32	8.18**	2,106	7.93**
D-Type[c]	6,96	21.31***	6,102	39.34***
D-Type[d]	2,32	16.23***	2, 33	10.29***

Note. [a] The analyses are based on the relative frequencies of complex viewing patterns observed for each subject/item and each respective condition
[b] The analyses of extensiveness are based on the relative frequencies of viewing patterns without regressions observed for each subject/ item under the respective conditions
[c] overall effect of the factor D-Type
[d] effect of the factor D-Type within the group of one-dimensional differences
** $p < .01$, *** $p < .001$

3.2.3.2.3 Summary

The analysis of viewing patterns supports the main hypotheses. Corresponding to exhaustive as opposed to self-terminating search strategies for *same* versus *different* stimuli, the viewing patterns under the *same* condition were significantly more complex than those under the *different* condition. Within the *different* condition, differences in size as a relative feature were associated with significantly more regressions than differences in the more codable features object class and color. Recall, however, that this effect of codability may be confounded with effects of discriminability.

The comparison of viewing patterns between the difference types with one as opposed to two or more difference dimensions provides evidence for a self-terminating search strategy: In the conditions with two- or three-dimensional difference types, the frequencies of simple viewing patterns and of patterns without regressions were nearly identical to those registered under the highly codable one-dimensional difference types (color and object class). This seems to be due to the easy detectability of color and object class differences, which can be processed faster than differences in size. Obviously size, being part of the conditions SC, SO, and SCO, was not processed after a difference in color or object class had been detected, as it did not influence the viewing patterns under these conditions in any way.

3.2.3.3 Analyses of Viewing Times and Reaction Times

In order to conduct analyses of viewing *times*, I defined the following parameters:

VT(T)	:=	viewing time of the target object, defined as sum of the duration of the first gaze at the target object and all regressions to it;
VT(C)	:=	viewing time of the context object, defined as sum of the duration of the first gaze at the context object and all regressions to it (VT(C) was = 0 for direct viewing patterns without any gaze at the context object);
VTtot	:=	VT(T) + VT(C) (fixations on the area between target and context object were included if they had occurred prior to the first fixation on the plus/cross);

For each dependent variable, processing times within a given condition that deviated by more than two standard deviations from the respective participant’s and item’s mean were replaced by estimates following the procedure recommended by Winer (1971). The proportion of replaced values was below 5 % for all variables. Figure 1 displays the results for the *same* condition and all *different* conditions.

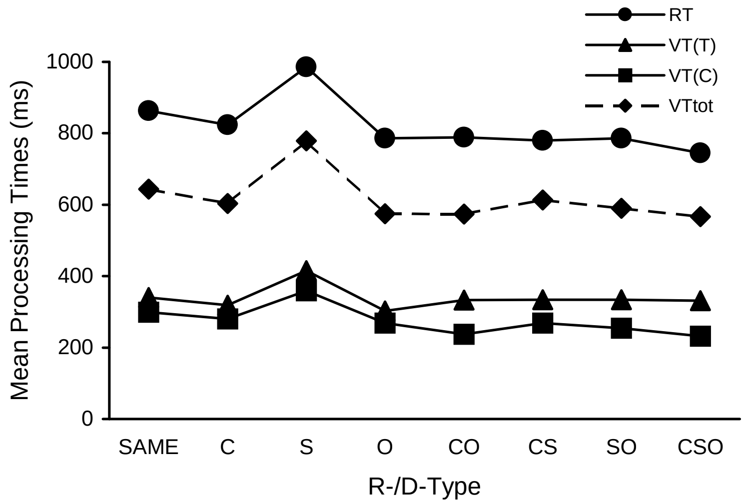


Figure 3. Mean values for RT, VTtot, VT(T), and VT(C), displayed for the *same* condition and each of the *different* conditions. Dashed lines for VTtot indicate that VTtot is no direct measure of processing time but represents the sum of VT(T) and VT(C).

As outlined above, only the viewing times preceding the first fixation on the icon at the bottom of the screen were evaluated. The similarity of the results obtained for reaction times and total viewing times (dashed lines in Figure 1) can be regarded as an index of the high validity of total viewing times as an indicator for the processing time for the objects. Recall that I had included the second task to be able to differentiate between overall processing times (measured as total viewing times) and reaction times (measured as the moment of the button press). In an ANOVA over the difference between total viewing times and reaction times including all *different* conditions and the *same* condition, I obtained no significant effects of condition. The time period between total viewing time and reaction time can thus be interpreted as (constant) motor latency.

3.2.3.3.1 Effects of R-Type

As Table 3 shows, participants reacted significantly faster to *different* than to *same* object pairs ($F1(1,16) = 10.97$, $p < .01$; $F2(1,214) = 23.58$, $p < .01$). The total time taken to explore the whole display was significantly shorter for *different* than for *same* object pairs ($F1(1,16) = 10.24$; $p < .01$; $F2(1,214) = 16.11$, $p < .01$). Considering the viewing times of the target and the context objects separately, the effect of the factor R-Type reached significance only for the context object ($F1$(1,16)

= 16.39, $p < .01$; $F2(1,214) = 20.78$, $p < .01$), but not for the target object (both *F*s < 1).

Table 3. Mean processing times and standard deviations by subjects for each response type

	RT	VT(T)	VT(C)	VTtot
same				
M	863	340	299	644
SD	142	64	67	111
different				
M	800	337	264	605
SD	142	84	81	129

These results are in line with the findings on the exploration patterns for *same* vs. *different* object pairs, showing fewer simple viewing patterns and more extensive viewing patterns with regressions for the *same* than for the *different* condition. Subjects were faster in judging two objects as being “different”, than they were in judging two objects as being “same”: For a “different” decision they did not have to scan both objects exhaustively with regard to all dimensions, but could make a decision as soon as they had found a difference. Because of the larger proportion of simple patterns under the *different* condition, the context object was looked at less often than under the *same* condition, yielding shorter mean viewing times of the context object. I had expected to find a similar effect of the smaller number of regressions under the *different* condition on the viewing times of the target object. Contrary to that prediction, however, the extensiveness of the viewing patterns was not to directly correlated to the viewing times of the objects in the display. I assumed that this might be due to differences between the individual duration of single gazes within a complex sequence of gazes at both objects. More detailed analyses revealed structural differences between the patterns with vs. without regressions. The findings are attached in Appendix A on “Structural Differences between Viewing Patterns”.

3.2.3.3.2 Effects of D-Number

The factor D-Number exhibited significant effects on all dependent variables, except for the viewing times of the target object ($F1(2,32) > 16$, $p < .001$; $F2(2,106) > 7$, $p < .01$ for RT, VT(C), and VTtot). As Table 4 shows, reaction times and viewing times were significantly longer for one- than for two- and three-dimensional differences. Paired comparisons revealed significant differences between the processing times of one- and two-dimensional difference types ($t1(16) > 4.6$, $p < .001$; $t2(71) > 2.5$, $p < .01$ for RT, VT(C), and VTtot) and of one- and three-dimensional difference types ($t1(16) > 5.9$, $p < .001$; $t2(71) > 3.4$, $p < .001$ for RT, VT(C), and VTtot). Two- and three-dimensional difference conditions did not differ significantly. This finding supports the notion of a self-terminating search strategy in visual discrimination. The detection times of multidimensional differences appeared to be determined by an easily detectable dimension and were independent of the number of differences involved (cf. Table 4)

Table 4. Mean processing times and standard deviations by subjects for one-, two-, and three-dimensional difference types

	RT	VT(T)	VT(C)	VTtot
1-dim. differences				
M	868	348	305	656
SD	146	73	75	125
2-dim. differences				
M	785	332	254	592
SD	145	91	84	135
3-dim. differences				
M	745	331	232	567
SD	136	99	101	139

I had predicted that the differences between viewing patterns in terms of complexity and extensiveness should be related to differences in processing times. The relative complexity of the observed viewing patterns was correlated with the viewing times of the context object (see above), but there were no effects of the relative extensiveness of the patterns on the viewing times of the target object. For further analyses of this result, please refer to Appendix A ("Structural Differences Between Viewing Patterns".)

3.2.3.3.3 Effects of D-Type

The main effect of the factor D-Type was significant for RT, VT(T), VT(C), and VTtot ($F1(6,96) > 7.7$, $p < .001$; $F2(6,102) > 5.7$, $p < .001$ for all variables). As Table 5 and Figure 1 show, the processing times for size differences were slower than for all other difference types. The remaining difference types did not differ substantially from each other.

Table 5. Mean processing times and standard deviations by subjects for each difference type

	RT	VT(T)	VT(C)	VTtot
C				
M	824	319	280	603
SD	154	75	80	135
S				
M	986	415	359	779
SD	190	79	73	131
O				
M	786	303	269	575
SD	139	95	110	151
CO				
M	789	333	237	574
SD	185	95	107	120
CS				
M	780	334	269	613
SD	133	76	80	134
SO				
M	786	334	254	590
SD	148	126	102	167
CSO				
M	745	331	232	567
SD	136	99	101	139

I conducted separate Analyses of Variance for three groups of difference types, namely the group of *one-dimensional differences* (S, C, O), the *two-dimensional differences* (CO, CS, SO), and finally the group of *more-than-one-dimensional differences*, consisting of the difference types CO, CS, SO, CSO. As I had expected – given the self-terminating search effect reported above – I did not obtain any significant results with regard to the two last-mentioned groups, but only for the group of one-dimensional differences ($F1(2,32) > 11.8$, $p < .001$; $F2(2,34) > 6.2$, $p < .001$ for RT, VT(T), VT(C), and VTtot).

Within the group of one-dimensional differences the predicted effects of the different degrees of codability of absolute and relative dimensions were confirmed: Size was processed significantly more slowly than color ($t1(16) > 4.7$, $p < .001$; $t2(32) > 2.5$, $p < .01$ for all variables) and object class ($t1(16) > 3.9$, $p < .01$; $t2(32) > 3.3$, $p < .01$ for all variables). Contrary to my initial expectation, color and object class did not differ significantly from each other (cf. Table 5). For the "same"-"different" decision process, the absoluteness of a dimension seemed to be more crucial than the color vs. form aspect. Dunnet tests ($p < .05$) that were conducted to determine the relation between the individual one-dimensional difference conditions and the two- and three-dimensional conditions, revealed significant differences between size and the multidimensional conditions only. All multidimensional differences were processed as fast as one-dimensional color or object class differences.

3.2.3.3.4 Additional Analyses of Codability Effects

As outlined at the beginning of this chapter, Bindra et al. (1968) introduced the notion of codability in order to explain the inhomogeneous pattern of results for *same* vs. *different* stimuli and the *fast-"same" phenomenon*. They showed that the observed discrepancies were due to the fact that *same* stimuli were processed more slowly than stimuli differing in absolute dimensions but faster than stimuli differing in relative dimensions. In line with Bindra and colleagues, I obtained significantly slower processing times for stimulus pairs differing in size than for identical stimulus pairs (Dunnet tests, $p < .05$ for RT, VT(T), VT(C), VTtot; see also Figure 1). In the remaining comparisons (*same* vs. C / O / CO / CS / SO / SCO), the "same" processing times were slower than the respective "different" processing times. For

RTs, Dunnet tests ($p < .05$) revealed significant differences between the *same* condition and each of the remaining stimulus conditions except for color (O, CO, CS, SO, CSO). The analyses of VT(C) and VTtot showed significant differences between *same* and CO and *same* and CSO only. For the viewing times of the target object, I obtained no significant differences between the *same* condition and the remaining difference types.

These findings are in line with a parallel processing mode (Allport, 1971; Bamber, 1969; Bindra et al., 1968; Donderi & Case, 1970; Donderi & Zelnicker, 1969; Downing & Gossman, 1970; Egeth, 1966; Hawkins, 1969): If the identity check of all dimensions for a “same” decision had been conducted in a serial manner, the processing times would have been even longer than those for a “different” decision in size-discrepant shapes.

3.2.3.3.5 Summary

The results of the analysis of viewing times and reaction times clearly confirm the assumption of a self-terminating search strategy that is based upon a codability effect. Color and object class as absolute dimensions were processed significantly faster than size as a relative dimension, and they therefore determined the processing times for multidimensional differences. Effects of color vs. form processing did not influence processing times of color and object class differences (cf. Boucart & Humphreys 1992, 1994, 1997). What seems to determine the reaction time to a difference in object class is the absoluteness of the object class dimension, but not the fact that it is a form dimension.

3.3 Visual Discrimination of Multidimensional Stimuli

The present findings provide evidence from both viewing patterns and processing times in support of the hypothesis that conjunctive “same”-“different” decisions on two objects are made through parallel, self-terminating processes. Participants looked for differences as long as necessary to detect one and checked the dimensions exhaustively only if the stimuli were identical. Thus, for multidimensional stimuli, decision times as well as viewing times and viewing patterns were determined by the difference dimension that was easiest to detect (self-terminating search). Ease of detection is closely connected to codability: Differences in absolute and highly codable dimensions, such as color or object class, were detected faster than differences in size as a relative dimension. Note, however, that the effects of codability may be confounded with discriminability effects. The comparison of color and object class differences did not reveal any effects of color vs. form processing. What did seem to be relevant for the decision process was the absoluteness of the object class difference but not the fact that it was a difference in form. The relative processing times registered for *same* as opposed to *different* object pairs support the notion of a parallel processing mode.

In the next chapter, I will outline the use of the findings from Experiment 1 in view of experimental findings on the referential communication paradigm. I will extrapolate the present findings and argue that they account for overspecifications in referential noun phrase descriptions as perceptually grounded phenomenon.

4 Determinants of Referential Overspecifications

In the referential communication task, speakers have to refer to multidimensional objects in the context of other multidimensional objects by specifying a set of features that clearly distinguish the intended object from the surrounding objects. In chapter 2, I introduced the empirical phenomenon of referential overspecifications: Speakers often utter features of the object to be specified that are redundant in view of a minimally contrastive specification (Eikmeyer & Ahlsén, 1998; Herrmann & Deutsch, 1976; Schriefers & Pechmann, 1988; Pechmann, 1989). Following Whitehurst, I want to argue that it is the availability of features that determines the form of an object specification: "While contrastive descriptions are efficient in terms of words they may be inefficient in terms of the effort to appropriately analyze the stimulus array." (Whitehurst, 1976, p. 478). Provided that the analysis of distinctive features in a referential communication task can be reduced to multiple "same"-"different" decisions, I will account for the high frequency of overspecifications in referential communication on the basis of the effects of self-terminating search and codability obtained in Experiment 1 and the incrementality of speech production processes (Pechmann, 1989; Schriefers & Pechmann, 1988).

I consider the case of *minimal* specifications first. Here, the process of referring to an object in a referential communication task can be reduced to three main stages of the production process:

1. detecting differences between target object and context object(s),
2. evaluating detected differences with regard to their distinctiveness, and
3. verbalizing the minimally distinctive features by means of a complex noun phrase, e.g. "the large green lamp".

The "same"-"different" experiment presented in chapter 3 revealed the main characteristics of the first stage. In chapter 2, findings on how speech production processes work and how the verbalization stage might be modeled were presented (Bock & Levelt, 1994; Dell, 1986; Schade & Eikmeyer, 1998; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; Schade, 1999). Yet, there are no detailed accounts of the second stage, i.e. the evaluation of stimulus dimensions with regard to their relevance.

A review of the literature on “same”-“different” judgments reveals a number of investigations on the processes involved in making “same”-“different” decisions on one dimension while disregarding a second dimension (Ballesteros & Manga, 1996; Besner & Coltheart, 1976; Bundesen & Larsen, 1975; Dixon & Just, 1978; Jolicoeur & Besner, 1987; Krueger, 1973; Miller & Bauer, 1981; Sekuler & Nash, 1972; Watanabe, 1988a). In order to model the processes underlying such complex decision tasks, Krueger (1978) presented the “noisy-operator theory”, and Miller and Bauer (1981) developed the “relevance rechecking model”, which was modified by Watanabe (1988a). These models all assume two basic stages: The first serves for difference detection only, whereas the second involves decisions on the relative relevance of the detected features. Krueger and Watanabe assumed that during the first stage, some irrelevant features could be filtered out, though not all. While Krueger explained the insufficient early filtering on the basis of noise, Watanabe presented a more detailed analysis of the mechanisms involved. He investigated the influence of the relation between relevant and irrelevant stimulus dimensions in terms of their relative degree of *integrality* (Garner, 1974; Garner & Felfoldy, 1970; Lockhead, 1972; Watanabe, 1988a, 1988b). In two highly integral dimensions, like orientation and form, the irrelevant dimension exerts a strong influence on the judgment of the relevant dimension. In contrast, two dimensions with a low degree of integrality (*separable* dimensions, e.g. color and form) can easily be attended to selectively and judged independently of each other. In his experiments, Watanabe (1988a, 1988b) found low degrees of integrality for color and size and for color and form, but high degrees of integrality for orientation and size. He specified the relevance rechecking model by Miller and Bauer as follows: If the dimensions involved in a “same”-“different” decision task with one relevant and one irrelevant dimension are separable, “the information coming from this irrelevant dimension should be filtered out at the first stage” (Watanabe, 1988a, p. 141).

The (modified) relevance rechecking model can be used to derive predictions on the availability of stimulus dimensions for the description of a multidimensional target object in the context of other multidimensional objects. Note, however, that in the experiments cited above on the effect of irrelevant differences on “same”-“different” judgments of relevant dimensions, subjects were explicitly instructed which dimension they had to disregard. For the production of minimal specifications of multidimensional objects in a referential communication task, the evaluation of relevant vs. irrelevant dimensions is part of the naming task. Nevertheless, the (modified) relevance rechecking model can be used to approximate the evaluation processes in the analysis of complex object displays in referential communication, as I want to propose it here.

4.1 Experiment 2

To accomplish a precise assessment of the relation between the processes of detection and naming, I ran a naming experiment on the basis of the stimulus material used in Experiment 1.

4.1.1 Predictions

Taking into account Watanabe's (1988a, 1988b) findings in combination with the self-terminating search effect and the large impact of codability on the "same"-"different" decision latencies obtained in Experiment 1, I derived the following hypotheses:

1) Differences in relative dimensions, such as size, that co-occur with differences in absolute dimensions, such as color or object class, should be filtered out at the first stage. After being filtered, these dimensions are – as I will term it – *functionally invisible* in view of higher order processes. Note, however, that they are – in principle – visible and perceivable.
2) Differences in absolute dimensions have to be rechecked at the second stage with regard to their relevance.

In view of experiments on referential communication, *minimal* specifications can only be produced on the basis of both stages within the relevance rechecking model. This implies a rather large cognitive effort. Although a minimal specification would fulfill basic communication rules, such as Grice's (1975) maxim to be minimal, it is usually not necessary to produce *minimal* specifications but to produce *unambiguous* specifications. Thus, Hypothesis 2 can be modified as follows:

2a) It may cost less effort to specify irrelevant differences in absolute dimensions than to explicitly ignore them (“principle of least effort”, “economy principle”; cf. Whitehurst, 1976; Pechmann, 1994).

Based on hypotheses 1 and 2a) I predicted that overspecifications of size as a relative dimension should occur rather seldomly, whereas color as an absolute dimension should be overspecified more frequently. Previous experiments had shown precisely this pattern of results (cf. Eikmeyer & Ahlsén, 1998, for a review); however, these experiments provide only little evidence on the procedural origin of referential overspecifications.

4.1.2 Method

4.1.2.1 Participants

7 male and 10 female students of the University of Bielefeld, who had not attended Experiment 1, took part in the experiment. They were all were right-handed and spoke German as their mother tongue. The experiment took about 50 minutes and the subjects were paid DM 10,- for their participation.

4.1.2.2 Stimuli and Design

To provide as much accordance as possible with Experiment 1, the stimulus material and the randomization procedure were the exactly the same as in Experiment 1. The design was identical, too, though not all of the factors were relevant for the purpose of the present experiment.

4.1.2.3 Apparatus

The experiment was controlled by a Compaq Pentium 4000 computer. The items were presented on a Sony Triniton 20’’ monitor. All verbal reactions were recorded on DAT-tapes. Reactions to the “plus”-“cross” decision task were registered with a pushbutton panel.

4.1.2.4 Procedure

Participants were asked to name the target object in a way that a listener would be able to identify it on the display. The target object was marked by the fixation point preceding the object display. The whole set of items of the detection task was used, including all *same* items. Participants were asked to say “same” when the two objects were identical and to name the target object when the objects differed. The “plus”-“cross” decision task was now carried out using a pushbutton panel to avoid interference between the naming task and the “plus”-“cross” decision task. Participants were instructed to push the left button if there was a plus and to do nothing if there was a cross.

Prior to the experiment, participants received written instructions on the naming task and the “plus”-“cross” decision task. They were asked to do each task as fast and accurately as possible. A set of nine *same* and nine *different* trials of all types and numbers was included in a practice block preceding the experimental blocks and all participants first practiced the sequential execution of the two tasks. On the first 1000 ms of each trial the fixation point was presented in the upper left or right corner of the screen depending on the position of the target object. Immediately following, the two stimulus items were presented together with the symbol at the bottom of the screen for 4000 ms. Verbal reactions were recorded on a DAT-tape and verbal reaction times were measured by means of a voice key. Reactions that took longer than 2500 ms were registered as time-outs. The pushbutton reactions to the “plus”-“cross” decision task were monitored, but were not analyzed any further.

4.1.3 Results

All verbal reactions were transcribed and coded as to the number and the types of attributes specified in the noun phrases. As Table 6 shows, color was overspecified substantially more often than size. Color was also overspecified when it was merely present, but not varied between the objects (Conditions S, O, SO). However, the presence of an irrelevant difference in color in conditions CO and SCO lead to substantially higher rates of overspecifications, about 80%, compared to the conditions without irrelevant color variation (cf. Table 6). The comparison of conditions C and CO with regard to the frequencies of color overspecifications revealed a significant difference ($F1(1,17) = 8.74$, $p < .001$; $F2(1,46) = 35.88$, $p < .001$). Size, in contrast, was hardly ever overspecified.

In condition CS, when minimal specifications included either color or size (Det-C-N, Det-S-N), color was specified more often than size (Det-C-N: 128 utterances; Det-S-N: 17 utterances). A Wilcoxon Signed Ranks Test for related samples showed that this difference was highly significant (p = .00016, T^+ = 152, z = 3.57, N = 18; because of the sample size (N > 15) the sum of ranks was transformed to a z-value; cf. Siegel & Castellan, 1988).

170 of all utterances included both color and size specifications. Most of them were utterances of the type SCN (152 utterances; 89.41%); the rest were inverted order phrases (18 CSN-phrases; 10.59%).

Table 6. Percentages of specification types and overspecifications for each Difference Type (D-Type)

D-Type	Minimal Specification	Example		Specification Types (%) Min. Spec.	Overspecifications C	S	C & S
C	the black ball	●	○	100	–	–	–
S	the large ball	●	●	61.8	38.2	–	–
O	the ball	●	★	33.5	66.5	–	–
CO	the ball	●	☆	19.8	80.2	–	–
CS	the black ball the big ball	●	○	68.1 9.0	22.9		–
SO	the ball	●	★	30.7	59.8	2.0	6.5
CSO	the ball	●	☆	20.1	74.3	0.2	5.4

4.2 Early Perceptual Origin of Referential Overspecifications

In sum, the findings presented above support the hypotheses developed on the basis of the relevance rechecking model and the principle of "least effort": Size-differences, co-occurring with differences in absolute dimensions, such as color or object class, are filtered out early during the detection process and are thus functionally invisible in view of the formulation process. The relative detectability of color as opposed to size differences directly influences the selection of prenominal adjectives and thus determines in large parts the form of the object specifications. The filter mechanisms for size and color are based on purely perceptual effects, i.e., the formation of referential overspecifications originates already on the level of visual perception.

As indicated in section 2.3.1.1, Pechmann (1989, 1994) and Schriefers and Pechmann (1988) stressed the importance of the incremental character of speech production processes for an account of referential overspecifications. If speakers always waited for the results of a complex evaluation of all dimensions with regard to their relevance for a minimal object specification, they would have to postpone the initiation of speech production processes, too. Thus, by planning and producing their utterances incrementally, speakers are able to initiate articulation processes earlier and to thereby produce fluent utterances. Pechmann (1989) suggested that speakers use color as an absolute dimension strategically: As soon as the first piece of information is available, linguistic encoding processes are initiated, while, at the same time, the relevant contextual alternatives are inspected in more detail. Thus, it is “characteristic of such a strategy that the speaker articulates features of the target before he has determined whether they are distinguishing or not” (Pechmann, 1989, p. 98)

The incremental processing mode in speech production seems to be an elegant explanation for referential overspecifications. Yet it cannot account for the canonical order of adjectives in complex noun phrases: In canonical order phrases, size is named before the color, although color is detected earlier than size (see chapter 3). In Experiment 3, presented in the next chapter, I addressed this inconsistency. In this experiment, I also assessed effects of task-difficulty on the occurrence of inverted adjective orders (see section 2.3.2.3).

5 Determinants of Prenominal Adjective Order

In section 2.3.2, I had introduced the phenomenon of the canonical adjective order in complex noun phrases and had presented two approaches to account for this phenomenon, the visuo-semantic and the procedural approach. The findings from Experiments 1 and 2 provide evidence for both approaches: Following the argument of Pechmann (1989, 1994) and Schriefers and Pechmann (1988), the large number of color overspecifications is due to the incremental mode of processing between conceptual preparation and grammatical encoding. At the same time, the data are readily explicable in the framework of the visuo-semantic approach on the basis of the relative detectability of color and size differences on the one hand and the relative effort associated with filtering out irrelevant color information on the other.

Experiment 3 was designed to investigate the canonical order effect on the basis of the results from Experiments 1 and 2. By means of eye movement analyses, I wanted to track the time course of the evaluation of distinctive features of the target object and the subsequent linguistic encoding processes more complex situations than in the two-objects situation.

5.1 Extrapolation to the Multiple-Objects Situation

In Experiment 2, only two objects were used. In this situation, it is enough to name either color or size or none of them; there is no condition in that both dimensions have to be named in order to minimally specify the target object (see Chapter 4, Table 6). The aim of Experiment 3 was to extrapolate from this two-object situation to a more complex situation, including relevant differences in color and size. In view of the quality of the eye tracker data, however, it was necessary to keep the number of objects within the display as small as possible. As the focus of Experiment 3 lay on the relative order of color and size adjectives, I decided to leave object class differences out of consideration. In German, the object class is specified definite noun phrases anyway, independent of other object classes being present or not. There is no such construction as “the red one” in English, and although it is correct to say “der/die/das Rote” in order to refer to one of several colored exemplars of an object, such an utterance would be elliptic.

5.2 Experiment 3

Three exemplars of an object with different colors and/or sizes were presented on each trial. In the following, the object exemplars presented in the display will be referred to as *objects*. The superordinate object type will be termed *object class.* Like in Experiment 2, one of the three objects was marked as target object by means of a fixation point preceding the object display at the position of the target object. Participants were asked to name the target object either a) in such a way that a listener would be able to identify it (*neutral* instruction group) or b) in a minimal way, i.e. to name the minimally distinctive features alone (*minimal* instruction group).

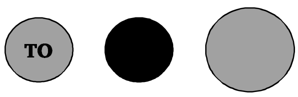
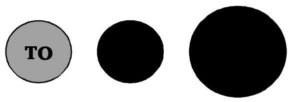
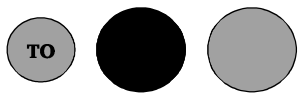
By including a minimal instruction group, I wanted to assess indirectly the relative facilitation that is achieved by applying the principle of least effort, i.e. by overspecifying redundant features instead of producing minimal specifications. As the results from Experiments 1 and 2 have shown, there is a strong effect of visual perception on the form of the object specifications. According to the findings obtained so far, size discrepancies coinciding with color discrepancies in one object are often filtered out early, and color is specified although it is not minimally distinctive. Under the minimal instruction, these overspecifications of color are incorrect, i.e. participants will have to evaluate each detected difference with regard to its relevance. By including a minimal instruction group, I varied the relative task difficulty between participants. On the basis of previous findings on the effects of task difficulty on the occurrence of inverted adjective orders (cf. chapter 2.3.2.3), I predicted that the increased task demands in the minimal instruction group should lead to the production of more inverted adjective orders. The analysis of the eye movements and processing times in the minimal instruction group might then allow inferences about the procedural origin of such inversions of the canonical adjective order.

5.2.1 Overview of the Experiment

Three experimental conditions were used that differed with regard to the minimal specification of the target object: In all conditions, color and size variations were present in the display. Between conditions, the combination of the target object with color and/or size-discrepant context objects was varied systematically (cf. Table 7):

- In condition SCO, both color and size had to be specified in order to refer unambiguously to the target object.
- In condition CO, color was the minimally distinctive feature of the target object; size was varied irrelevantly.
- In condition SO, size was the minimally distinctive feature of the target object, and color was varied irrelevantly.

Table 7. Example of the construction of object displays for conditions SCO, CO and SO

	SCO			CO (S irrelevant)			SO (C irrelevant)		
Minimal speci-fication	der kleine rote Ball (the small red ball)			der rote Ball (the red ball)			der kleine Ball (the small ball)		
Example									
Objects	TO	CF	SF	TO	CF	SF	TO	CF	SF
TO vs. CO	TO	– C	– S	TO	– C	– CS	TO	– CS	– S

Note. All objects belong to the same object class; all objects differ with regard to color and/or size.

Sixteen objects of different semantic categories were chosen for the construction of experimental items. Half of the object names were monosyllabic; the other objects had disyllabic names. To rule out gender effects, masculine, female and neuter exemplars were selected of each semantic category. The objects with their German names are listed in Table 8, together with the respective semantic categories (cf. Mannhaupt, 1983) and gender types.

Table 8. Objects and object names in German

Sem. Category	male		feminine		neuter	
Tools	hammer	Hammer	saw	Säge	knife	Messer
Furniture	chair	Stuhl	lamp	Lampe	bed	Bett
Animals	dog	Hund	cat	Katze	pig	Schwein
Toys	ball	Ball	doll	Puppe	wheel	Rad
Garment	shoe	Schuh	cap trousers	Mütze Hose	shirt	Hemd

For the construction of experimental items, nine objects of each object class were constructed by means of systematic variations of color and size. Three colors (red, blue, yellow) and three sizes (small, medium, large) were used; their German names are given in Table 9. In order to model size as a relative dimension, the target object was assigned a medium size only, and size discrepant context objects were assigned either a smaller or a larger size. In accordance with experiments 1 and 2, the ratio between the sizes of the target object and any size discrepant context object was 4 : 5 and 5 : 4, respectively. Objects were scaled to fit into frames of 3,01° x 3,14° (large), 2,41° x 2,51° (medium) and 1,92° x 2,01° (small) with an approximate distance of 60 cm from the screen.

Table 9. Colors, sizes, and their names

Colors		Sizes	
red	rote	small	kleine
blue	blaue	large	große
yellow	gelbe		

Note. Size and color names are given in their inflected form as it is used in singular, nominative, definite noun phrases. In contrast to the infinite forms of the adjectives, the inflected forms are all disyllabic. The forms are the same for all gender types.

In order to create the experimental items, three objects of one object class were combined. One of the three objects was constructed as the *target object* (TO in the following); the other two objects were context objects (cf. Table 7). In relation to the target object, the first context object was color discrepant in all conditions and the second context object was size discrepant from the TO. They will be called *color foil* (CF) and *size foil* (SF) in the following. However, depending on the conditions, the CF and the SF could incorporate additional differences in size or color:

- In condition SCO, the CF was color-discrepant from the TO, and the SF was size-discrepant from the TO. In this constellation, both color and size had to be named in order to specify the target object minimally (cf. Table 7).
- In condition CO, the CF was color-discrepant from the TO. The SF differed from the TO in terms of both color and size. As can be inferred from Table 7, it would be sufficient to name the color of the TO in order to refer to it minimally; size was thus varied irrelevantly in this condition.
- In condition SO, the SF was size-discrepant from the TO and the CF differed from the TO in terms of both color and size. As Table 7 shows, it would be enough to name the size of the TO for a minimal specification, i.e., in this condition color was varied irrelevantly.

As outlined above, color and size were varied in all conditions. However, it never occurred that more than two colors and two sizes were present in a given display: If both context objects differed from the TO in terms of color (size), the colors (sizes) of the context objects were identical (cf. conditions CO (SO) in Table 7).

The three objects were aligned from left to right. In order to vary the relative position of the TO within the object alignment, three *main* array types were defined, namely one with the TO on the left, one with the TO in the middle and one with the TO on the right. Within these main array types, the positions of the context objects were varied, too, resulting in a balanced set of six array types (see Figure 5). The relations between TO, CF and SF in the respective conditions and array types are illustrated in Figure 5. Each object occurred twice in each position within the array, with the TO in the middle in array types 3 and 4, the CF in the middle in array types 1 and 6, and the SF in the middle in array types 2 and 5. Note that array types 1 and 6, 2 and 5, and 3 and 4 are each mirror images of each other.

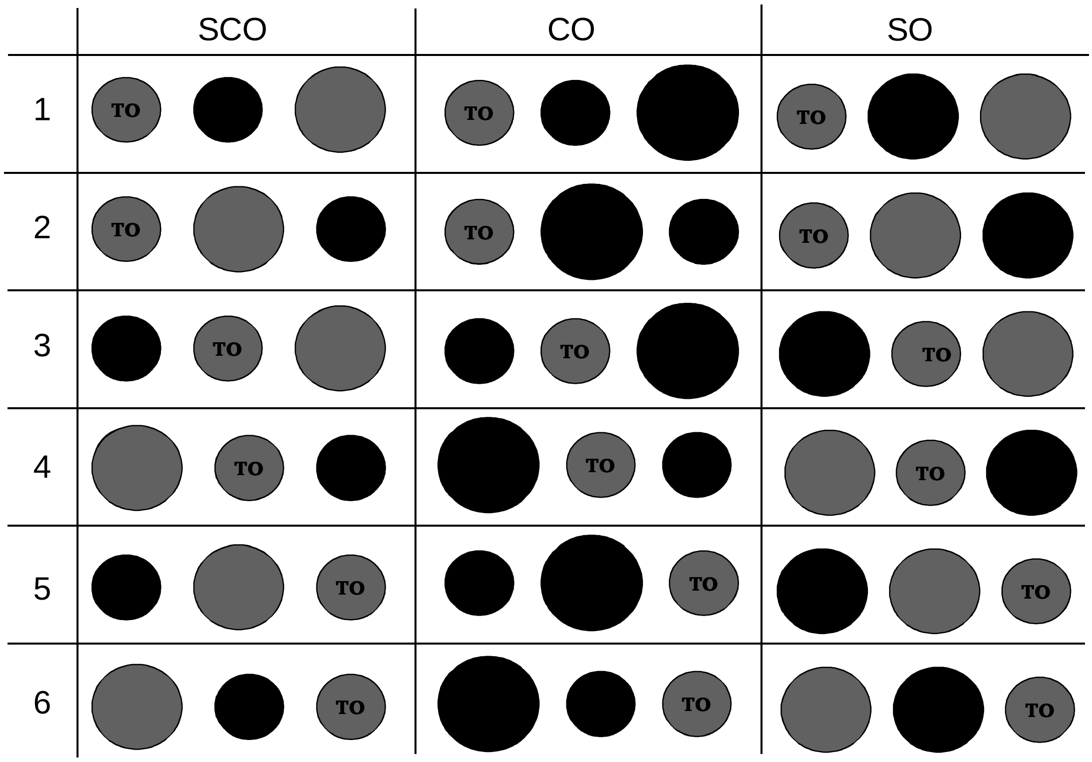


Figure 5. Overview of Array Types after the complete randomization in each condition

Under each condition, three items were constructed from each object class. Each item was then assigned to a different *main* array type. As two array types were available for each *main* array type, eight combinations of selected array types were possible. The combinations were assigned randomly to the 16 object classes, such that each combination of array types occurred twice within one condition.

5.2.2 Hypotheses and Predictions

As outlined above, the findings of experiments 1 and 2 had revealed a strong influence of early visual perception processes on the selection of prenominal adjectives for a multidimensional object specification. In particular, the effects of codability and self-terminating search and the functional invisibility of size turned out to be important determinants of the form of the object specifications (cf. sections 3 and 4). Based on these features, the following hypotheses should hold with regard to the present investigation:

H1 The easy detectability of differences in color (→ codability effect) leads to an early perceptual grouping of the three objects of the display with regard to the color of the TO: Only those objects that are of the same color as the TO will be considered; color discrepant objects will not be evaluated any further (→ self-terminating search; functional invisibility of size; see Experiments 1 and 2). Within the remaining objects of the same color as the TO, size discrepancies will then be considered to generate an unambiguous conceptual representation of the target object specification.

H2a Size differences, co-occurring with color differences in one object, will be filtered out early during perception (→ functional invisibility of size; see Experiment 2).

H2b Size differences co-occurring with color differences in one object will be disregarded if no minimal specification is required (→ principle of least effort; see Experiment 2). Otherwise, color-discrepant objects have to be explicitly rechecked with regard to their size.

H3 Due to increased task difficulty, there should be more inverted adjective order phrases in the minimal instruction group than in the neutral instruction group (see section 2.3.2.3).

On the basis of these hypotheses, the following predictions were derived for the neutral und the minimal instruction group:

5.2.2.1 Neutral Instruction Group

Because of the early perceptual grouping (→ H1), color discrepant context objects should be associated with shorter viewing times and fewer glances at them than context objects of the same color as the TO. The latter should be viewed longer for a more detailed comparison in view of differences in size. Context objects that differ from the TO in both color and size should be regarded as briefly as color discrepant objects because the size discrepancy should be filtered out early (→ H2a + H2b). Therefore, no overspecifications of size were expected in condition CO. Because of the high degree of codability of the color dimension and its easy detectability, reaction times under condition CO should be considerably shorter than under conditions SCO and SO. In the latter two conditions, size was a minimally distinctive feature and should be detected more slowly than the color differences in condition CO (→ H1). There should be a high rate of color overspecifications in condition SO, as the detected color difference in the CF will not be rechecked with regard to its relevance (→ H2a + H2b).

Because of the linear alignment of the objects, the object in the middle position of the array should be fixated more often. Such effects of the array type on the number of glances and the viewing times should occur for the TO in array types 3 and 4, for the CF in array types 1 and 6 and for the SF in array types 2 and 5 (cf. Figure 5).

5.2.2.2 Minimal Instruction Group

In the minimal instruction group, participants were not allowed to produce overspecifications. In condition CO, this should not be associated with additional effort compared to the neutral instruction group, as it should be easy to filter out the irrelevant size discrepancy in the SF in this condition (H1, H2a + H2b; see above). In condition SO, however, the CF, which is color- and size-discrepant from the TO, will have to be inspected more thoroughly than in the neutral instruction group, as for a minimal specification of the TO the relevance of the detected color difference in CF has to be rechecked (H2a + H2b). This additional effort might be regarded as indirect evidence for what is gained by applying the principle of least effort and producing an overspecified utterance.

The minimal instruction should increase the difficulty of the task. On the basis of previous findings on the effect of task difficulty on the occurrence of inverted adjective order phrases significantly higher rates of non-canonical adjective order phrases were predicted for the minimal in comparison with the neutral instruction group (→ H3; cf. Pechmann, 1994; Pechmann & Zerbst, 1990, 1995).

Parallel to the neutral instruction group, effects of the array type on the number of glances and the viewing times should occur for the TO in array types 3 and 4, for the CF in array types 1 and 6 and for the SF in array types 2 and 5 (cf. Figure 5).

5.2.3 Method

5.2.3.1 Participants

34 subjects took part in the experiment. All participants were undergraduate students at the University of Bielefeld and were native speakers of German. The instruction groups consisted of 17 participants each and were matched with regard to age and gender. The experiment took about 45 minutes and all participants were paid DM 8,- for taking part.

5.2.3.2 Stimuli

The construction of the stimulus material for each condition and array type has been described above. As in Experiments 1 and 2, a fixation point was presented prior to the object display at the position of the upcoming TO (left, middle, right). Participants were asked to look at the fixation point and to name the object that would appear at the position of the fixation point. In order to be able to distinguish the exploration processes associated with the naming task from those that were uncritical in view of utterance generation processes, a second task was added to the naming task. As in Experiments 1 and 2, the three objects were aligned at the top of the screen and an additional symbol (plus or cross) was integrated at the bottom of the display. Participants were asked to first name the TO (minimally) and to then decide, whether the icon at the bottom of the screen was a plus or a cross. As in Experiment 2, the reaction to the “plus”-“cross” decision task was given non-verbally by means of a pushbutton panel. A plus was assigned to one fourth of the items of each condition; the rest was combined with a cross. In order to force subjects to make a distinct eye movement away from the object display and to prevent preview effects on the icon at the bottom of the screen, the icon and the three objects were positioned in maximum distance: The three objects were aligned at the top of the screen and the symbol was centered at the bottom of the screen. The positions of the three objects were determined irrespectively of the size differences by the midpoints of the respective imagery frames (see above). These frames were positioned in such a way that the distance between the three objects amounted to 2,53°. The distance between the left- and rightmost objects from the edges of the screen was fixed to 1,53°, viewed from a mean distance of 60 cm from the screen.

5.2.3.3 Design

Three nested within-subjects factors were included in the experiment, namely *Condition* (3) with the levels SCO, CO and SO, *Number of Syllables* (2) and *Array* (6). The factor *Instruction* (2) with the levels neutral and minimal was varied between subjects. The experimental items were assigned proportionally to four blocks of 36 items, each block including 12 items of each experimental condition. A practice block with 12 items was constructed to precede the experiment. Four additional practice blocks with three items each were created to precede the individual blocks during the experiment.

The order of main experimental blocks was varied individually for each subject within a given instruction group. Within each block, the order of items was pseudo-randomized for each subject. The objects included in two successive items always differed with regard to object class and semantic category. In addition, the position pattern, i.e. the alignment of TO, CF and SF, was never the same for two successive items.

5.2.3.4 Apparatus

The experiment was controlled by a Compaq Pentium 4000 computer. The items were presented on a Sony Trinitron 20’’ monitor. During the experiment, the participants’ eye movements were recorded via an SMI HW-Eye-Link-HM eye tracking system with a sampling frequency of 250 Hz. For the analysis of the data recorded by the eye tracker, onset and offset times and pixel coordinates of all fixations were extracted. Verbal reactions were recorded on a DAT-tape and verbal reaction times were measured by means of a voice key. The pushbutton reactions to the “plus”-“cross” decision task were monitored, but were not analyzed any further.

5.2.3.5 Procedure

Prior to the experiment, participants were asked to read an instruction on the naming task and the “plus”-“cross” decision task. They learned about the composition of the display and the index function of the preceding fixation point. Participants were asked to do the two tasks one after the other, and to do each task as fast and accurately as possible. They were explicitly informed that the display time of each item would be sufficient to do the two decision tasks successively. Then the headband of the eye tracker was mounted and the system was calibrated. In an initial practice block, all participants practiced the sequential execution of the two tasks and familiarized themselves with the structure of the display. Each trial started with the presentation of the fixation point in the upper left/middle/right position of the display for 1500 ms, depending on the respective position of the target object. Immediately afterwards, the three objects were presented together with the symbol at the bottom of the screen for 5000 ms in the neutral instruction group or for 5300 ms in the minimal instruction group. The voice key was activated on the onset of the stimulus display and was triggered as soon as it registered a sound. When the naming latency was longer than 3000 ms, the reaction was coded as time-out. Reactions to the “plus”-“cross” decision tasks were monitored during the experiment, but they were not included in the later analysis.

5.2.3.6 Analysis

All verbal reactions were transcribed and coded with regard to their form. Only those noun phrases were further analyzed that specified the target object correctly according to the respective condition. In the neutral condition, this included all overspecifications of the target object; in the minimal condition overspecifications were coded as errors. On the basis of the DAT-recordings of the utterances, erroneous trigger reactions of the voice key were corrected for by means of a digital audio editor. In order to prepare the eye tracker data for statistical analyses, stimulus areas were defined within the display, with one stimulus area each for the objects on the left, middle and right position of the object alignment at the top of the screen and for the icon at the bottom of the screen. All fixations lying within the contours of an object or less than 1.2° away from it were scored as fixations on that object. For each area, all fixations were extracted from the onset until the end of the recording of the eye tracker. The onset and offset times and the coordinates of each fixation were extracted, and fixation times were computed. Onset times of fixations beginning before stimulus onset and ending during the stimulus presentation were recoded by zero. Only those fixations that preceded the first fixation on the icon at the bottom of the screen were used for the analysis.

Depending on the array types involved, the display areas were coded as belonging to the TO, the CF or the SF. For each object, the fixations extracted from all items were collected. The resulting fixation data for all objects were merged. When participants had turned away from an object, the viewing duration of the current fixation block was computed as difference between offset of the last fixation and the onset of the first fixation. As in Experiment 1, such blocks of consecutive fixations on an object will be termed “gazes” and their duration “gaze duration” (see chapter 3; Experiment 1). Viewing patterns were defined on the basis of the order of gazes on the TO, the CF and the SF. As there was a large diversity of observed viewing patterns, they could not be assessed directly for a statistic analysis of the eye tracker data, like it had been conducted in Experiment 1. Therefore, several meta-variables were computed: For TO, CF and SF, the individual number of glances at these objects was extracted (N(TO), N(CF) and N(SF) in the following), and the overall number of glances during the exploration (N(tot)) was computed as their sum. In addition, the variable N(obj) was introduced to indicate how many of the three objects of the display were fixated during the exploration process. Beyond the analysis of viewing patterns, analysis of processing times were conducted. Therefore, the viewing time of each object, VT(TO), VT(CF) and VT(SF), was computed as sum of all gaze durations.

As expected, the results of the statistical analyses of viewing times were parallel to those obtained in the analyses over the meta-variables to capture the viewing patterns. To keep the main body of the text readable, I will include the statistical analyses for *viewing times* only; the statistical analysis of viewing patterns via the meta-variables defined above is given in Appendix B.

In order to specify more exactly the processes associated with planning and generating the object specification and articulating the utterance, separate analyses were run over viewing patterns and processing times *before utterance onset* and *during articulation*. All glances that were registered before utterance onset were subsumed under the first period of time, also including glances that started before utterance onset and ended during articulation. All other glances were subsumed under the articulation period (cf. Figure 6).[28]

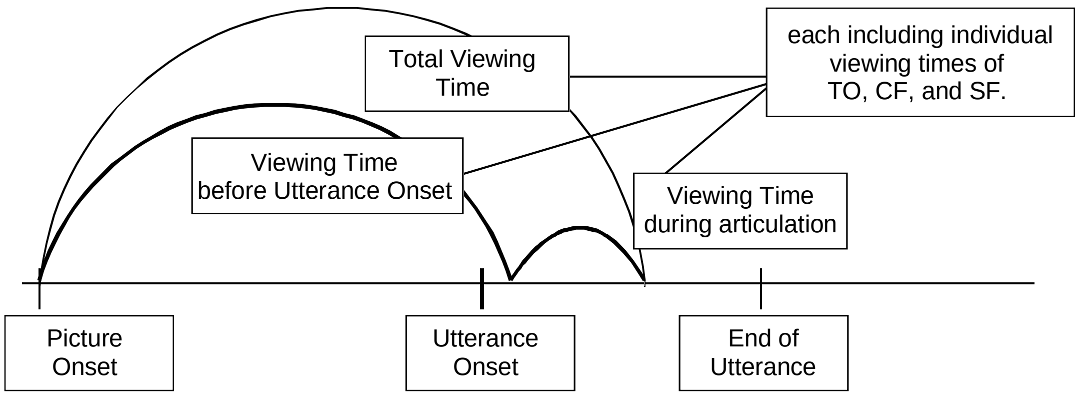


Figure 6. Timing of a trial: Definition of dependent variables

For the statistical analyses of viewing patterns and processing times, ANOVAS were run over the variables defined above. Separate ANOVAS were run for each instruction group, including the factors *Condition* (3), *Number of Syllables* of the name of the object class (2), and *Array* (6). In the neutral instruction group, AN*C*OVAS were run, integrating the variable *utterance length* as a covariate. As Table 9 shows, the inflected forms of the adjectives in the noun phrase were all disyllabic. Therefore, utterance length was coded as number of adjectives in a noun phrase. In addition, ANCOVAS were run over the whole set of data, including the between-subjects factor *Instruction*. The factor *Number of Syllables* of the noun was included in all AN(C)OVAS. Since neither the main effect of this factor nor its interactions with the other factors turned out to be significant in any of the analyses, it will not be mentioned in the following presentation of results. All AN(C)OVAS were run over subjects and items and additional F_{min} statistics were computed to control for artifacts induced by the separate consideration of *F1*- and *F2*-values (see chapter 3; cf. H.H. Clark, 1973; Raaijmakers, Schrijnemakers & Gremmen, 1999). When presenting analyses of processing times, I will report both *F1*- and *F2*-statistics. AN(C)OVA results presented below as being significant yielded significant F_{min} statistics. In order to inspect the data with regard to differences between levels

[28] The results of the analysis of the viewing times before utterance onset and the analysis of total viewing times (including all gazes until participants turned to the icon at the bottom of the screen; see Figure 6) turned out to be parallel: All significant effects that appeared in the first also emerged in the latter.

of individual factors, paired t-tests between (adjusted) means were run for subjects and items. In the neutral instruction group, subject and item means were adjusted for effects of utterance length and corresponding t-values were computed according to Winer (1971). AN(C)OVA results and paired comparisons between levels of the factor Condition will be presented in the text; for the sake of readability, however, the results of paired t-tests between array types will be given in tabular form in Appendix B.

As described in Experiment 1, only the data obtained until the first gaze at the symbol at the bottom of the screen were evaluated. As outlined above, separate analyses were run over viewing patterns and processing times *before utterance onset* and *during articulation*. Note that all glances that were registered before utterance onset were subsumed under the first period, including glances that started before utterance onset and ended during the articulation. All other glances were subsumed under the articulation period.

5.2.4 Results

The data from Experiment 3 were analysed with regard to error rates, specification types and utterance structures, reaction times, viewing patterns, and viewing times.

5.2.4.1 Error Analysis

5.2.4.1.1 Neutral Instruction Group

10.2 % of the answers were coded as errors: The data of 247 trials had to be excluded because subjects had not reacted in time (8 time-outs, 0.3 %), had hesitated before or during the articulation (133 trials, 5.5 %), had misnamed the color, size, or object class of the target object (93 trials, 3.9 %), or had underspecified the target object (13 trials, 0.5 %). In addition, 38 trials (1.6 %) had to be discarded because of technical errors.

5.2.4.1.2 Minimal Instruction Group

The error rate was significantly higher in the minimal than in the neutral instruction group ($F1(1,32) = 6.43$, $p < .05$; $F2(1,143) = 16.73$; $p < .001$). Note, however, that in the minimal instruction group, overspecifications were coded as errors, too. Overall, 15.2 % of the data in the minimal instruction group were excluded as errors. The data of 7.7 % were excluded as time-outs (15 trials, 0.6 %), as hesitations (103 trials, 4.2 %), as misnamings (53 trials, 2.2 %) or as underspecifications (17 trials, 0.7 %). 8.2 % of the answers were coded as overspecifications of color (130 trials, 5.3 %) or size (70 trials, 2.9 %). For one participant of the minimal instruction group the exclusion of erroneous answers led to a rate of less than 40% correct answers in condition SO. Therefore, the data from this subject had to be excluded yielding an overall rate of 18.1 % of missing values of for the analyses of processing times.

5.2.4.2 Specification Types and Utterance Structures

5.2.4.2.1 Neutral Instruction Group

In the neutral instruction group, overspecifications of size in condition CO occurred in 26.7 % of the trials. In contrast, 87 % of all specifications under condition SO included an overspecification of color, i.e., color was overspecified significantly more often than size ($F1(1,16) = 30.41$; $p < .001$; $F2(1,94) = 79.01$; $p < .001$). Only 0.9 % of all specifications including color and size specifications were inverted adjective order phrases. These results are in line with previous findings and agree with the predictions.

5.2.4.2.2 Minimal Instruction Group

In the minimal instruction group no overspecifications were allowed. Nevertheless, overspecifications of color did occur, accounting for nearly one third of all errors under this condition (see above). Many participants reported after the experiment that they had had serious difficulties in realizing that in some cases (in condition SO), the color was irrelevant and size alone had to be specified. As illustrated in Figure 7, participants obviously needed some time to overcome the strong saliency of the irrelevant color discrepancy. A more detailed analysis of the rate of overspecifications of color over the four experimental blocks revealed a significant decrease over time ($\chi^2_{0.001,\,3} = 92.35$).

Compared to the neutral instruction group, the rate of inverted adjective orders in specifications including both dimensions increased significantly from 1.1 % to 8.1 % ($F1(1,32) = 32.94$, $p < .05$; $F2(1,47) = 45.47$; $p < .001$). This finding can be interpreted as an effect of increased task difficulty (cf. Pechmann, 1994; Pechmann & Zerbst, 1990, 1995 for comparable findings). I will address the effects of the high saliency of the color dimension and the influences of task difficulty in more detail in the General Discussion (see chapter 6).

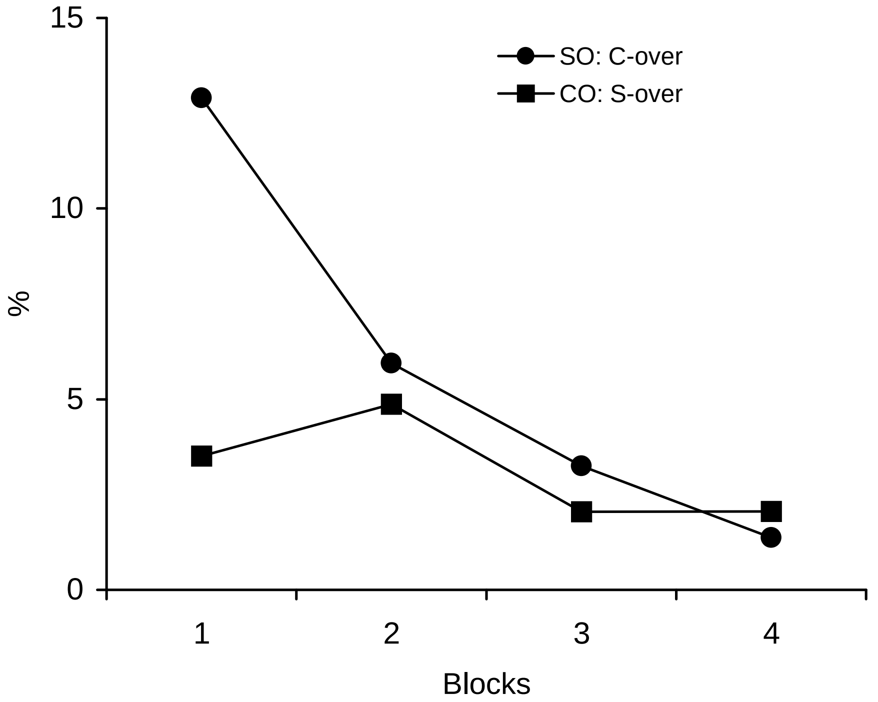


Figure 7. Rates of overspecifications of color and size in conditions SO and CO

5.2.4.3 Reaction Times

5.2.4.3.1 Neutral Instruction Group

An ANCOVA over reaction times including the factors Condition, Syllable and Array revealed a significant main effect for Condition ($F1(2,31) = 4.57$, $p < .05$, $F2(5,107) = 13.27$, $p < .001$). Participants reacted faster in conditions SO and CO than in condition SCO (cf. Figure 8); the difference between SO and SCO was significant ($t1(31) = 3.18$, $p < .01$; $t2(140) = 2.37$, $p < .01$). Whereas the observed reaction times in conditions SCO and CO corresponded to the predictions, the reaction times under condition SO were unexpectedly short. Because of the easy detectability of the color discrepancies in condition CO, I had predicted that reaction times under condition CO should be shorter than under conditions SCO and SO. In the latter two, reaction times were predicted to be similar, since in both conditions, size, being a less codable dimension than color, was a distinctive feature. Contrary to the prediction, though, the RTs under condition SO were shorter than under conditions SCO and CO. I do not have an explanation for this effect at this point.

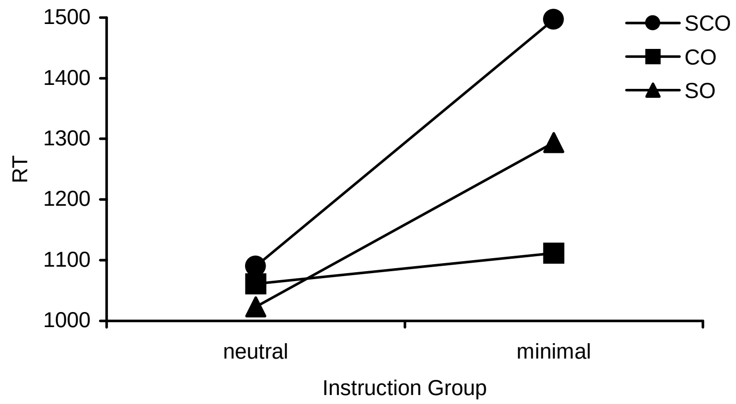


Figure 8. Mean reaction times (ms) by conditions for the neutral and the minimal instruction group.

5.2.4.3.2 Minimal Instruction Group

The main effect of Condition was highly significant ($F1(2,30) = 46.06$, $p < .001$, $F2(2,108) = 182.56$, $p < .001$). Paired comparisons between conditions revealed significantly faster reaction times in condition CO than in conditions SCO ($t1(30) = 8.92$, $p < .001$; $t2(94) = 19.42$, $p < .001$) and SO ($t1(30) = 5.34$, $p < .001$; $t2(94) = 9.82$, $p < .001$). The latter two differed significantly, too ($t1(30) = 5.24$, $p < .001$; $t2(94) = 10.38$, $p < .001$). Note, however, that the required reaction to condition SCO in the minimal instruction group included two adjectives, whereas the minimal specification of the target object in conditions SO and CO included one adjective only. Thus, the significant difference between conditions SCO and SO may in part be due to effects of utterance length.

5.2.4.3.3 Neutral vs. Minimal Instruction Group

The ANCOVA over both instruction groups including the between-subjects factor Instruction and the within-subjects factors Condition and Array did not reveal a significant main effect of any of these factors. The interaction of Instruction and

Condition, however, turned out to be highly significant ($F1(2,61) = 23.63$, $p < .001$; $F2(2,107) = 95.14$, $p < .001$): As predicted, the instruction type did not affect reaction times in condition CO. Reaction times in SO and SCO, however, were slower in the minimal than in the neutral instruction group (see Figure 8).

5.2.4.4 Viewing Patterns

As indicated above, I will confine to presenting a descriptive analysis of the viewing patterns. The results of the detailed statistical analysis is included in Appendix B. For the descriptive analysis of the patterns, tree diagrams of successive transitions between the objects were drawn for each condition. Since the relative position of the TO in the array might have affected the viewing patterns, separate trees were drawn for each main array type. The trees are organized as follows (see Figures 9a/b to 11a/b): There are six to eight strata that represent successive transitions from one object to the next. In each stratum, the proportion of transitions between objects is represented as lines of proportional thickness with thicker lines for higher percentages of transitions within the stratum. Recall that the fixation point at trial onset specified the location of the TO. Therefore all subjects looked at the TO first and all tree diagrams have the TO at the topmost position. The proportion of transitions to the symbol at the bottom of the screen is represented in each stratum. Note that the transitions within a stratum are not computed as conditional probabilities. They are computed as relative frequencies of the overall number of transitions observed in a stratum. To illustrate the relation between the overall numbers of transitions observed in different strata, each stratum is shaded from white to dark gray according to the overall number of patterns including a transition on that stratum. This implies an illustration of the mean length of patterns observed under a given condition, too. A pattern consisting of five transitions (including the transition to the symbol at the bottom of the screen) will thus be included in strata 1 to 5.

For each pattern, the number of transitions until utterance onset was computed. In the trees in Figures 9 to 11, the strata that are framed in black include 25 % to 75 % of all utterance onsets. The rest of the utterance onset was registered in an earlier or a later stratum.

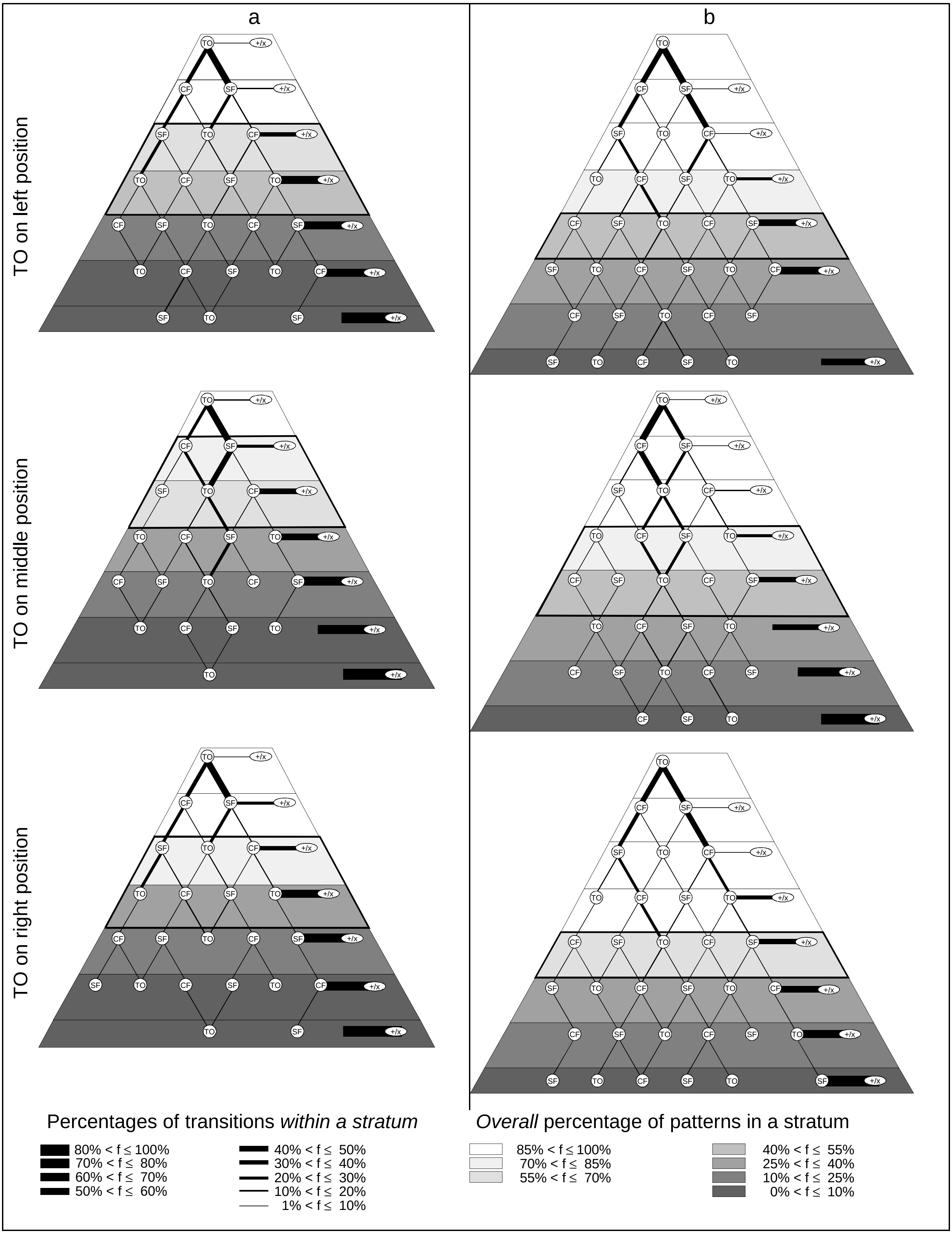


Figure 9. Viewing Patterns in condition SCO in the neutral instruction group (a) and the minimal instruction group (b)

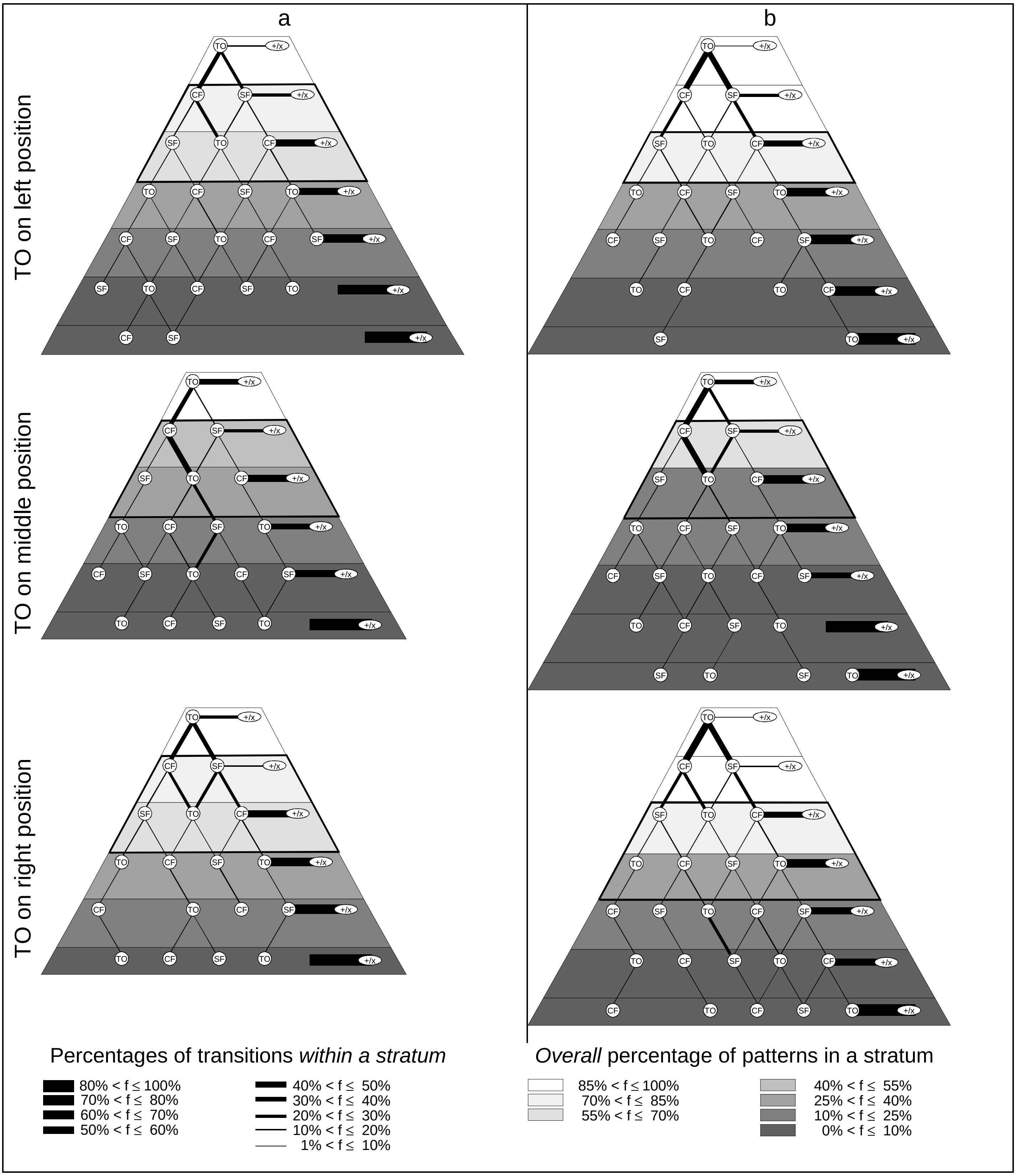


Figure 10. Viewing Patterns in condition CO in the neutral instruction group (a) and the minimal instruction group (b)

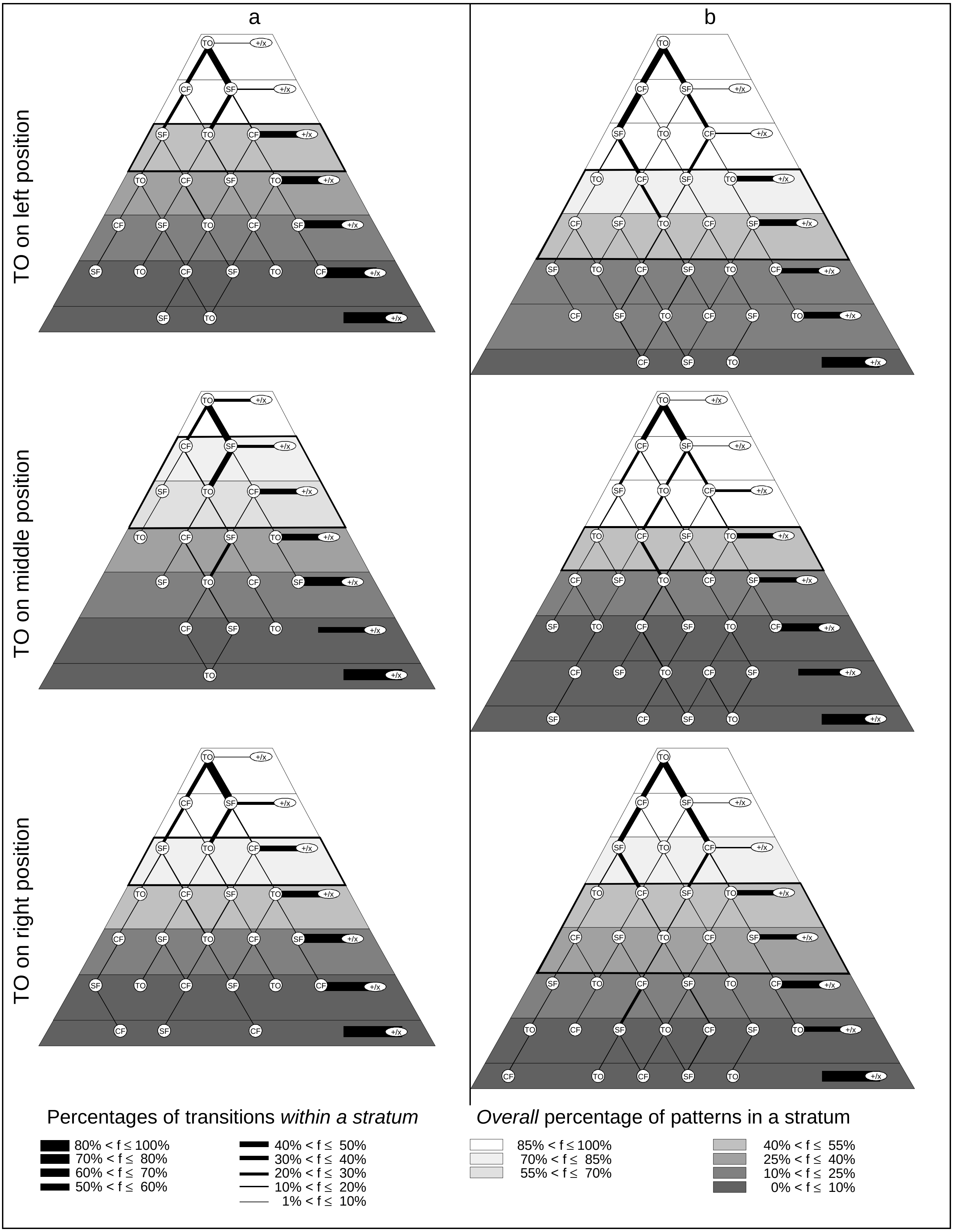


Figure 11. Viewing Patterns in condition SO in the neutral instruction group (a) and the minimal instruction group (b)

5.2.4.4.1 Neutral Instruction Group

In all conditions, only few glances were registered after utterance onset. As can be inferred from the color of the later strata, transitions within the object display *during* articulation were registered for less than 40 % of all patterns. Thus, in most of the cases participants must have turned to the “plus”-“cross” decision task soon after utterance onset.

In condition SCO, participants first turned to the SF, which had the same color as the TO, and then to the TO in most of the cases (Figure 9a). When they had turned to the CF first, they did not turn to the TO until they had also looked at the SF. When the TO was in the middle of the display, they first turned to the TO, then to the SF, and then back to the TO. Overall, there was no effect of the position of the TO on the viewing patterns; however, participants tended to fixate fewer objects before utterance onset, when the TO was in the middle of the array, than when it was on the left or right position.[29] The results for condition SO were parallel to those for condition SCO (cf. Figure 11a). However, on average, the number of glances before utterance onset was lower in condition SO than in condition SCO. This might in part account for the finding that reaction times were shorter under condition SO, than under condition SCO. In condition CO (Figure 10a), the overall viewing patterns and the patterns observed until utterance onset were shorter than in conditions SCO and SO. They were shortest when the TO was in the middle of the display. In condition CO, there was no preference to fixate either the CF or the SF first.

As predicted for conditions SCO and SO, there seem to be perceptual grouping effects at a very early stage of processing: Because of the easy (peripheral) detectability of the color discrepancy in the CF, participants either did not look at that object at all, or only once before turning to the SF, which had the same color as the TO in conditions SCO and SO. This implies, however, that in condition SO, size differences, co-occurring with a color difference in the CF, were filtered out early,

---

[29] The analysis of the meta-variables, defined to capture the basic characteristics of the viewing patterns, showed that there were more effects of the Array type on viewing patterns than just the effect of the middle position of the TO in Array types 3 and 4. The results of the more detailed analyses of the viewing patterns are described in Appendix B.

too, which accounts for the high frequency of overspecifications of color in condition SO (see above).

5.2.4.4.2 Minimal Instruction Group

Parallel to the neutral instruction group, there were only few patterns of a length that exceeded the strata including most of the utterance onsets. There was no preference to turn to the SF or to the CF first in either condition (cf. Figures 9b, 10b, and 11b). In conditions SCO and SO, participants tended to first look at each object at least once before they started to speak. When the TO was in the middle of the display, they often first looked at either the CF or the SF, returned to the TO and then looked at the other context object (SF or CF, respecitvely). In condition CO (Figure 10b), the overall viewing patterns and the patterns observed until utterance onset were shorter than in conditions SCO and SO and were shortest when the TO was in the middle of the display.

5.2.4.4.3 Neutral vs. Minimal Instruction Group

In all conditions, the viewing patterns were longer for the minimal than for the neutral instruction group. In condition CO, there were no other differences between the instruction groups. In conditions SCO and SO, in contrast, the utterance onsets were registered substantially later for the minimal instruction group. As predicted, participants in the neutral instruction group obviously grouped the objects with regard to the color of the TO at a very early stage of processing and preferred to first turn to the SF rather than to the CF. By contrast, there was no evidence for grouping on the basis of early color information in the minimal instruction group. Rather, participants in the minimal group looked at each object (in the order of their appearance) and evaluated all objects before they started to speak. In particular, they thus evaluated the color discrepant objects with regard to potential size differences to find out whether the detected color discrepancy was relevant for a minimal description of the TO. This additional effort observed under the minimal instruction can be interpreted as indirect evidence for what is gained by applying the principle of least effort in the neutral instruction group: Here, participants overspecified the color in condition SO because they apparently did not evaluate the color discrepancy in the SF in view of its relevance.

5.2.4.5 Viewing Times before Utterance Onset

5.2.4.5.1 Neutral Instruction Group

The results of the ANCOVAS of the viewing times of each object before utterance onset are given in Table 8. There were significant main effects of Condition and Array on the viewing times of the CF and the SF. The viewing times of the TO displayed a significant effect of Array only. The interaction of Condition and Array was significant for VT(TO) and VT(SF).

Table 8. Effects of Condition (C) and Array (A) for the neutral instruction group: ANOVA results for the viewing times of each object before utterance onset

	Subjects		Items	
	df	*F1*	*df*	*F2*
VT(TO)				
C	2,31	1.03	2,107	.88
A	5,79	16.49***	5,107	21.59***
C x A	10,159	2.89**	10,107	2.55**
VT(CF)				
C	2,31	11.68***	2,107	8.27***
A	5,79	33.78***	5,107	76.51***
C x A	10,159	1.43	10,107	1.80
VT(SF)				
C	2,31	3.93*	2,107	2.15*
A	5,79	21.49***	5,107	48.89***
C x A	10,159	1.98*	10,107	2.26*

Note. * $p < .05$; ** $p < .01$, *** $p < .001$

As Figure 12a shows, the viewing times of the TO were significantly longer in array types 3 and 4, when the TO was in the middle of the object display, than in the other array types (see Appendix C, Table C1). Table C2 of Appendix C summarizes the results of the separate analyses of each array type in each condition for VT(TO). The interaction is obviously due to the fact that there was a significant difference between array types 3 and 4 for condition SO, but not for conditions SCO and CO.

The viewing times of the CF were significantly longer in condition SCO than in condition SO ($t1(31) = 1.91$, $p < .05$; $t2(140) = 1.95$, $p < .05$). The CF was regarded significantly longer in condition CO than in conditions SCO ($t1(31) = 2.71$, $p < .01$; $t2(140) = 10.51$, $p < .005$) and SO ($t1(31) = 4.29$, $p < .005$; ; $t2(140) = 10.47$, $p < .005$). The short viewing times observed under SCO and SO correspond to the finding that many viewing patterns registered under these conditions included at most one, if not even no glance at the CF before utterance onset (cf. Figure 12a, Figures 9a and 11a). As Figure 12a and Table C1 in Appendix C show, the CF was looked at significantly longer in array types 1 and 6, when it was positioned in the middle of the object display.

The viewing times of the size- (and color-) discrepant SF were significantly shorter in condition CO than in conditions SCO ($t1(31) = 2.77$, $p < .005$; $t2(140) = 6.61$, $p < .005$) and SO ($t1(31) = 3.09$, $p < .005$; $t2(140) = 6.16$, $p < .005$). This is due to the more elaborate evaluation of size differences between the TO and the SF in the latter conditions that was not necessary for condition CO. In that condition, the SF was also color discrepant form the TO. The middle position of the SF in array types 2 and 5 lead to a significant increase in viewing times (see Figure 12a and Appendix C, Table C1). The remaining array types differed among each other, too; however, there were no significant differences within mirror pairs of array types (1-6, 2-5, and 3-4). In order to explore the nature of the interaction of Condition and Array, individual analyses of array types were conducted for each condition. As Table C2 of Appendix C shows, condition CO displayed a mere effect of the middle position of SF in array types 2 and 5. In conditions SCO and SO, in contrast, the viewing times observed under the remaining array types differed significantly, too, and the effect of the middle position of the SF in array types 2 and 5 was less pronounced (cf. Figure 12a). This is due to the fact that the SF, which differs from the TO in conditions SCO and SO in size, but not in color, always had to be analyzed independently of the array type in order to correctly specify the TO.

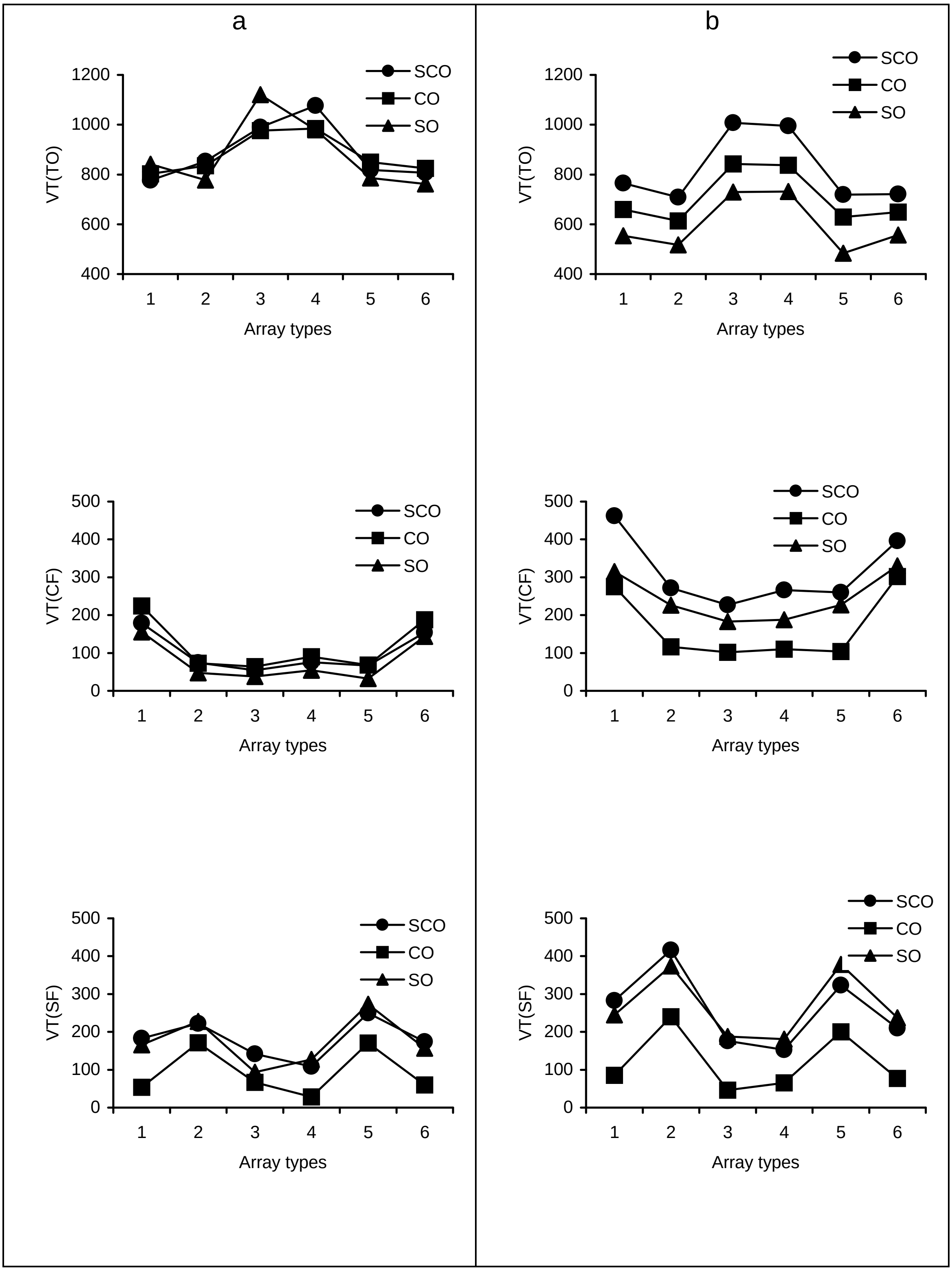


Figure 12. Mean viewing times (ms) of TO, CF, and SF before utterance onset by conditions and array types for the neutral instruction group (a) and the minimal instruction group (b).

As Figure 12a shows, the TO was looked at significantly longer than the CF and the SF. In order to analyze the relation between the viewing times of the CF and the SF in each condition, an additional ANCOVA was run over VT(CF) and VT(SF) including the within-subject factors *Object* and *Condition*. The effect of Object was significant ($F1(1,15) = 12.41$, $p < .005$; $F2(1,140) = 39.08$, $p < .001$). As the effects of Condition on VT(CF) and VT(SF) differed (see above), there was no significant main effect of Condition, but its interaction with Object was highly significant ($F1(2,31) = 35.6$, $p < .001$; $F2(2,140) = 21.59$, $p < .001$). In condition CO, the viewing times of the CF and the SF did not differ from each other, as both objects differed from the TO in the relevant color dimension. As predicted on the basis of the self-terminating search effect, the additional size discrepancy in the SF did not influence VT(SF). Paired t-tests showed that the SF was looked at significantly longer than the CF in conditions SCO ($t1(16) = 3.61$, $p < .01$; $t2(47) = 7.45$, $p < .001$) and SO ($t1(16) = 5.6$, $p < .001$; $t2(47) = 6.89$, $p < .001$). This corresponds to the prediction that – because of its color discrepancy from the TO – the CF is filtered out early in these conditions, whereas the SF and its potential size difference from the TO is explored in more detail.

5.2.4.5.2 Minimal Instruction Group

The viewing times of TO, CF and SF before utterance onset displayed significant effects of Condition and Array (cf. Table 9). The viewing times of the TO were significantly shorter in condition SO than in conditions SCO ($t1(15) = 2.75$, $p < .05$; $t2(94) = 7.21$, $p < .001$) and CO ($t1(15) = 2.74$, $p < .05$; $t2(94) = 4.14$, $p < .001$). This might be due to the increased task difficulty in the minimal instruction group that affords a more thorough analysis of the object display, particularly in condition SO (cf. Figure 12b): Here, the “functionally invisible” size difference in the SF had to be detected to find out that the additional color discrepancy in the SF is not relevant for a minimal specification of the TO. In condition SO, participants therefore had to spend more time on evaluating the SF which may well have been at the expense of the energy they would normally spend on the TO. VT(TO) was significantly longer in condition SCO than in condition CO ($t1(15) = 3.95$, $p < .001$; $t2(94) = 3.58$, $p < .001$). Note that this may be largely caused by the longer utterances to be produced in SCO. The effect of Array turned out to be a mere effect of the middle position of the

TO in array types 3 and 4, leading to a significant increase of the viewing times in these array types (cf. Appendix C, Table C3).

Table 9. Effects of Condition (C) and Array (A) for the minimal instruction group: ANOVA results for the viewing times of each object before utterance onset

	Subjects		Items	
	df	*F1*	*df*	*F2*
VT(TO)				
C	2,30	11.46***	2,108	74.98***
A	5,75	30.04***	5,108	44.74***
C x A	10,150	.67	10,108	.95
VT(CF)				
C	2,30	29.91***	2,108	126.31***
A	5,75	26.74***	5,108	73.29***
C x A	10,150	1.63	10,108	4.08**
VT(SF)				
C	2,30	59.59***	2,108	94.62***
A	5,75	22.76***	5,108	53.43***
C x A	10,150	1.11	10,108	1.61

Note. $p < .05$; ** $p < .01$, *** $p < .001$

As Figure 12b shows, the viewing times of the color- (and size-) discrepant CF were significantly smaller for condition CO than for conditions SCO ($t1(15) = 8.89$, $p < .001$; $t2(94) = 7.14$, $p < .001$) and SO ($t1(15) = 6.85$, $p < .001$; $t2(94) = 4.96$, $p < .001$). This can be regarded as task-specific effect: Beyond the salient color difference between the CF and the TO, subjects had to carefully check the CF for possible size discrepancies in conditions SCO and SO. The significant effect of Array on VT(CF) was grounded on significant differences between the array types including the CF at the middle position (1 and 6) and the remaining array types (2 to 5) (cf. Appendix C, Table C3).

The size- (and color-) discrepant SF was looked at significantly shorter in condition CO than in conditions SCO ($t1(15) = 6.84$, $p < .001$; $t2(94) = 6.75$, $p < .001$) and SO ($t1(15) = 21.3$, $p < .001$; $t2(94) = 7.82$, $p < .001$). In condition CO, participants had to detect the color discrepancy in the SF only, as the additional size discrepancy was irrelevant in view of a minimal specification of the TO. In conditions SCO and SO, in contrast, the additional size discrepancy had to be evaluated for a minimally distinctive object description. The main effect of Array on the viewing times of the SF was due to a significant increase of VT(SF) in array types 2 and 5, when the SF was in the middle of the display (cf. Appendix C, Table C3). Within the remaining array types, VT(SF) was significantly shorter for array types 3 and 4 than for array types 1 and 6, because in array types 3 and 4 the TO was in the middle and TO and SF were always "neighbored" (see Figure 5). By contrast, the TO and its SF were "disconnected" in array types 1 and 6 (cf. Appendix C, Table C3).

In order to assess differences between the viewing times of the CF and the SF, an ANOVA with the within-subject factors Object and Condition was conducted. There were significant main effects of Object ($F1(1,15) = 9.16$, $p < .01$; $F2(1,141) = 4.24$, $p < .05$) and Condition ($F1(2,30) = 60.28$, $p < .001$; $F2(2,141) = 60.25$, $p < .001$). Their interaction was not significant. The comparison of VT(CF) and VT(SF) within experimental conditions revealed that in conditions CO and SO, the CF was viewed significantly longer than the SF (CO: $t1(15) = 6.03$, $p < .001$; $t2(47) = 2.34$, $p < .05$; SO: $t1(15) = 2.91$, $p < .05$; $t2(47) = 2.16$, $p < .05$).

5.2.4.5.3 Neutral vs. Minimal Instruction Group

The ANCOVA over both instruction groups including the factors Instruction, Condition, and Array (cf. Table 10) revealed significant main effects of Condition and Array for VT(TO) and VT(SF). This finding is in line with previous results from the separate analyses within each instruction group (see above). The main effect of Instruction was significant for VT(CF) only, with significantly longer viewing times of the CF in the minimal instruction group. The interaction of Instruction and Condition was significant for VT(CF) and VT(SF): In conditions SCO and SO, the viewing times increased significantly in the minimal instruction group, whereas they remained about the same in condition CO (cf. Figure 12).

Table 10. Effects of Instruction (I), Condition (C) and Array (A): ANOVA results for the viewing times of each object before utterance onset

	Subjects		Items	
	df	*F1*	*df*	*F2*
VT(TO)				
I	1,30	1.72	1,107	5.42*
C	2,61	6.28**	2,107	20.47***
A	5,154	41.87***	5,107	59.82***
I x C	2,61	2.31	2,107	1.60
VT(CF)				
I	1,30	14.49***	1,107	18.53***
C	2,61	9.50	2,107	.07
A	5,154	57.08***	5,107	126.02***
I x C	2,61	34.04***	2,107	126.03***
VT(SF)				
I	1,30	3.72	1,107	2.88
C	2,61	44.80***	2,107	72.28***
A	5,154	42.94***	5,107	99.32***
I x C	2,61	7.81**	2,107	10.65***

Note. * $p < .05$; ** $p < .01$, *** $p < .001$

5.2.4.6 Viewing Times During Articulation

Graphic displays of the mean viewing times during articulation were scaled according to mean utterance length (~ 600 ms). As Figure 13 shows, participants turned to the symbol at the bottom of the screen rather early during the articulation process.[30]

5.2.4.6.1 Neutral Instruction Group

There were no significant effects of Condition for either of the variables VT(TO), VT(CF) and VT(SF) during articulation. Array turned out to be significant for VT(CF) ($F1(5,79) = 4.96$, $p < .001$; $F2(5,107) = 9.71$, $p < .001$), and VT(SF) ($F1(5,79) = 2.8$, $p < .05$; $F2(5,107) = 2.36$, $p < .05$).

Paired comparisons between array types for VT(CF) revealed significantly longer viewing times for array types 1 and 6 than for array types 2, 3 and 5, which is obviously due to the middle position of CF in array types 1 and 6. Similarly, VT(SF) was significantly larger in array type 5 than in array types 1, 3, 4 and 5 (cf. Figure 13a and Appendix C, Table C4).

An ANCOVA over VT(CF) and VT(SF) including the factors Object and Condition showed a significant effect of Object ($F1(1,15) = 7.71$, $p < .05$; $F2(1,140) = 13.04$, $p < .001$). The main effect of Condition turned out not to be significant, however, the interaction Object x Condition was significant ($F1(2,31) = 4.78$, $p < .05$; $F2(2,140) = 7.03$, $p < .001$). Paired comparisons revealed no significant differences between VT(CF) and VT(SF) in condition CO but significantly longer viewing times for the SF in conditions SCO ($t1(16) = 2.72$, $p < .05$; $t2(47) = 4.02$, $p < .001$) and SO ($t1(16) = 2.21$; $p < .05$; $t2(47) = 2.52$, $p < .05$). This may account for both, the significant effect of Object and its interaction with Condition (cf. Figure 13a).

[30] In previous experiments (see Meyer et al., 1998; Levelt & Meyer, 2000), the participant's gaze was lead off the object display by including another object to be named. Here, longer viewing times during articulation were observed. Thus, the short viewing times of the object display during articulation in the present experiment might be in part due to the selection of a non-linguistic task as distractor.

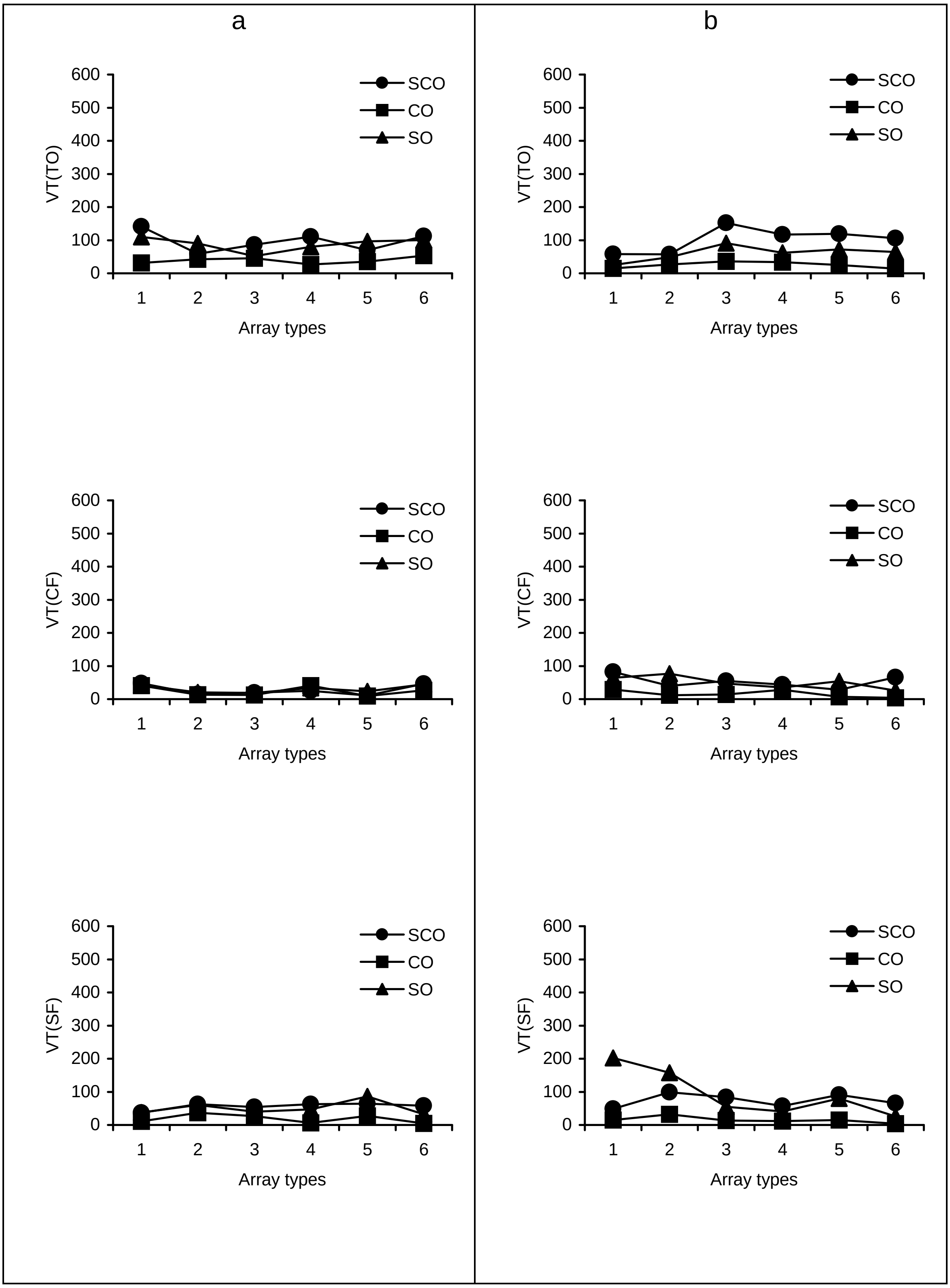


Figure 13. Mean viewing times (ms) during articulation by conditions and array types for the neutral instruction group (a) and the minimal instruction group (b).

5.2.4.6.2 Minimal Instruction Group

The results of the ANOVA over the viewing times of individual objects during articulation are summarized in Table 10. They all displayed a significant main effect of Condition. Array was significant for VT(TO) only.

Paired comparisons between conditions for each variable revealed significantly shorter viewing times in condition CO than in conditions SO and SCO for VT(TO), VT(CF), and VT(SF) (SCO-CO: $t1(15) > 3.99$, $p < .001$; $t2(94) > 3.99$, $p < .001$; SO-CO: $t1(15) > 2.65$, $p < .05$; $t2(94) > 3.06$, $p < .01$ for VT(TO), VT(CF), and VT(SF)). A more detailed analysis of the effects of array types on VT(TO) is given in Table C5 of Appendix C. There was no relation between the identified significant differences and the systematic construction of array types (cf. Figure 13b).

Table 11. Effects of Condition (C) and Array (A) for the minimal instruction group: ANOVA results for the viewing times of each object during articulation

	Subjects		Items	
	df	*F1*	*df*	*F2*
VT(TO)				
C	2,30	8.25***	1,108	32.49***
A	5,75	3.78**	5,108	5.48***
C x A	10,150	1.20	10,108	1.21
VT(CF)				
C	2,30	8.43***	1,108	7.23***
A	5,75	1.30	5,108	1.17
C x A	10,150	2.09*	10,108	1.79
VT(SF)				
C	2,30	7.50***	1,108	8.41***
A	5,75	1.60	5,108	1.75
C x A	10,150	1.39	10,108	1.08

Note. * $p < .05$; ** $p < .01$, *** $p < .001$

A separate ANOVA over VT(CF) and VT(SF) revealed significant effects of the factors Object ($F1(1,15) = 6.34$, $p < .05$; $F2(1,141) = 10.05$, $p < .001$) and Condition ($F1(2,30) = 9.84$, $p < .001$; $F2(2,141) = 4.53$, $p < .05$). Paired comparisons between VT(CF) and VT(SF) within each condition revealed significantly longer viewing times of the SF in condition SO ($t1(15) = 2.21$, $p < .05$; $t2(47) = 2.22$, $p < .05$). Both objects were looked at significantly shorter in condition CO than in conditions SCO and SO (see above), which accounts for the significant effect of Condition in the ANOVA (cf. Figure 13b).

5.2.4.6.3 Neutral vs. Minimal Instruction Group

The ANCOVA over the data from both instruction groups revealed no significant effects of Instruction. There was a significant interaction of Instruction and Array for VT(TO) ($F1(5,154) = 3.69$, $p < .01$; $F2(5,107) = 6.23$, $p < .001$), which is probably due to the short viewing times of the TO in array types 1 and 2 in the minimal instruction group (cf. Figure 13). For VT(CF) there was a significant main effect of Array ($F1(5,154) = 3.74$, $p < .01$; $F2(5,107) = 3.47$, $p < .01$) and a significant interaction of Instruction and Condition ($F1(2,61) = 4.81$, $p < .05$; $F2(2,107) = 5.9$, $p < .01$): Whereas there were no significant differences between conditions in the neutral instruction group, conditions SCO and SO were associated with significantly larger viewing times of the CF than condition CO in the minimal instruction group (cf. Figure 13). For VT(SF), there was a main effect of condition ($F1(2,61) = 7.48$, $p < .001$; $F2(2,107) = 6.98$, $p < .001$) with the viewing times in conditions SCO and SO being significantly larger than in condition CO for both instruction groups (see above).

5.2.5 Discussion

Experiments 1 and 2 provided evidence in support of strong perceptual influences on the choice of prenominal adjectives and the occurrence of redundant color specifications. The present findings allow more detailed conclusions as to how these influences become effective in a complex referential communication task: Because of the easy (peripheral) detectability of color differences in the object display, participants of the neutral instruction group were able to selectively attend to the objects that had the same color as the TO. The analysis of viewing patterns and viewing times showed that when color discrepant objects were looked at at all before utterance onset, they were viewed only briefly and at an early stage of processing. After that, they were disregarded and only objects of the same color as the TO were inspected. As predicted, size differences co-occurring with color differences in one object were functionally invisible in view of the later object specification processes. Because of the self-terminating search effect (see Experiments 1 and 2), only the color discrepancy in the CF was evaluated for the object description while the additional size difference was filtered out. Following the principle of least effort, participants did not recheck the detected color discrepancy in the CF with regard to its relevance and produced overspecified object descriptions in most of the cases. In order to explore the effort of overruling this principle of least effort and to find out what is actually gained by its application, the minimal instruction group was contrasted with the neutral instruction group. Particularly in condition SO, which included an irrelevant variation of color, participants had obvious difficulties in specifying the TO minimally. The data show that during the course of the first experimental block, subjects had to overcome the strong perceptual influence of the early classification of the objects according to their color and had to check each object with regard to discrepancies in both color and size. Corresponding to the prediction that the minimal instruction should increase the difficulty of the task, the overall error rates were higher under the minimal than under the neutral instruction group. Along with the aggravated task demands, the rate of inverted adjective orders in condition SCO increased significantly. As outlined above, participants of the minimal instruction group looked at each object at least once, before they produced an utterance. This extensive visual exploration is associated with longer overall viewing times in the minimal instruction group – compared to the neutral instruction group – and probably induced serious time pressure on the linguistic processing of the visual input. Following Pechmann (1994) and Pechmann and Zerbst (1990, 1995), the occurrence of inverted adjective orders under time pressure may result from applying the incremental mode of processing, i.e. from naming the adjectives according to their order of detection (color before size).

Most of the perceptual processes described above apparently occurred at an early stage of processing, as the descriptive analysis of viewing patterns has shown (see section 5.2.4.4). Similarly, the linguistic planning of the whole noun phrase seems to be fully completed before articulation is initiated. There are no systematic effects of utterance length on the viewing times during articulation, as they would have to be expected on the basis of the findings by Meyer and colleagues (Levelt & Meyer, 2000; Meyer et al., 1998; Meyer & van der Meulen, 2000): The longer the utterance, the longer should be the phonological encoding processes associated with it. Accordingly, the viewing times of the target object to be named should depend on the complexity of the encoding processes. Instead, the analyses of viewing patterns and viewing times showed that during articulation, not only the target object, but also the context objects are viewed. This is in line with Eberhard (2000), who proved that when formulating a contrastive description, speakers tend to fixate the contextual alternative, instead of the target object. In addition, it may be plausible that, due to the higher affordances of the referential communication task compared to the naming tasks used by Meyer and colleagues, participants widely preplanned the object description prior to articulation. These preplanning processes might be comparable to the apprehension processes observed by Griffin & Bock (2000) in a sentence production task.

6 General Discussion

The present findings provide important insights in the perceptual and linguistic encoding processes involved in the production of complex referential expressions. In the General Discussion of the results, I will first give a concise overview of the main results obtained in the complete series of experiments and subsequently discuss these findings in view of perceptual, linguistic, and procedural aspects of referential communication.

6.1 Summary of Main Findings

In Experiment 1, I assessed the processes underlying the comparison of the target object to its contextual alternatives, using the experimental paradigm of conjunctive "same"-"different" decisions. The results provide evidence for a parallel, self-terminating comparison process. To make a "different" decision, participants looked for differences as long as necessary to detect one. As a consequence, processing times for multidimensional stimuli were determined by the difference dimension that was easiest to detect (self-terminating search), which is the dimension with the highest degree of codability: Differences in absolute and highly codable dimensions, such as color or object class, were detected faster than differences in size as a relative dimension. The detection times for color and object class differences did not differ from each other; what seemed to be relevant for the decision process was the absoluteness of the object class difference and not the fact that it was a difference in form.

The results of Experiment 1 did not provide any evidence on the evaluation of detected differences with regard to their relevance in view of higher order processes, such as referring unambiguously to multidimensional objects in a referential communication task. Reviewing the literature on "same"-"different" judgments, I evaluated investigations on the processes involved in making "same"-"different" decisions on one dimension while disregarding a second dimension. On the basis of the relevance rechecking model (Miller & Bauer, 1981; Watanabe, 1988a) and the self-terminating search and codability effects obtained in Experiment 1, I derived hypotheses on the evaluation of detected differences in a referential communication task. I predicted that due to the self-terminating search effect, size differences, co-occurring with color differences in one object, would be filtered out early in the detection process. Therefore, irrelevant differences in size should not be overspecified in a referential communication task. By contrast, irrelevant differences in color would have to be rechecked with regard to their relevance, in order not to produce an overspecified utterance. Following Whithurst's principle of least effort (see Whitehurst, 1976; Pechmann, 1989, 1994), I predicted that in these cases, it would cost less effort to produce an overspecified utterance, than to explicitly ignore the irrelevant difference in color. In a parallel assessment of the detection materials, participants overspecified the objects' color substantially more often than their size. I concluded that, apparently, referential overspecifications of color have an "early" visual, rather than a "late" linguistic origin, since they are largely attributable to mechanisms of visual discrimination. The results can be regarded as evidence for an incremental processing mode between the stages of perceptual analysis and conceptualization.

Experiment 3 was designed to assess in more detail the determinants of prenominal adjective order. The results show that perceptual grouping processes on the basis of early color information and the functional invisibility of size differences, co-occurring with color differences in one object, account for the high frequency of redundant color specifications. The data suggest that the information about color and size of the target object is available early enough for a simultaneous grammatical encoding of both dimensions under consideration of prenominal adjective ordering rules. The occasional occurrence of non-canonical adjective orders (color before size) under increased task demands is due to less effective perceptual grouping processes and time pressure. The referential noun phrase descriptions are then built up incrementally according to the order of detection of distinctive features.

6.2 Perceptual Grouping on the Basis of Early Color Information

Participants seem to group the object display with regard to color first, filtering out color discrepant objects in a very early stage of processing (Baylis & Driver, 1992; Duncan, 1980; Duncan & Humphreys, 1989; see also Cohen & Shoup, 1997; Miller & Bauer, 1981; Treisman, 1982; Treisman & Gelade, 1980). Within the remaining set of objects of one color, they then evaluate potential size differences. The early filtering of color discrepant objects leads to frequent color overspecifications (resul-ting from disregarded size differences in color- and size-discrepant context objects; see Experiments 1 and 2). Supportive evidence of grouping processes is provided by the fact that in condition SCO of Experiment 3, VT(SF) was shorter when the objects of the same color (TO and SF) appeared next to each other. Parallel results were obtained in the analysis of viewing patterns before utterance onset (see Appendix B). This is in line with previous findings that spatial proximity of objects of the same color may be a powerful grouping factor (Baylis & Driver, 1992; Fox, 1998; Han, Humphreys, & Chen, 1999; Wertheimer, 1923; see also Jiang, Olson & Chun, 2000, for the importance of spatial configuration in color visual short term memory).

In the minimal instruction group, participants had to check each difference between the TO and its context objects in view of its relevance. Perceptual grouping was rather counterproductive, as it led to an early filtering of color discrepant objects that had to be rechecked with regard to their size, in order to avoid an overspecified and thus erroneous object specification (e.g. in condition SO). The analysis of viewing patterns revealed that participants viewed each object at least once before starting to speak. Nevertheless, suppressing the early perceptual color grouping of the display seems to be rather difficult, which is apparent in the high frequency of color overspecifications and participants' descriptions of their difficulties to register more than the color difference in color- and size-discrepant context objects.

This 'functional invisibility' of size, as I termed it in the present work, is apparently closely related to attentional capture. Participants reported that they simply did not 'see' the size difference in color- and size-discrepant context objects. This is in line with findings that little or nothing is known about unattended stimuli on surprise retrospective questioning ('inattentional blindness'; see Mack & Rock, 1998; Simons, 2000), and that early perceptual grouping and segmentation may involve preattentive processes (Mack, Tang, Tuma, Kahn, & Rock, 1992; Moore & Egeth, 1997). Note, though, that "the poor knowledge shown may reflect poor explicit memory, rather than the absence of on-line processing when the unattended stimulus was presented" (Driver, Davis, Russell, Turatto, & Freeman, 2001: 67). As the experiments presented in this work do not allow any more detailed inferences as to these issues I will not go into detail on aspects of attentional capture or awareness at this point, (see Driver et al., 2001, for a concise overview). It can be concluded, though, that the initial perceptual processes of grouping and structuring the display are a dominant feature of visual perception with strong impacts on the form of the specification of the target object.

6.3 Semanto-Syntactic Principles vs. Procedural Constraints: A Trade-Off

The viewing patterns and viewing times registered before utterance onset in the neutral instruction group show that the perceptual grouping processes occur rather early in the visual exploration process. Thus, in the normal and unhindered production process, the information about both color and size should be available early enough to be incorporated in semanto-syntactic encoding processes. These encoding processes seem to be sensitive to adjective ordering rules, which will be considered in more detail in the following sections on representational principles underlying prenominal adjective order.

As the analyses of viewing patterns and viewing times in the minimal instruction group showed, participants looked at more objects of the display and took more time to evaluate the display with regard to the minimally distinctive features of the TO. Therefore, the conceptualization of the relevant adjectives took longer. In some cases, participants apparently did not take the time to evaluate the size information before initiating the syntactic encoding of the noun phrase and thus could not apply the rules of canonical adjective ordering (size before color). Therefore, they aligned the adjectives incrementally, according to the order of detection (color before size) in these cases. This accounts for the high rate of inverted adjective order phrases in the minimal instruction group.

Taken together, the data from both instruction groups provide evidence for the application of adjective ordering rules during the generation of complex noun phrases. Situational factors, such as task difficulty and time pressure, however, may force speakers to do without the ordering rules and to apply a purely incremental processing mode by mentioning the relevant dimensions in a non-canonical order according to their order of detection. Therefore, the process of generating complex noun phrases can be characterized as a trade-off between visuo-semantic and procedural constraints.

6.4 Representational Principles Underlying Canonical Adjective Order

The descriptive accounts of prenominal adjective ordering on the basis of visuo-semantic characteristics of dimensional adjectives suggest that speakers have some knowledge of adjective ordering rules. It is not clear, however, how these rules might be represented in the speakers' minds, "so, eventually, we are faced with the problem with the rules behind the rules – those rules that guide a person in utilizing rules to construct or comprehend a string of attributive modifiers" (Sichelschmidt, 1986: 146).

According to early studies in the framework of the visuo-semantic approach, perceptual factors of the object and semantic properties of the relevant differences determine the form of the object specifications (see section 2.3.2.1). Thus, the representation of ordering rules may be based on rather abstract and feature-based knowledge of the order of 'types' of (dimensional) adjectives. The ordering of color and size adjectives in complex noun phrases could be regarded as an instantiation of a more general rule that absolute features are positioned nearer to the noun than relative features (Bache 2000; Eichinger, 1991; Greenberg, 1963; Hetzron, 1978; Seiler, 1978). There are substantial similarities between the knowledge about these feature-based rules of adjective ordering and the knowledge of selection restrictions that govern the combination of verbs and their complements. Therefore, the representational basis of canonical adjective orders might be comparable to the "grammatically relevant subsystem" proposed by Pinker (1989) for the acquisition of verb meanings and the semanto-syntactic restrictions associated with some of them. Kemmerer (2000) presents evidence in favor of this representational view on the basis of selective impairments of the knowledge about the linear order of adjectives in a noun phrase on the one hand, and the knowledge about their semantic features that are invisible to syntax on the other hand.

Richards (1979) provides evidence for the existence of an internal representation of adjective rules from a developmental point of view. In a cross-sectional study with English speaking children, she found that three-year-olds display strong ordering preferences for prenominal adjective order that are given up at age four and five but eventually return at age six. Instead of focusing on the form of the object descriptions, four- and five-year-olds concentrate on the semantic contents of their descriptions, which improves considerably at that age. The acquisition of adjective ordering rules might thus proceed in steps from pure ordering imitation (at the age of three) via the acquisition of the adjectives meanings and their grammatically relevant semantic features (at age four to five) to rule-governed competence in the field of prenominal adjective ordering (at age six). Similar developmental sequences from imitation via exploration and finally rule generation to rule-based linguistic competence are present in various areas of language acquisition (see, e.g., Pinker, 1984, 1986; Shore, 1995; Tomasello, 2000).

6.5 The Role of Monitoring in Referential Communication

Pechmann (1989) tracked participants' fixation behavior during a referential communication task using more complex display types than the ones used in the present work. He found that speakers began to speak long before they had seen all objects. Thus, changes of the conceptualization of the target object may have to be made ad hoc, during articulation, on the basis of new context information. Studies that evaluated underspecified utterances and their overt repairs, such as "the red ball, the small one" (cf. Eikmeyer & Ahlsén, 1998) provide evidence for monitoring processes during the production of object specifications that control for *ambiguities* in the object specifications. It is less obvious, however, whether there are corresponding monitoring processes for *redundancies* in object specifications. Schriefers and Pechmann (1988; see also Pechmann, 1984, 1989, 1994) assume that referential overspecifications occur because of the incremental transfer of the early color information from conceptualization to formulation and that this transfer of information is not monitored in view of redundancy. Dale and Reiter (1995; pp. 249f.) assume that speakers might believe that "extra-modifiers may be helpful" for the identification of the target object (see also Mangold & Pobel, 1988).

The findings presented in this work, however, suggest yet another account: Overspecifications of color apparently originate already at the level of visual processing of multidimensional differences (see Experiment 2 and 3), and the suppression of overspecifications, as it was required in the minimal instruction group in Experiment 3, is associated with considerable effort. Therefore, it seems to be more plausible that in most cases speakers simply do not realize that the detected difference in color is irrelevant in view of a minimal description of the target object.

7 Conclusion

In the experiments presented in this work, eye movements were used to track the time course of the processes involved in the production of complex object specifications in a referential communication task. The results provide evidence for a strong perceptual influence on the form of complex object specifications. The analyses of viewing patterns and processing times even allow inferences as to the nature of these influences: Perceptual grouping processes on the basis of early color information largely determine the conceptualization of the target object description. They are apparently that rapid that they make the information on both color and size of the target object available early enough to enable the speaker to encode both dimensions according to canonical adjective ordering rules. Only in case of increased task demands, when the perceptual analysis of the object display is more complex and time-consuming, the features of the target object are encoded in the (temporal) order of their detection (color before size).

There are detailed and comprehensive models and simulations of the speech production process (Dell, 1986; Levelt, Roelofs & Meyer, 1999; Schade & Eikmeyer, 1998). However, these models do not incorporate a component for the visual and conceptual processes preceding the linguistic encoding processes, but tend to treat them as 'lead in' processes lying outside the domain of the model. Other models explicitly deal with the visual-perceptual processes involved in object recognition but do not cover the linguistic planning processes involved in generating complex utterances (Humphreys et al., 1995). The current findings provide an empirical basis for more complete accounts of speech production and its relation to (visual) perception. They open up the potential to model utterance generation processes from the early visual extraction of multidimensional features of the input over conceptualization and linguistic encoding processes to the articulation of a referring expression.

8 References

Allopenna, P.D., Magnuson, J.S., & Tanenhaus, M.K. (1998). Tracking the time course of word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419-439.

Allport, D.A. (1971). Parallel encoding within and between elementary stimulus dimensions. *Perception and Psychophysics, 10,* 104-108.

Bache, C. (2000). *Essentials of Mastering English: A Concise Grammar*. Berlin: Mouton.

Ballesteros, S., & Manga, D. (1996). The effect of variation of an irrelevant dimension on same-different visual judgments. *Acta Psychologica, 92,* 1-16.

Bamber, D. (1969). Reaction times and error rates for “same”-“different” judgments of multidimensional stimuli. *Perception and Psychophysics, 6,* 169-174.

Baylis, G.C., & Driver, J. (1992). Visual parsing and response competition: The effect of grouping factors. *Perception and Psychophysics, 51,* 145-162.

Besner, D., & Coltheart, M. (1976). Mental size scaling examined. *Memory and Cognition, 4,* 525-531.

Bindra, D., Donderi, D.C., & Nishisato, S. (1968). Decision latencies of “same” and “different” judgments. *Perception and Psychophysics, 3,* 121-130.

Blanken, G. (1998). Lexicalisation in speech production: Evidence from form-related substitutions in aphasia. *Cognitive Neuropsychology, 15,* 321-360.

Bock, J.K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89,* 1-47.

Bock, J.K. (1996). Language Production: Methods and Methodologies. *Psychonomic Bulletin and Review, 3,* 395-421.

Bock, J.K., & Levelt, W.J.M. (1994). Grammatical encoding. In: M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics.* New York: Academic Press.

Bock, J.K., Loebell, H., & Morey, R. (1992). From conceptual roles to syntactic relations: Bridging the syntactic cleft. *Psychological Review, 99,* 150-171.

Boucart, M., & Humphreys, G.W. (1992). Global shape cannot be attended without object identification. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 785-806.

Boucart, M., & Humphreys, G.W. (1994). Attention to orientation, size, luminance, and color: Attentional failure within the form domain. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 61-80.

Boucart, M., & Humphreys, G.W. (1997). Selection by color and form in vision. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 136-153.

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.

Brennen, T., David, D., Fluchaire, I., & Pellat, J. (1996). Naming faces and objects without comprehension: A case study. *Cognitive Neuropsychology, 13,* 93-110.

Brooks, B.A., Impelman, D.M., & Lum, J.T. (1981). Backward and forward masking associated with saccadic eye movements. *Perception and Psychophysics, 30,* 62-70.

Brown, C.M., & Haagort, P. (1999). *The Neurocognition of Language.* Oxford: Oxford University Press.

Brunel, N., & Ninio, J. (1997). Time to detect the difference between two images presented side by side. *Cognitive Brain Research, 5*, 273-282.

Bundesen, C., & Larsen, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 214-220.

Bush, G. (1990). *Presidential Proclamation 6158 on the Project on the Decade of the Brain.* See http://lcweb.loc.gov/loc/brain/proclaim.html.

Byrne, B. (1979). Rules of prenominal adjective order and the interpretation of “incompatible” adjective pairs. *Journal of Verbal Learning and Verbal Behavior, 18*, 73-78.

Camaioni, L., & Ercolani, A.P. (1988). The role of comparison activity in the development of referntial communication. *International Journal of Behavioral Development, 11,* 403-413.

Caramazza, A. (1998). The interpretation of semantic category-specific deficits: What do they reveal about the organization of conceptual knowledge in the brain? *Neurocase, 4,* 265-272.

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the 'tip-of-the-tongue' phenomenon. *Cognition, 64,* 309-343.

Carlson-Radvansky, L.A., & Irwin, D.E. (1995). Memory for structural information across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1441-1458.

Carter, R.C. (1982). Visual search with color. *Journal of Experimental Psychology, Human Perception and Performance, 8,* 127-136.

Cave, K.R. (1999). The FeatureGate model of visual selection. *Psychological Research, 62*, 182-194.

Cave, K.R., & Wolfe, J.M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology, 22,* 225-271.

Chertkow, H., Bub, D., & Caplan, D. (1992). Constraining theories of semantic memory processing: Evidence from dementia. *Cognitive Neuropsychology, 9,* 327-365.

Clark, E.V. (1973). What’s in a word? On the child’s acquisition of semantics in his first language. In: T.E. Moore (Ed.), *Cognitive Development and the Acquisition of Language.* New York: Academic Press.

Clark, E.V. (1978). From gesture to word: On the natural history of deixis in language acquisition. In: J.S. Bruner & A. Garton (Eds.), *Human Growth and Development.* New York: Oxford University Press.

Clark, H.H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335-359.

Clark, H.H., Carpenter, P.A., & Just, M.A. (1973). On the meeting of semantics and perception. In: W.G. Chase, (Ed.), *Visual Information Processing.* New York: Academic Press.

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

Cohen, A., & Shoup, R. (1997). Perceptual dimensional constraints in response selection processes. *Cognitive Psychology, 32,* 128-181.

Coltheart, M. (1980). Iconic memory and visual persistence. *Perception and Psychophysics, 27,* 138-227.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100,* 589-608.

Coltheart, M., Inglis, L., Cupples, L., Michie, P., Bates, A., & Budd, B. (1998). A semantic subsystem of visual attributes. *Neurocase, 4,* 353-370.

Cooper. R.M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6,* 84-107.

Corbetta, M., & Shulman, G.L. (1998). Human cortical mechanisms of visual attention during orienting and search. *Philosophical Transactions of the Royal Society of London Series B, 353,* 1353-1362.

Dale, R., & Reiter, E. (1995). Computational interpretations of the Gricean Maxims in the generation of referring expressions. *Cognitive Science, 18,* 233-263.

Damasio, A.R., McKee, J., & Damasio, H. (1979). Determinants of performance in color anomia. *Brain and Language, 7,* 74-85.

Danks, J.H., & Schwenk, M.A. (1972). Prenominal adjective order and communication context. *Journal of Verbal Learning and Verbal Behavior, 11,* 183-187.

Danks, J.H., & Glucksberg, S. (1971). Psychological scaling of adjective orders. *Journal of Verbal Learning and Verbal Behavior, 10,* 63-67.

Dell, G.S. (1985). Positive Feedback in hierarchical connectionist models: Applications to language production. *Cognitive Science, 9,* 3-23.

Dell, G.S. (1986). A spreading activation theory of lexical access in speech production. *Psychological Review, 83,* 283-321.

Dell, G.S., Chang, F., & Griffin, Z.M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science, 23,* 517-542.

Dell, G.S., Schwartz, M.F., Martin, N., Saffran, E.M., & Gagnon, D.A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104,* 801-838.

de Renzi, A., & Vignolo, L.A. (1962). Token Test: A sensitive test to detect receptive disturbances in aphasics. *Brain, 85,* 665-678.

Deubel, H., & Schneider, W.X. (1996). Saccade selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 30,* 1827-1837.

Deutsch, W. (1994). Mit dem Zeitgeist gegen den Zeitgeist – Die Anfänge des Forschungsprogramms „Objektbenennung“ [With the spirit of the age against the spirit of the age: The beginnings of the research project ‘object naming’]. In: H.-J. Kornadt, J. Grabowski, & R. Mangold-Allwinn (Eds.), *Sprache und Kognition: Perspektiven moderner Sprachpsychologie* (pp. 15-35). Heidelberg: Spektrum.

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition, 11,* 159-184.

Dixon, P., & Just, M.A. (1978). Normalization of irrelevant dimensions in stimulus comparisons. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 36-46.

Donderi, D., & Case, B. (1970). Parallel visual processing: Constant same-different decision latency with two to fourteen shapes. *Perception and Psychophysics, 8,* 373-375

Donderi, D., & Zelnicker, D. (1969). Parallel processing in visual same-different decisions. *Perception and Psychophysics, 5,* 197-200.

Downing, B.D. (1970). Response probabilities and “same-different” reaction times. *Perception and Psychophysics, 9,* 213-215.

Downing, B.D., & Gossman, J.R. (1970). Parallel processing of multidimensional stimuli. *Perception and Psychophysics, 8,* 57-60.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology, 92*, 53-78.

Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition, 80,* 61-95.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87,* 272-300.

Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review, 93,* 433-458.

Eberhard, K.M. (2000). *Watching Speakers Speak: Using Eye Movements to Study Language Production.* Invited paper presented at the 70th Annual Midwestern Psychological Association Conference, Chicago, IL (see also www.nd.edu/~keberhar/MPA/).

Eberhard, K.M., Spivey-Knowlton, M.J., Sedivy, J.C., & Tanenhaus, M.K. (1995). Eye movements as a window to real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research, 24,* 409-436.

Egeth, H.E. (1966). Parallel vs. serial processes in multidimensional stimulus discrimination. *Perception and Psychophysics, 1,* 245-252.

Eichinger, L.M. (1991). Woran man sich halten kann: Grammatik und Gedächtnis [What to adhere to: Grammar and memory]. *Jahrbuch Deutsch als Fremdsprache, 17,* 203-220.

Eikmeyer, H.-J., & Ahlsén (1998). The cognitive process of referring to an object: A comparative study of German and Swedish. In: T. Haukioja (Ed.), *Papers from the 16th Scandinavian Conference of Linguistics* (Publications of the Department of Finnish and General Linguistics of the University of Turku, vol. 60, pp. 75-86).

Eikmeyer, H.-J., & Schade, U. (1991). Sequentialization in connectionist language-production models. *Cognitive Systems, 3,* 128-138.

Eikmeyer, H.-J., Schade, U., Kupietz, M., & Laubenstein, U. (1999). A connectionist view of language production. In: R. Klabunde & C. von Stutterheim (Eds.), *Representations and Processes in Language Production* (pp. 205-236). Wiesbaden: DUV.

Entus A., & Bindra, D. (1970). Common features of the “repetition” and “same-different” effects in reaction time experiments. *Perception and Psychophysics, 7,* 143-148.

Ertel, S. (1971). Pränominale Adjektivfolgen und semantische Tiefenstruktur [Prenominal adjective sequences and semantic deep structure]. *Studia Psychologica, XIII,* 127-136.

Farell, B. (1985). “Same”-“different” judgments: A review of current controversies in perceptual comparisons. *Psychological Bulletin, 98,* 419-456.

Feldman, J. (1999). The role of objects in perceptual grouping. *Acta Psychologica, 102,* 137-163.

Ferreira, F., & Henderson, J.M. (1998). Linearization strategies during language production. *Memory and Cognition, 26,* 88-96.

Flores d’Arcais, G.B., & Schreuder, R. (1987). Semantic activation during object naming. *Psychological Research, 49,* 153-159.

Foorman, B.R. (1983). English-, Spanish-, and Japanese-speaking children’s performance on perceptual and communication tasks. *Communication and Cognition, 16,* 381-401.

Foorman, B.R., & Kinoshota, Y. (1983). Linguistic effects on children's encoding and decoding performance in Japan, the United States, and England. *Child Development, 54,* 69-77.

Ford, W., & Olson, D. (1975). The elaboration of the noun phrase in children's description of objects. *Journal of Experimental Child Psychology, 19,* 371-382.

Fox, E. (1998). Perceptual grouping and selective visual attention. *Perception and Psychophysics, 60,* 1004-1021.

Fromkin, V.A. (1971). The non-anomalous nature of anomalous utterances. *Language, 47,* 27-52.

Fussell, S.R., & Krauss, R.M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology, 62*, 378-391.

Garmiza C., & Anisfeld, M. (1976). Factors reducing the efficiency of referent-communication in children. *Merrill-Palmer Quarterly, 22,* 125-136.

Garner, W.R. (1974). *The Processing of Information and Structure.* Potomac, MD: Erlbaum.

Garner, W.R., & Felfoldy, G.L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1,* 225-241.

Garrard, P., Perry, R., & Hodges, J.R. (1997). Disorders of semantic memory. *Journal of Neurology, Neurosurgery, and Psychiatry, 62*, 431-435.

Garrett, M.F. (1975). The analysis of sentence production. In: Bower, G. (Ed.), *Psychology of Learning and Motivation, Vol. 9* (pp.177-220). Cambridge: Cambridge University Press.

Garrett, M.F. (1988). Processes in language production. In: Newmeyer, F.J. (Ed.), *Linguistics: The Cambridge Survey, Vol. III* (pp. 69-96). Cambridge: Cambridge University Press.

Gazzaniga, M.S. (Ed.) (2000). *The New Cognitive Neurosciences.* Cambridge, MA: MIT Press.

Goodglass, H. (1998). Stages of lexical retrieval. *Aphasiology, 12*, 287-298.

Gordon, R.D., & Irwin, D.E. (1996). What's in an object file? Evidence from priming studies. *Perception and Psychophysics*, 58, 1260-1277.

Gordon, R.D., & Irwin, D.E. (2000). the role of physical and conceptual properties in preserving object continuity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 136-150.

Greenberg, J.H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In: J.H. Greenberg (Ed.), *Universals of Language* (pp. 73-113). Cambridge, MA: MIT Press.

Grice, H.P. (1975). Logic and Conversation. In: P.Cole & J. Morgan (Eds.), *Speech Acts* (pp. 41-58). New York: Academic Press.

Griffin, Z.M., & Bock, J.K. (2000). What the eyes say about speaking. *Psychological Science, 11,* 274-279.

Grill, D.P. (1971). Variables influencing the mode of processing of complex stimuli. *Perception and Psychophysics, 10,* 51-57.

Han S., Humphreys, G.W., & Chen, L. (1999). Uniform connectedness and classical Gestalt principles of perceptual grouping. *Perception and Psychophysics, 61,* 661-674.

Harley, T.A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science, 8,* 191-219.

Harley, T.A. (1993). Connectionist approaches to language disorders. *Aphasiology, 7,* 221-249.

Harley, T.A. (1998). The semantic deficit in dementia: Connectionist approaches to what goes wrong in picture naming. *Aphasiology, 12,* 299-318.

Harley, T.A. (1999). Will one stage and no feedback suffice in lexicalisation? *Behavioral and Brain Sciences, 22,* 45.

Hartsuiker, R.J., Kolk, H.H.J., & Huiskamp, P. (1999). Priming word order in sentence production. *The Quarterly Journal of Experimental Psychology, 52A,* 129-147.

Haviland, E., & Lempers, J.D. (1984). Children's Classification Skills as a predictor of their referential communicative encodings. *The Journal of Genetic Psychology, 146,* 413-422.

Hawkins, H.L. (1969). Parallel processing in complex visual discrimination. *Perception and Psychophysics, 5,* 56-64.

Henderson, J.M. (1992a). Identifying objects across saccades: Effects of extrafoveal preview and flanker object context. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 521-530.

Henderson, J.M. (1992b). Visual attention and eye movement control during reading and picture viewing. In: K. Rayner (Ed.), *Eye Movements and Visual Cognition: Scene Perception and Reading* (pp. 260-283). New York: Springer.

Henderson, J.M., Dixon, P., Petersen, A., Twilley, L.C., & Ferreira, F. (1995). Evidence for the use of phonological representations during transsaccadic word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 82-97.

Henderson, J.M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 417-429.

Henderson, J.M., & Ferreira, F. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47,* 201-221.

Henderson, J.M., Pollatsek, A., & Rayner, K. (1987). Effects of foveal priming and extrafoveal preview on object identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 449-463.

Henderson, J.M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception and Psychophysics, 45,* 196-208.

Herrmann, T. (1982). *Sprechen und Situation.* Berlin: Springer.

Herrmann, T., & Deutsch, W. (1976). *Psychologie der Objektbenennung* [Psychology of object naming]. Bern: Huber.

Herrmann, T., & Grabowski, J. (1994). *Sprechen: Psychologie der Sprachproduktion* [Speaking: The psychology of speech production]. Heidelberg: Spektrum Verlag.

Hetzron, R. (1978). On the relative order of adjectives. In: H. Seiler (Ed.), *Language Universals* (pp. 165-184). Tübingen: Gunter Narr Verlag.

Hirst, G. (1999). What exactly are lexical concepts? *Behavioral and Brain Sciences, 22,* 45-46.

Hoffman, J.E. (1979). A two-stage model of visual search. *Perception and Psychophysics, 25,* 319-327.

Hoffman, J.E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics, 57,* 787-795.

Hoffmann, J., and Kämpf, U. (1985). Mechanismen der Objektbenennung – parallele Verarbeitungskaskaden [The mechanisms of object naming – parallel processing cascades]. *Sprache und Kognition, 4,* 217-230.

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

Horton, W.S., & Keysar, B. (1998). Speaking with common ground: From principles to processes in pragmatics: A reply to Polichak and Gerrig. *Cognition, 66,* 191-198.

Humphreys, G.W., Lamote, C., & Lloyd-Jones, T.J. (1995). An interactive activation approach to object processing: Effects of structural similarity, name frequency, and task in normality and pathology. *Memory, 3,* 535-586.

Humphreys, G.W., Price, C.J., & Riddoch, M.J. (1999). From objects to names: A cognitive neuroscience approach. *Psychological Research, 62,* 118-130.

Humphreys, G.W., & Riddoch, M.J. (1988). On the case of multiple semantic systems: A reply to Shallice. *Cognitive Neuropsychology, 5,* 143-150.

Humphreys, G.W., & Riddoch, M.J. (1999). Impaired development of semantic memory: Separating a semantic from structural knowledge and diagnosing a role for action in establishing stored memories for objects. *Neurocase, 5,* 519-532.

Humphreys, G.W., Riddoch, M.J., & Price, C.J. (1997). Top-down processes in object identification: Evidence from experimental psychology, neuropsychology and functional neuroanatomy. *Proceedings of the Royal Society of London / B, 352,* 1275-1282.

Humphreys, G.W., Riddoch, M.J., & Quinlan, P.T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5,* 67-103.

Hupet, M., & Chantraine, Y. (1992). Changes in repeated references: Collabration or repetition effects? *Journal of Psycholinguistic Research, 21,* 485-496.

Irwin, D.E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology, 23,* 420-456.

Irwin, D.E. (1992). Visual memory within and across fixations. In: K. Rayner (Ed.), *Eye Movements and Visual Cognition: Scene Perception and Reading* (pp. 146-165). New York: Springer.

Irwin, D.E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology, 36,* 1-27.

Jacobs, A.M., & Grainger, J. (1994). Models of visual word recognition – sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1311-1334.

Jiang, Y., Olson, I.R., & Chun, M.M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26,* 683-702.

Jolicoeur, P., & Besner, D. (1987). Additivity and interaction between size ratio and response category in the comparison of size-discrepant shapes. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 478-487.

Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329-354.

Kahneman, D., Treisman, A., & Gibbs, B.J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 175-219.

Kanwisher, N., & Driver, J. (1992). Objects, attributes, and visual attention: Which, what, and where. *Current Directions in Psychological Science, 1,* 26-31.

Kelter, S., Grötzbach, H., Freiheit, R., Höhle, B., Wutzig, S., & Diesch, E. (1984). Object identifdication: The mental representation of physical and conceptual attributes. *Memory and Cognition, 12,* 123-133.

Kemmerer, D. (2000). Selective impairment of knowledge underlying prenominal adjective ordering: Evidence for the autonomy of grammatical semantics. *Journal of Neurolinguistics, 13,* 57-82.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11,* 201-258.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14,* 185-209.

Klix, F., & Metzler, P. (1982). Structural coding of pictures in human memory and its relation to the representation of concepts. In: F. Klix, J. Hoffmann, & E. van der Meer (Eds.), *Cognitive Research in Psychology*. Amsterdam: North Holland.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35,* 1897-1916.

Kroll, J.F., & Potter, M.C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behavior, 23,* 29-66.

Krueger, L.E. (1973). Effect of stimulus frequency on speed of “same”-“different” judgments. In: S. Kornblum (Ed.), *Attention and Performance IV* (pp. 497-506). New York: Academic Press.

Krueger, L.E. (1978). A theory of perceptual matching. *Psychological Review, 85,* 278-304.

LaHeij, W., Mak, P., Sander, J., & Willeboordse, E. (1998). The gender-congruency effect in picture-word tasks. *Psychological Research, 61,* 209-219.

Lapointe, S.G., & Dell, G.S. (1989). A synthesis of some recent work in sentence production. In: G.N. Carlson & M.K. Tanenhaus (Eds.), *Linguistic Structure in Language Processing* (pp. 107-156). Dordrecht: Kluwer.

Larsen, A., & Bundesen, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 1-20.

Levelt, W.J.M. (1981). The speaker’s linearization problem. *Philosophical Transactions of the Royal Society London Series B, 295,* 305-315.

Levelt, W.J.M. (1982). Linearization in describing spatial networks. In: S. Peters & E. Saarinen (Eds.), *Processes, Beliefs, and Questions.* Dordrecht: Reidel.

Levelt, W.J.M. (1983). Monitoring and self-repair in speech. *Cognition, 14,* 41-104.

Levelt, W.J.M. (1989). *Speaking: From Intention to Articulation.* Cambridge, MA: MIT Press.

Levelt, W.J.M. (1999). Producing spoken language: A blueprint of the speaker. In: C.M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 83-122). Oxford/New York: Oxford University Press.

Levelt, W.J.M., & Meyer, A.S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology, 12,* 433-452.

Levelt, W.J.M., Praamstra, P., Meyer, A.S., Helenius, P., & Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience, 10,* 553-567.

Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1-75.

Levelt, W.J.M., Schriefers, H., Vorberg, D., Meyer, A.S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98,* 122-142.

Lindsay, R.K., & Lindsay, J.M. (1966). Reaction time and serial versus parallel information processing. *Journal of Experimental Psychology, 71,* 294-303.

Liversedge, S.P., & Findlay, J.M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4,* 6-14.

Lockhead, G.R. (1972). Processing dimensional stimuli: A note. *Psychological Review, 79,* 410-419.

Loftus, E.F. (1975). Spreading activation within semantic categories: Comments on Rosch's "Cognitive Representation of Semantic Categories". *Journal of Experimental Psychology: General, 104,* 234-240.

Logothetis, N.K., & Sheinberg, D.L. (1996). Visual object recognition. *Annual Reviews of Neuroscience, 19,* 577-621.

Lounasmaa, O.V., Hämäläinen, M., Hari, R., & Salmelin, R. (1996). Information processing in the human brain: Magnetoencephalographic approach. *Proceedings of the National Academy of Sciences of the United States of America, 93,* 8809-8815.

Mack, A., & Rock, I. (1998). *Inattentional Blindness.* Cambridge, MA: MIT Press.

Mack, A., Benyu, T., Tuma, R., Kahn, S., & Rock, I. (1992). Perceptual organization and attention. *Cognitive Psychology, 24,* 475-501.

Mangold, R. (1986). *Sensorische Faktoren beim Verstehen überspezifizierter Objektbenennungen.* Frankfurt: Peter Lang.

Mangold-Allwinn, R. (1994). Das Forschungsprogramm „Objektbenennung": Neuere Entwicklungen [The research program "object naming": Recent developments]. In: H.J. Kornadt, J., Grabowski, & R. Mangold-Allwinn (Eds.), *Sprache und Kognition: Perspektiven moderner Sprachpsychologie* (pp. 209-232). Heidelberg: Spektrum Verlag.

Mangold-Allwinn, R., Barattelli, S., Kiefer, M., & Koelbing, H.G. (1995). *Wörter für Dinge: Von flexiblen Konzepten zu variablen Objektbenennungen* [Words for things: From flexible concepts to variable object names]. Opladen: Westdeutscher Verlag.

Mangold, R., & Pobel, R. (1988). Informativeness and instrumentality in referential communication. *Journal of Language and Social Psychology, 7,* 181-191.

Mannhaupt, H.-R. (1983). Produktionsnormen für verbale Reaktionen zu 40 geläufigen Kategorien [Productive norms for verbal reactions to 40 current categories]. *Sprache und Kognition, 4,* 264-278.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* San Francisco: Freeman.

Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., & Ungerleider, L.G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science, 270,* 102-105.

Martin, J.E. (1969a). Semantic determinants of preferred adjective ordering. *Journal of Verbal Learning and Verbal Behavior, 8,* 697-704.

Martin, J.E. (1969b). Some competence-process relationships between noun phrases with prenominal and postnominal adjectives. *Journal of Verbal Learning and Verbal Behavior, 8,* 471-480.

Martin, J.E., & Molfese, D.L. (1972). Preferred adjective ordering in very young children. *Journal of Verbal Learning and Verbal Behavior, 11,* 287-292.

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81,* 899-917.

Matin, E., Shao, K.C., & Boff, K.R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception and Psychophysics, 53,* 372-380.

McCarthy, R.A., & Warrington, E.K. (1994). Disorders of semantic memory. *Philosophical Transactions of the Royal Society of London Series B, 346,* 89-96.

McConkie, G.W., & Currie, C.B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 563-581.

Meyer, A.S. (1996). Lexical access in phrase and sentence production: Results from picture-word interference experiments. *Journal of Memory and Language, 35,* 477-496.

Meyer, A.S. (1997). Conceptual influences on grammatical planning units. *Language and Cognitive Processes, 12,* 859-863.

Meyer, A.S., Sleiderink, A.M., & Levelt, W.J.M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66,* B25-B33.

Meyer, A.S., & van der Meulen, F.F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin and Review, 7,* 314-319.

Meyer, A.S., van Elswijk, G., & Tily, H. (2001). *Processing of peripheral information during multiple object naming.* Paper presented at ‘The XII Conference of the European Society for Cognitive Psychology’. Edinburgh, Scotland.

Miller, J. (1978). Multidimensional same-different judgments: Evidence against independent comparisons of dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 411-422.

Miller, J. (1982). Discrete versus continuous stage models of human information processing: In search of partial output. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 273-296.

Miller, J., & Bauer, D.W. (1981). Irrelevant differences in the “same”-“different” task. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 196-207.

Miozzo, M., & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 907-922.

Moore, C.M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 339-352.

Mordkoff, J.T., Yantis, S., & Egeth H.E. (1990). Detecting conjunctions of color and form in parallel. *Perception and Psychophysics, 48,* 157-168.

Mummery, C.J., Patterson, K., Hodges, J.R., & Price, C.J. (1998). Functional neuroanatomy of the semantic system: Divisible by what? *Journal of Cognitive Neuroscience, 10,* 766-777.

Nickerson, R.S. (1965). Response times for “same”-“different” judgments. *Perceptual and Motor Skills, 20,* 15-18.

Nickerson, R.S. (1967). “Same”-“different” response times with multi-attribute stimulus differences. *Perceptual and Motor Skills, 24,* 543-554.

Nickerson, R.S. (1969). ‘Same’-‘different’ response times: A model and a preliminary test. *Acta Psychologica, 30,* 257-275.

Nickerson, R.S. (1971). ‘Same’-‘different’ response times: A further test and a ‘counter and clock’ model. *Acta Psychologica, 35,* 112-127.

Nickerson, R.S. (1972). Binary-classification of reaction time: A review of some studies of human information-processing capabilities. *Psychonomic Monograph Supplements, 4,* 275-318.

Nickerson, R.S. (1973). Frequency, recency, and repetition effects on *same* and *different* response times. *Journal of Experimental Psychology, 101,* 330-336.

Olson, D.R. (1970). Language and thought: Aspects of a cognitive theory of semantics. *Psychological Review, 77,* 257-273.

O’Seaghdha, P.G., & Marin, J.W. (1997). Mediated semantic-phonological priming: Calling distant relatives. *Journal of Memory and Language, 36,* 226-252.

Pechmann, T. (1984). *Überspezifikation und Betonung in referentieller Kommunikation* [Overspecification and stress in referential communication]. Unpublished doctoral dissertation, Universität Mannheim, Germany.

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics, 27,* 89-110.

Pechmann, T. (1994). *Sprachproduktion: Zur Generierung komplexer Nominalphrasen* [Speech production: On generating complex noun phrases]. Opladen: Westdeutscher Verlag.

Pechmann, T., & Deutsch, W. (1982). The development of verbal and nonverbal devices for reference. *Journal of Experimental Child Psychology, 34,* 330-341.

Pechmann, T., & Zerbst, D. (1990). Zum Einfluß kognitiver Prozesse auf die Produktion komplexer Nominalphrasen [On the influence of cognitive processes on the production of complex noun phrases]. In: S.W. Felix, S. Kanngießer & G. Rickheit (Eds.). *Sprache und Wissen: Studien zur Kognitiven Linguistik.* Opladen: Westdeutscher Verlag.

Pechmann, T., & Zerbst, D. (1994). Zur allmählichen Verfertigung von Nominalphrasen beim Reden: Experimentelle Untersuchungen zur inkrementellen Sprachproduktion [On the gradual creation of noun phrases in speaking: Experimental investigations on incremental speech production]. In: S.W. Felix, C. Habel, & G. Rickheit (Eds.), *Kognitive Linguistik: Repräsentationen und Prozesse.* Opladen: Westdeutscher Verlag.

Pechmann, T., & Zerbst, D. (1995). Zum Einfluß der syntaktischen Struktur auf den Zeitverlauf von Sprachproduktionsprozessen [On the influence of syntactic structure on the time-course of speech production processes]. *Sprache und Kognition, 14*, 64-74.

Peterson, R.R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 539-557.

Pinker, S. (1984). *Language Learnability and Language Development.* Cambridge, MA: Harvard University Press.

Pinker, S. (1986). Productivity and conservatism in language acquisition. In: W. Demopoulos & A. Marras (Eds.), *Language Learning and Concept Acquisition: Foundational Issues.* Norwood, N.J.: Ablex.

Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure.* Cambridge, MA: MIT Press.

Piaget, J. (1926). *The Language and Thought of the Child.* New York: Harcourt, Brace.

Piaget, J. (1954). *The Construction of Reality in the Child.* New York: Basic Books.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56-115.

Pollatsek, A., Rayner, K., & Collins, W.E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General, 113,* 426-442.

Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3-25.

Posner, M.I. (1992). Attention as a cognitive and neural system. *Current Directions in Psychological Science, 1,* 11-14.

Price, C.J., Moore, C.J., Humphreys, G.W., Frackowiak, R.S.J., & Friston, K.J. (1996). The neural regions sustaining object recognition and naming. *Proceedings of the Royal Society of London / B, 263,* 1501-1507.

Prinzmetal, W. (1981). Principles of feature integration in visual perception. *Perception and Psychophysics*, *30*, 330-340.

Quinlan, P.T. (1991). Differing approaches to two-dimensional shape recognition. *Psychological Bulletin, 109,* 224-241.

Raaijmakers, J.G.W., Schrijnemakers, J.M.C., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41,* 416-426.

Rayner, K. (1984). Visual selection in reading, picture perception and visual search: A tutorial review. In: H. Bouma & D. Bowhuis (Eds.), *Attention and Performance*, *X*. Hillsdale, NJ: Erlbaum.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41,* 211-236.

Rayner, K. (1992). *Eye movements and visual cognition: Scene Perception and Reading.* New York: Springer.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372-422.

Remington, R.W. (1980). Attention and saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 726-744.

Richards, M.M. (1975). The pragmatic communication rule of prenominal adjective ordering: A critique. *American Journal of Psychology, 88,* 201-215.

Richards, M.M. (1977). Ordering Preferences for congruent and incongruent English adjectives in attributive and predicative contexts. *Journal of Verbal Learning and Verbal Behavior*, *16*, 489-503.

Richards, M.M. (1979). Adjective ordering in the language of young children: An experimental investigation. *Journal of Child Language, 6,* 253-277.

Robinson, G., & Cipolotti, L. (2001). The selective preservation of color naming in semantic dementia. *Neurocase, 7,* 65-75.

Roelofs, A. (1992). A spreading activation theory of lemma retrieval in speaking. *Cognition, 42,* 107-142.

Roelofs, A. (1997a). A case for nondecomposition in conceptually driven word retrieval. *Journal of Psycholinguistic Research, 26,* 33-67.

Roelofs, A. (1997b). The WEAVER model of word-form encoding in speech production. *Cognition, 64,* 249-284.

Rohlfing, K.J., Belke, E., Rehm, M., & Goecke, K.U. (submitted). Situiertheit: Das Zusammenspiel von Kontext und Situation [Situatedness: The interaction of context and situation]. *Sprache und Kognition.*

Rosch, E.H. (1973). On the internal structure of perceptual and semantic categories. In: T.E. Moore (Ed.), *Cognitive Development and the Acquisition of Language.* New York: Academic Press.

Rosch, E.H. (1975a). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104,* 192-233.

Rosch, E.H. (1975b). The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 303-322.

Rosch, E.H., & Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 8,* 573-605.

Rugg, M.D. (1999). Functional neuroimaging in cognitive neuroscience. In: C.M. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 15-36). Oxford: Oxford University Press.

Saffran, E.M., & Schwartz, M.F. (1994). Of cabbages and things: Semantic memory from a neuropsychological perspective – a tutorial review. In: C. Umiltà & M. Moscovitch (Eds.), *Attention and Performance XV* (pp. 507-536). Cambridge, MA: MIT Press.

Salmelin, R., Hari, R., Lounasmaa, O.V., & Sams, M. (1994). Dynamics of brain activation during picture naming. *Nature, 368,* 463-465.

Sanders, A.F. (1993). Performance theory and measurement through chronometric analysis. In: D.E. Meyer & S. Kornblum (Eds.), *Attention and Performance, XIV* (pp. 689-699). Cambridge, MA: MIT Press.

Santee, J.L., & Egeth, H.E. (1980). Selective attention in the speeded classification and comparison of multidimensional stimuli. *Perception and Psychophysics, 28,* 191-204.

Sartori, G., & Job, R. (1988). The oyster with four legs: A neuropsychological study on the interaction of visual and semantic information. *Cognitive Neuropsychology, 5,* 105-132.

Schade, U. (1992). *Konnektionismus: Zur Modellierung der Sprachproduktion* [Connectionism: On modelling speech production]. Opladen: Westdeutscher Verlag.

Schade, U. (1999). *Konnektionistische Sprachproduktion* [Connectionist speech production]. Wiesbaden: DUV.

Schade, U., & Eikmeyer, H.-J. (1998). Modelling the production of object specifications. In: J. Grainger & A. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Schriefers, H. (1990). Lexical and conceptual factors in the naming of relations. *Cognitive Psychology, 22,* 111-142.

Schriefers, H. (1992). Lexical access in the production of noun phrases. *Cognition, 45,* 35-54.

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 841-850.

Schriefers, H. (2000). *Grammatical Gender in the Production of Noun Phrases.* Invited talk presented at the University of Bielefeld, Germany.

Schriefers, H., de Ruiter, J.P., & Steigerwald, M. (1999). Parallelism in the production of noun phrases: Experiments and reaction time models. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25,* 702-720.

Schriefers, H., Meyer, A.S., & Levelt, W.J.M. (1990). Exploring the time course of lexical accesss in language production: Picture-word interference studies. *Journal of Memory and Language, 29,* 86-102.

Schriefers, H., & Pechmann, T. (1988). Incremental production of referential noun-phrases by human speakers. In: M. Zock & G. Sabah (Eds.), *Advances in Natural Language Generation* (pp. 172-179). London: Pinter.

Schriefers, H., & Teruel, E. (2000). Grammatical gender in noun phrase production: The gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1368-1377.

Sedivy, J.C., Tanenhaus, M.K., Chambers, C.G., & Carlson, G.N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71,* 109-147.

Seiler, H. (1978). Determination: A functional dimension for interlanguage comparison. In: H. Seiler (Ed.), *Language Universals* (pp. 301-328). Tübingen: Gunter Narr Verlag.

Sekuler, R., & Nash, D. (1972). Speed of size scaling in human vision. *Psychonomic Science, 27,* 93-94.

Seymour, P.H.K. (1973). A model for reading, naming and comparison. *British Journal of Psychology, 64,* 35-49.

Shallice, T. (1988). Specialisation within the semantic system. *Cognitive Neuropsychology, 5,* 133-142.

Shore, C.M. (1995). *Individual Differences in Language Development.* Sage Series on Individual Differences and Development, Vol. 7. Thousand Oaks, CA: Sage Publications.

Sichelschmidt, L. (1986). Optionality and choice in the analysis of adjective order: Comments on Ney. *Studia Linguistica, 40,* 135-148.

Sichelschmidt, L. (1989). *Adjektivfolgen: Eine Untersuchung zum Verstehen komplexer Nominalphrasen* [Adjective sequences: An investigation on understanding complex noun phrases]. Opladen: Westdeutscher Verlag.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric Tests for the Behavioral Sciences* (2nd ed.). New York: McGraw Hill.

Simons, D.J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences, 4,* 147-155.

Snodgrass J.G., & Townsend, J.T. (1980). Comparing parallel and serial models: Theory and implementation. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 330-334.

Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174-215.

Sobin, N.J. (1984). On code-switsching inside NP. *Applied Psycholinguistics, 5,* 293-303.

Sodian, B. (1988). Children's attribution of knowledge to the listener in a referential communication task. *Child Development, 59,* 378-385.

Sonnenschein, S. (1982). The effects of redundant communications on listeners: When more is less. *Child Development, 53,* 717-729.

Sonnenschein, S. (1984). The effects of redundant communication on listeners: Why different types may have different effects. *Journal of Psycholinguistic Research, 13*, 147-166.

Sonnenschein, S. (1985). The development of referential skills: Some situations in which speakers give redundant messages. *Journal of Psycholinguistic Research, 14,* 489-508.

Sonnenschein, S. (1986). Development of referential communication: Deciding that a message is uninformative. *Developmental Psychology, 22,* 164-168.

Sonnenschein, S. (1988). The development of referential communication: Speaking to different listeners. *Child Development, 59,* 694-702.

Stachowiak, F.J. (1978). Some universal aspects of naming as a language activity. In: H. Seiler (Ed.), *Language Universals* (pp. 207-228). Tübingen: Gunter Narr Verlag.

Tanenhaus, M.K., Magnuson, J.S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29,* 557-580.

Tanenhaus, M.K., & Spivey-Knowlton, M.J. (1996). Eye-tracking. *Language and Cognitive Processes, 11,* 583-588.

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M., & Sedivy, J.C. (1995a). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632-1634.

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M., & Sedivy, J.C. (1995b). Using Eye movements to study language comprehension: Evidence for visually mediated incremental interpretation. In: T. Inui & J.L. McClelland (Eds.), *Attention and Performance, XVI* (pp. 457-478). Cambridge, MA: MIT Press.

Terrace, H.S. (1985). In the beginning was the "name". *American Psychologist, 40,* 1011-1028.

Teyssier, J. (1968). Notes on the syntax of the adjective in modern English. *Lingua, 20,* 225-249.

Tomasello, M. (2000). The item-based nature of children’s early syntactic development. *Trends in Cognitive Sciences, 4,* 156-160.

Treisman, A. (1977). Focused attention in the perception and retrieval of multidimensional stimuli. *Perception and Psychophysics, 22,* 1-11.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 194-214.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology, 40 A,* 201-237.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97-136.

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General, 114,* 285-310.

van Turennout, M., Hagoort, P., & Brown, C.M. (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 787-806.

van Turennout, M., Hagoort, P., & Brown, C.M. (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science, 280,* 572-574.

Warrington, E.K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology, 27*, 635-657.

Watanabe, T. (1988a). Effect of irrelevant differences as a function of the relations between relevant and irrelevant dimensions in the *same-different* task. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 132-142.

Watanabe, T. (1988b). Effect of the amount of the irrelevant disparity on *different* reaction times. *Japanese Psychological Research*, *30,* 11-17.

Weiß, P., & Mangold, R. (1997). Bunt gemeint doch farblos gesagt: Wann wird die Farbe eines Objektes nicht benannt? [Meant with color, but said without: When is the color of an object not named?]. *Sprache und Kognition, 16,* 31-47.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt: II [Investigations on the theory of the Gestalt: II]. *Psychologische Forschung, 4,* 301-350.

Whitehurst, G.J. (1976). The development of communication: Changes with age and modeling. *Child Development*, *47*, 473-482.

Whitehurst, G.J., & Merkur, A.E. (1977). The development of communication: Modeling and contrast failure. *Child Development, 48,* 993-1001.

Williams, J.D. (1972). Effects of practice with controlled stimulus pairs on same-different judgments. *Journal of Experimental Psychology, 96,* 73-77.

Winer, B.J. (1971). *Statistical Principles in Experimental Design.* New York: McGraw-Hill.

Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419-433.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24,* 295-340.

Zelinsky, G., & Sheinberg, D. (1995). Why some search tasks take longer than others: Using eye movements to redefine reaction times. In: J.M. Findlay, R. Walker, & R.W. Kentridge (Eds.), *Eye Movement Research: Mechanisms, Processes, and Applications* (pp. 325-336). Amsterdam: North Holland.

Zelinsky, G., & Sheinberg, D. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 244-262.

Zifonun, G., Hoffmann, L., & Strecker, B. (1997). *Grammatik der deutschen Sprache* [Grammar of the German language]. Berlin: de Gruyter.

Zhu, M. (1995). Children’s mastering of the quantitative principle in referential communication. *Psychological Science China, 18,* 193-196.

Zorzi, M., & Vigliocco, G. (1999). Compositional semantics and the lemma dilemma. *Behavioral and Brain Sciences, 22,* 60-61.

9 Appendix

Part A – Experiment 1: Structural Differences between Viewing Patterns

Part B – Experiment 3: Statistical Analysis of Viewing Patterns

Part C – Experiment 3: Effects of Array Type

9.1 Part A – Experiment 1: Structural Differences Between Viewing Patterns

In this section, I provide additional results on processing times and reaction times associated with specific viewing patterns in Experiment 1 (cf. chapter 3). I had predicted that qualitative differences between viewing patterns in terms of complexity and extensiveness should be related to quantitative differences in viewing times for target and context object. I obtained significant effects of the factors R-Type and D-Number on both the complexity and the extensiveness of the viewing patterns. As predicted, VT(C) correlated with the complexity of the viewing patterns: The more direct viewing patterns were observed, the shorter were the viewing times of the context object (see section 3.2.3). However, the viewing times of the target object were not related to the relative extensiveness of the observed viewing patterns: Viewing patterns with a regression to the target object yielded a similar VT(T) as viewing patterns without regressions to the target object. This might be due to structural differences between the viewing patterns. Each pattern may be associated with different ways of extracting information and memorizing parts of the display in visual short-term memory (cf. Carlson-Radvansky & Irwin, 1995; Irwin, 1991, 1992; Zelinsky & Sheinberg, 1995, 1997). In the analyses described above, the viewing time of an object was defined as the sum of the first gaze at the object and all regressions to it. However, in view of structural differences within viewing patterns, analyses of individual gaze durations should be a more appropriate approach to the structural and temporal differences between viewing patterns. I assumed that more complex viewing patterns were associated with shorter gaze durations for the target object, whereas viewing patterns of less complexity should be associated with longer gaze durations.

When testing this hypothesis, I could only include the first gaze at the target object (*VT1(T)* in the following), because for the case of simple viewing patterns I did not have any data on gaze durations for the context object. I compared simple and complex viewing patterns with and without regressions with respect to VT1(T). In addition, I compared total viewing times, overall reaction times, and viewing times of the target object between pattern types. The total viewing time for simple viewing patterns was computed on the basis of the former definition of viewing times with VT(C) := 0 and VTtot = VT(T). The main types of viewing patterns (simple patterns, complex patterns with/without regressions) were coded as levels of the factor *Pattern* (3), which was analyzed in ANOVAs over each of the dependent variables VT1(T), VT(T), VTtot and RT. Except for one subject, who never applied the simple viewing pattern, all subjects could be included in the analysis of the respective patterns. For the item analysis I had to exclude 81 items that were associated with only two of the three patterns.

Table A1. Mean durations of the first gaze at the target object, mean viewing times of the target object, mean total viewing times and reaction times obtained for each viewing pattern.

	complex pattern without regression	complex pattern with one/two regressions	simple pattern
VT1(T)			
M	284	262	447
SD	118	108	233
VT(T)			
M	284	484	447
SD	118	211	233
VTtot			
M	582	772	447
SD	199	254	233
RT			
M	774	915	789
SD	218	234	244

The main effect of the factor Pattern was significant for each of the variables defined above ($F1(2,30) > 16$, $p < .001$; $F2(2,270) > 17$, $p < .001$ for VT1(T), VT(T), VTtot and RT). However, as illustrated in Figure A1, its influence on the respective variables differed (see Table A1).

The duration of the first gaze at the target object (VT1(T)) was significantly shorter for the complex patterns with regressions than for those without (see Table A2). The simple viewing patterns were associated with significantly longer gaze durations than both types of complex viewing patterns. This is in line with the assumption that the gaze durations should become shorter with increasing complexity and extensiveness of the viewing pattern. The overall viewing time of the target object (VT(T)) was significantly longer for simple patterns than for complex patterns without regression. At the same time, however, VT(T) was significantly longer for complex patterns with regressions than for those without. This is obviously due to the fact that for complex patterns with regressions, VT(T) includes the durations of the first gaze at the target object and all regressions to it.

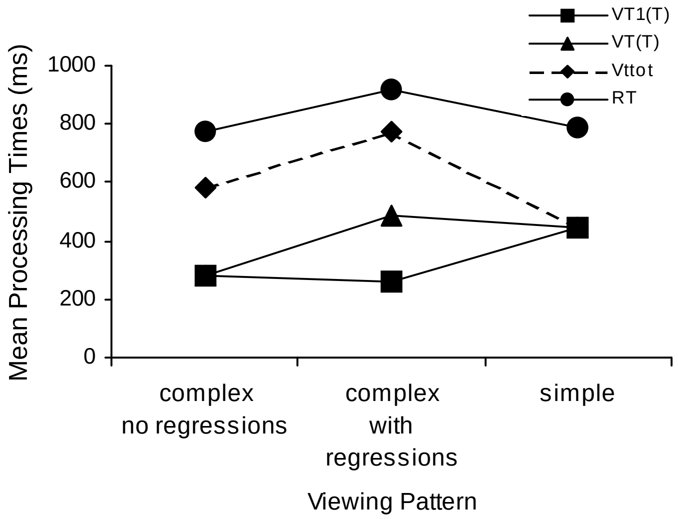


Figure A1. Mean values for the respective values VT1(T), VT(T), VTtot, and RT, displayed for each viewing pattern. Dashed lines for VTtot indicate that VTtot is no direct measure of processing time but represents the sum of VT(T) and VT(C). Note that VT1(T) and VT(T) are identical for complex patterns without regressions, as well as VT1(T), VT(T), and VTtot are identical for simple viewing patterns.

In sum, the analyses of the first gaze at the target object and its overall viewing time support the notion of structural differences between viewing patterns: The duration of the first gaze at the target became shorter the more explicit the viewing patterns got. This effect might occur because the very short first gazes at the target object were often too short to complete the processing, such that participants had to return to the target object later to extract all the information that they need. However, it may also be an effect of preplanning with regard to the complexity and extensiveness of the overall viewing pattern to follow.

Table A2. Paired comparisons between viewing patterns for the duration of the first gaze at the target object, the viewing time of the target object, total viewing time and reaction time.

	pattern without regression vs. pattern with regressions		pattern without regression vs. simple pattern		pattern with regressions vs. simple pattern	
	t1[a]	*t2*[b]	*t1*[a]	*t2*[b]	*t1*[a]	*t2*[b]
VT1(T)	2.90*	2.12*	6.06***	7.51***	6.26***	7.64***
VT(T)	12.94***	16.13***	6.06***	7.51***	1.89	2.65*
VTtot	6.81***	11.68***	6.43***	9.76***	12.03***	15.27***
RT	8.54***	7.46***	0.73	.65	5.43***	4.37***

Note. [a] $df = 15$
[b] $df = 135$
* $p < .05$; ** $p < .01$; *** $p < .001$

The longest reaction times and total viewing times were associated with the complex viewing patterns with regressions and the shortest with the simple viewing patterns (see Table A1 and Figure A1). However, the difference between the reaction times obtained for simple patterns as opposed to those for complex patterns without regression did not reach significance (see Table A2). The finding that complex viewing patterns without regressions and simple patterns did not differ in total reaction times is interesting with respect to the idea of an early response generation based on partial output of visual processing (Miller, 1982), as it implies that the time between the end of the extraction of visual information from the display and the button press was longer for the simple viewing patterns than for the complex viewing patterns, being associated with glances at *both* objects (see also Figure A1). This suggests that when using the complex viewing pattern without regressions, participants benefited from the additional information extracted during the longer and thorough exploration of the display. When finishing the exploration process, they were already apt to press the correct buttons. There is one caveat, however, that should be kept in mind when interpreting this pattern of results in terms of early response generation: The time between the end of the total viewing time and the reaction time coincided with the beginning of the processing of the second task. Subjects were told to first do the “same”-“different” decision and to decide then whether the symbol at the bottom of the screen was a plus or a cross. The earlier subjects had come to a decision with regard to the “same”-“different” judgment, the earlier they started to process the second task. If they had not pressed the button yet, the second task might have interfered with the first one and might thereby have caused longer latencies between the end of the exploration of the object display and the button press.

9.2 Part B – Experiment 3: Statistical Analysis of Viewing Patterns

9.2.1 Viewing Patterns Before Utterance Onset

As indicated in section 5.2.3.6, there was a large diversity of observed viewing patterns in Experiment 3, which could not be assessed directly for a statistic analysis. Therefore, several meta-variables were computed: For TO, CF and SF, the individual number of glances at these objects was extracted (*N(TO), N(CF)* and *N(SF)* and the overall number of glances during the exploration (*N(tot)*) was computed as their sum. Beyond that, the variable *N(obj)* was introduced to indicate how many of the three objects of the display were fixated during the exploration process.

9.2.1.1 Neutral Instruction Group

The results of the ANCOVAS over the meta-variables defined to analyze the observed viewing patterns are given in Table B1. There were significant main effects of Condition and Array for all variables. The interaction Condition x Array was significant for N(obj) and N(SF).

The main effect of Condition for N(tot), the total number of glances at the object display, was based on a significant difference between SCO, being associated with the largest number of glances at the object display, and CO with the smallest number of glances ($t1(31) = 2.98$, $p < .005$; $t2(140) = 9.01$, $p < .005$). The more detailed analysis of differences between array types (see Appendix C, Table C6) revealed significantly less glances in array types 3 and 4 than in array types 1, 2, 5, and 6 (cf. Figure B1a). This is probably due to the central position of the target object in array types 3 and 4, which makes it easier to structure the whole object display with regard to color discrepancies between the target object and the context objects on the basis of an early peripheral preview. Array types 1 and 6 were associated with significantly more glances at the object display than array types 2 and 5. As indicated in the analysis of viewing times, this may be due to the fact that in the latter array types, as well as in array types 3 and 4, the objects of the same color (TO and CF) are neighbored, whereas they are disconnected in array types 1 and 6 (see Figure 5 in section 5.2.1). This effect of neighborhood may be considered as additional evidence for perceptual grouping processes on the basis of early color information.

Table B1. Effects of Condition (C) and Array (A) for the neutral instruction group: ANOVA results for the number of glances at the objects and the number of viewed objects before utterance onset

	Subjects		Items	
	df	*F1*	*df*	*F2*
N(tot)				
C	2,31	5.43**	2,107	6.41**
A	5,79	28.30***	5,107	26.84***
C x A	10,159	.95	10,107	2.51*
N(obj)				
C	2,31	4.83**	2,107	5.67**
A	5,79	25.36***	5,107	38.29***
C x A	10,159	2.26*	10,107	2.01*
N(TO)				
C	2,31	17.02***	2,107	18.59***
A	5,79	10.65***	5,107	20.81***
C x A	10,159	1.01	10,107	.80
N(CO1)				
C	2,31	6.10**	2,107	5.94**
A	5,79	40.96***	5,107	89.20***
C x A	10,159	.71	10,107	.65
N(CO2)				
C	2,31	6.81**	2,107	7.95***
A	5,79	22.40***	5,107	54.65***
C x A	10,159	4.69***	10,107	4.49***

Note. * $p < .05$; ** $p < .01$, *** $p < .001$

The analysis of N(obj), the average number of objects of the display that are fixated during the exploration process, revealed a significantly higher number of fixated objects for condition SCO than for conditions CO ($t1(31) = 2.47$, $p < .01$; $t2(140) = 4.81$, $p < .005$) and SO ($t1(31) = 2.85$, $p < .005$; $t2(140) = 5.43$, $p < .005$). Paired comparisons between array types showed significant differences between array types 3 and 4 and the remaining array types (cf. Figure B1a and Appendix C, Table C6). As well as the corresponding finding for N(tot), this is due to the middle position of the target object in these conditions. Array types 1 and 6, with the objects of one color class being disconnected, were associated with significantly larger averages of fixated objects than array types 2 and 5, where the objects of the same color were neighbored (see above). Because of the significant interaction of Condition and Array, array types were compared separately for each condition (see Appendix C, Table C7). Effects of color neighborhood and of the relative position of the target object occurred under conditions SCO and SO only. In these conditions, array types 3 and 4 were associated with significantly smaller averages than the other array types, and array types 1 and 6 were associated with significantly larger averages than array types 2, 3, 4 and 5 and. In condition CO, there was merely an effect of the central position of the TO with significantly smaller averages in array types 3 and 4 than in the other array types. The latter, however, did not differ among each other (cf. Figure B1a).

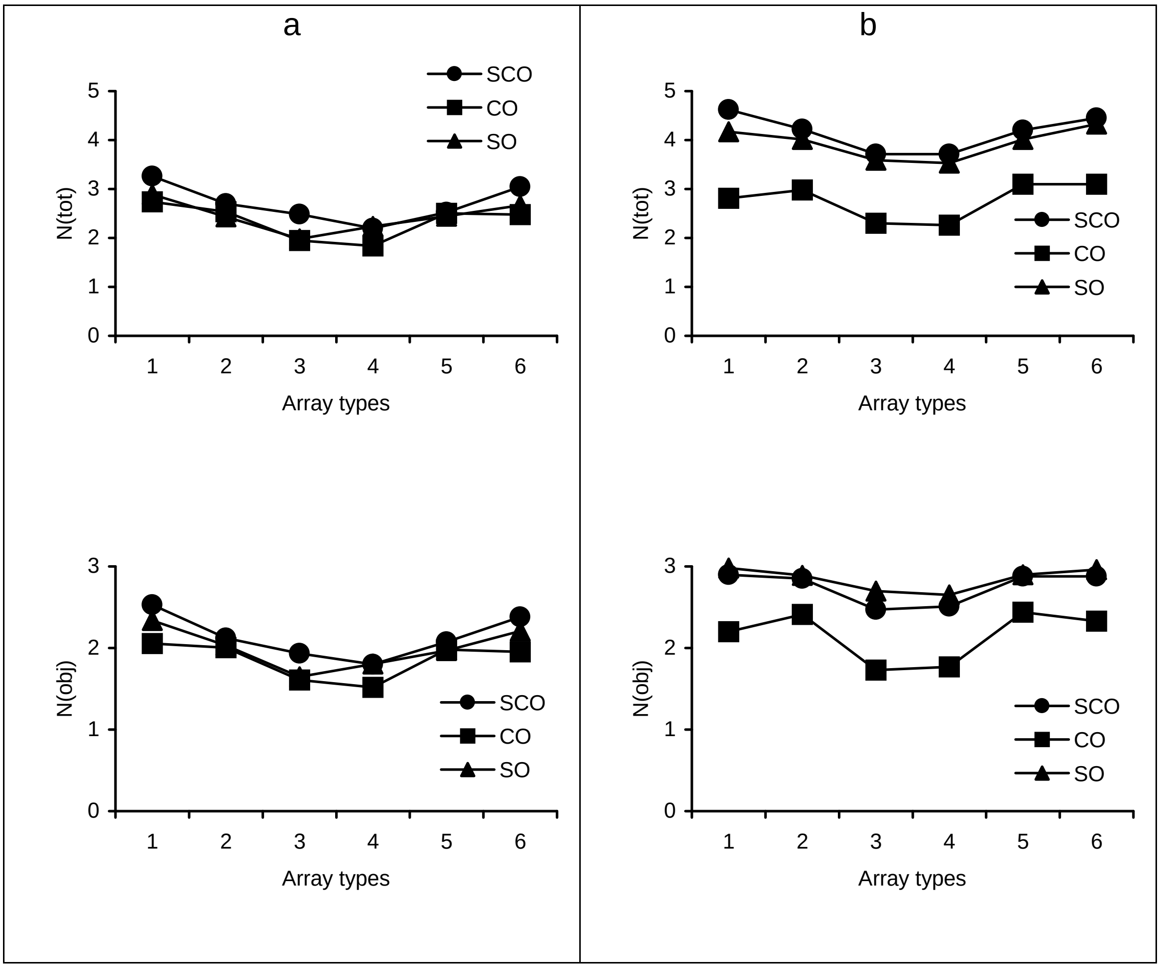


Figure B1. Mean overall number of glances and mean number of objects viewed before utterance onset by conditions and array types for the neutral instruction group (a) and the minimal instruction group (b).

N(TO), the number of glances at the TO was smallest in condition SO, differing significantly from conditions SCO (*t1*(31) = 3.84, $p < .005$; *t2*(140) = 3.10, $p < .01$) and CO (*t*(31) = 4.44, $p < .005$; *t2*(140) = 3.85, $p < .005$). The TO was looked significantly more often under condition CO than under condition SCO (*t1*(31) = 1.75, $p < .05$; *t2*(140) = 1.85, $p < .05$). As Figure B2a shows, the middle position of the TO in array types 3 and 4 did not lead to more glances at it on these positions. Instead, the effect of the factor Array is grounded in significantly fewer glances at the TO in the array types 2 to 5, as opposed to array types 1 and 6 (see also Appendix C, Table C6). This phenomenon may again be caused by the neighborhood effect: In contrast to conditions 2 and 5, the objects of the same color are disconnected in conditions 1 and 6.

The color- (and size-) discrepant CF was looked at significantly less often in condition SO than in conditions SCO (*t1*(31) = 2.37, $p < .025$; *t2*(140) = 3.84, $p < .005$) and CO (*t1*(31) = 2.64, $p < .01$; *t2*(140) = 2.98, $p < .005$). This is in line with the prediction that in condition SO, color discrepant objects are filtered out right from the start. As Figure B2a and Table C6 in Appendix C show, the CF was looked at significantly more often in array types 1 and 6, when it was positioned in the middle of the object display.

The number of glances at the size- (and color-) discrepant SF was significantly smaller under condition CO than under conditions SCO (*t1*(31) = 3.69, $p < .005$; *t2*(140) = 4.08, $p < .005$) and SO (*t1*(31) = 3.47, $p < .005$; *t2*(140) = 3.09, $p < .005$). Parallel to the findings for the CF, the main effect of Array was based on a significant increase of the number of glances at the SF, when it was in the middle position (array types 2 and 5). In addition, there were significant differences between the remaining position types: The number of glances was significantly higher in array types 1 and 6, than in array types 3 and 4 (cf. Appendix C, Table C6). This is probably due to the fact that in the latter array types the objects of the same color (TO and SF) were neighbored (array types 3 and 4) and were thus easier to group with regard to their color, than when they were disconnected (array types 1 and 6). The significant interaction of Condition and Array was investigated by means of paired comparisons between array types within individual conditions (see Appendix C, Table C7). In condition CO, the effect of Array was solely based on the significant increase of the number of glances at the SF in array types 2 and 5 with SF in the middle position. All other array types did not differ from each other. In conditions SCO and SO, in contrast, the effect of the middle position of SF in array types 2 and 5 was less pronounced. In addition, there were differences between array types 1 and 6 and array types 3 and 4 with significantly less glances at SF in conditions SCO and SO in the latter array types. Combining these findings with the analyses of the effects of Condition on the number of glances at SF, the results can be interpreted as follows: In condition CO, participants looked at the SF only seldomly, except for the cases in which it was positioned in the middle of the object alignment. Conditions SCO and SO, in contrast, included a relevant difference in size between the TO and SF. Therefore, participants looked at the SF significantly more often, independently of its relative position within the array, which accounts for the less pronounced effect of the middle position of SF in array types 2 and 5 under these conditions.

As Figure B2a shows, the target object was looked at more often than the CF and the SF, which is obviously due to the fact that the TO is the object to be named. In order to assess the differences between CF and SF with regard to the respective numbers of glances, an additional ANCOVA was run over the data extracted from CF and SF including the within-subject factors *Object* (2) and *Condition* (3). The effect of Object was significant ($F1(1,15) = 13.45$, $p < .005$; $F2(1,140) = 22.7$, $p < .001$). As the effects of condition on N(CF) and N(SF) differed (see above), Condition did not have a significant main effect in this analysis, but the interaction of Object and Condition was highly significant ($F1(2,31) = 42.57$, $p < .001$; $F2(2,140) = 14.36$, $p < .001$). Paired t-tests revealed significant differences between the number of glances at the two objects in conditions SCO ($t1(16) = 4.12$, $p < .001$; $t2(47) = 5.90$, $p < .001$) and SO ($t1(16) = 5.75$, $p < .001$; $t2(47) = 5.34$, $p < .05$). In these conditions, the CF was apparently filtered out early during the exploration process in favor of a more elaborate comparison of the SF and the TO. In condition CO, participants looked at the CF about as often as at the SF, as in both objects only the color discrepancy from the TO was relevant.

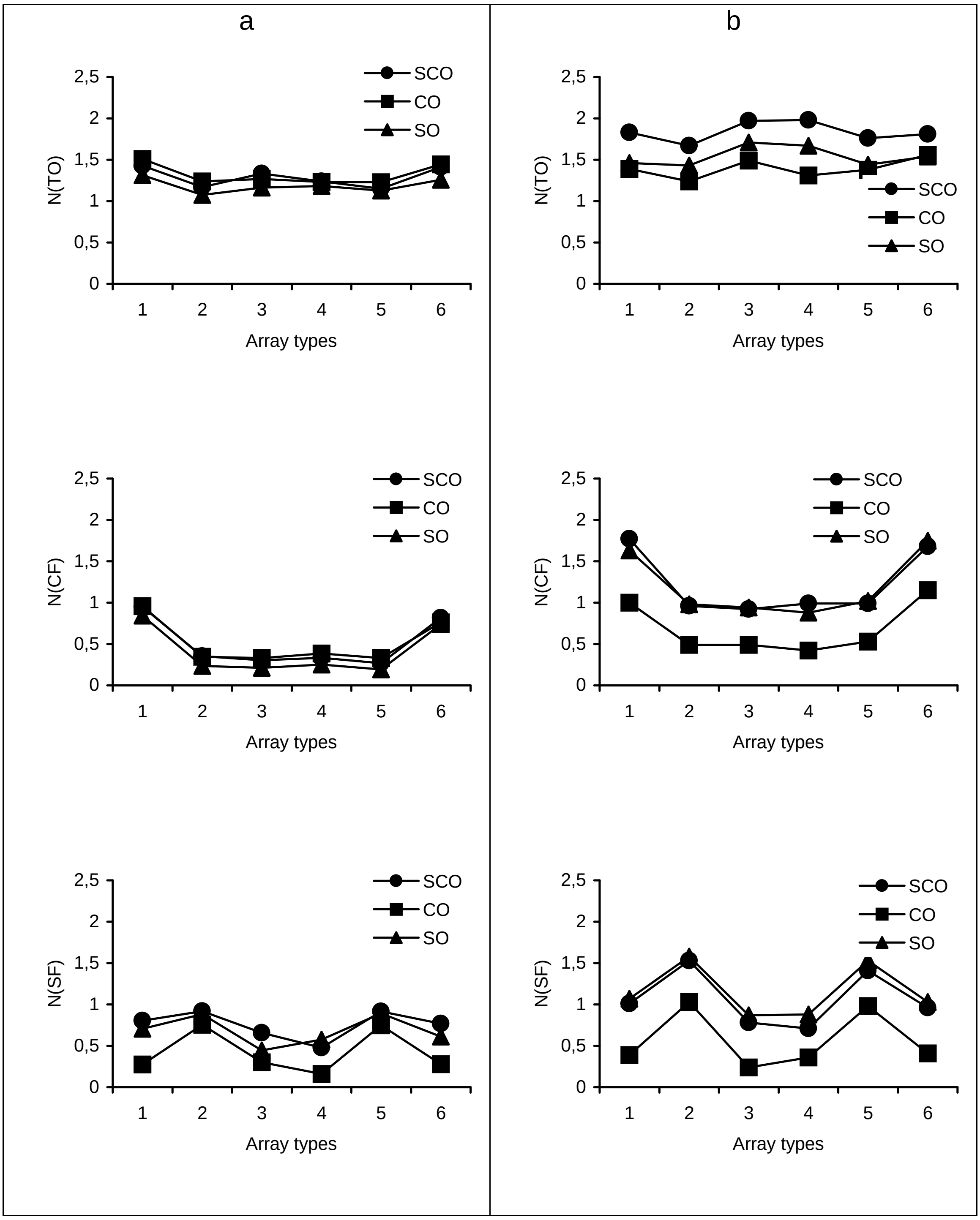


Figure B2. Mean number of glances at individual objects before utterance onset by conditions and array types for the neutral instruction group (a) and the minimal instruction group (b).

9.2.1.2 Minimal Instruction Group

ANOVAS over the meta-variables defined to analyze the observed viewing patterns for the minimal instruction group revealed significant effects of Condition and Array for all dependent variables (cf. Table B2). The interaction of these factors was significant for N(obj) only.

Table B2. Effects of Condition (C) and Array (A) for the minimal instruction group: ANOVA results for the number of glances at the objects and the number of viewed objects before utterance onset

	Subjects		Items	
	df	*F1*	*df*	*F2*
N(tot)				
C	2,30	105.31***	2,108	297.96***
A	5,75	26.54***	5,108	31.33***
C x A	10,150	1.19	10,108	1.41
N(obj)				
C	2,30	86.79***	2,108	298.11***
A	5,75	45.65***	5,108	43.34***
C x A	10,150	4.63***	10,108	3.23***
N(TO)				
C	2,30	30.90***	2,108	91.58***
A	5,75	5.48***	5,108	8.89***
C x A	10,150	1.84	10,108	2.21*
N(CF)				
C	2,30	101.90***	2,108	239.57***
A	5,75	86.28***	5,108	159.69***
C x A	10,150	1.50	10,108	2.25*
N(SF)				
C	2,30	166.22***	2,108	238.52***
A	5,75	73.76***	5,108	127.54***
C x A	10,150	.93	10,108	1.03

Note. * $p < .05$; ** $p < .01$, *** $p < .001$

As Figure B1b shows, participants needed on average 1.5 more glances at the object display in conditions SCO and SO, than in condition CO. This difference was highly significant (SCO-CO: $t1(15) = 12.04$, $p < .001$; $t2(94) = 14.65$, $p < .001$; SO-CO: $t1(15) = 13.31$, $p < .001$; $t2(94) = 13.31$, $p < .001$). N(tot) was larger under condition SCO than under condition SO ($t1(15) = 2.36$, $p < .05$; $t2(94) = 2.29$, $p < .05$), which is in part due to the fact that longer utterances had to be produced under condition SCO. The more detailed inspection of the differences between array types (see Appendix C, Table C8) revealed significantly less glances in array types 3 and 4 than in array types 1,2, 5 and 6. This effect is probably caused by the central position of the TO in array types 3 and 4, which makes it easier to judge peripherally which object might be relevant for the target object specification.

The analysis of N(obj) revealed parallel results as the analysis of N(tot) (see also Figure B1b): In condition CO, significantly less objects of the display were fixated than in conditions SCO ($t1(15) = 9.44$, $p < .001$; $t2(94) = 10.28$, $p < .001$) and SO ($t1(15) = 10.6$, $p < .001$; $t2(94) = 13.22$, $p < .001$). Conditions SCO and SO differed significantly, too ($t1(15) = 3.24$, $p < .01$; $t2(94) = 2.83$, $p < .01$). The main effect of Array was grounded in the fact that less objects were fixated in array types 3 and 4 with the TO in central position than in all other array types (cf. Appendix C, Table C8). In order to assess the significant interaction of Condition and Array for N(obj), individual analyses of the effect of Array were computed for each condition (see Appendix C, Table C9). Beyond the significant differences between array types 3 and 4 and the other array types in all conditions, SCO and SO displayed significant differences between the array types 1 and 6 on the one hand and array types 2 and 5 on the other. As outlined above, this may be due to the neighborhood effect in array types 2 to 5, which helped the perceptual grouping of color classes. In array types 3 and 4, this neighborhood effect was combined with the effect of the central position of the TO.

The number of glances at the TO was smallest in condition CO, which differed significantly from conditions SCO ($t1(15) = 6.76$, $p < .001$; $t2(94) = 11.73$, $p < .001$) and SO ($t1(15) = 3.3$, $p < .005$; $t2(94) = 4.29$, $p < .001$). The TO was fixated significantly more often under condition SCO than under condition SO ($t1(15) = 5.74$, $p < .001$; $t2(94) = 7.48$, $p < .001$). In part, this effect can be attributed to the relative length of the answers required in condition SCO: Following findings by Meyer et al. (1998), the viewing times of an object to be named correlate with the utterance length, as objects seem to be viewed, until the phonological form its name is found. The middle position of the TO in array types 3 and 4 lead to a significant increase of the number of glances at the TO (cf. Figure B2b and Appendix C, Table C8).

The CF was looked at significantly less often in condition CO than in conditions SCO ($t1(15) = 10.52$, $p < .001$; $t2(94) = 7.25$, $p < .001$) and SO ($t1(31) = 14.42$, $p < .001$; $t2(94) = 7.34$, $p < .001$). This corresponds to the prediction that the CF, being color-discrepant from the TO, was filtered out early during the exploration process. As Figure B2b and Table C8 of Appendix C show, the CF was looked at significantly more often in array types 1 and 6, when it was positioned in the middle of the object display.

Parallel to the findings for the CF, the number of glances at the SF was significantly smaller in condition CO than in conditions SCO ($t1(15) = 13.95$, $p < .001$; $t2(94) = 7.05$, $p < .001$) and SO ($t1(15) = 22.62$, $p < .001$; $t2(94) = 8.80$, $p < .001$). The latter two differed significantly for subjects, but not for items ($t1(15) = 2.29$, $p < .05$; $t2(94) = 1.47$, $p > .05$). The main effect of Array was based on a significant increase of the number of glances at the SF in array types 2 and 5, when it was in the middle position of the array (cf. Appendix C, Table C8). In addition, array types 1 and 6 differed significantly from array types 3 and 4, which may be due to the central position of the TO in the latter array types and the potential preview effects from TO on the SF in these array types.

As Figure B2b shows, N(TO) was larger than N(CF) and N(SF) in all conditions. In order to compare the CF and the SF with regard to the respective numbers of glances, an ANOVA was conducted including the within-subject factors *Object* (2) and *Condition* (3). The effect of Object was significant ($F1(1,15) = 35.86$, $p < .001$; $F2(1,141) = 5.7$, $p < .05$), as well as the effect of Condition ($F1(2,30) = 192.6$, $p < .001$; $F2(2,141) = 89.72$, $p < .001$). The interaction of these factors was not significant, as for both objects significantly less glances were registered in condition CO, than in conditions SCO and SO (see above). In all conditions, the SF was looked at significantly less often than the CF (SCO: $t1(15) = 3.44$, $p < .01$; $t2(47) = 2.32$, $p < .05$; CO: $t1(15) = 5.59$, $p < .001$; $t2(47) = 2.42$, $p < .05$; SO: $t1(15) = 2.71$, $p < .05$; $t2(47) = 2.91$, $p < .01$).

9.2.1.3 Neutral vs. Minimal Instruction Group

In an ANCOVA over both instruction groups including the factors Instruction, Condition, and Array, the differences between instruction groups were analyzed. As Table B3 shows, there were significant main effects of the factors Instruction, Condition and Array for N(tot) and N(obj). In addition, the interaction of Instruction and Condition was significant, which is due to the fact that the increase of N(tot) and N(obj) from the neutral to the minimal instruction group was more prominent in conditions SCO and SO than in condition CO. As Figure B1 shows, the exploration of the display under conditions SCO and SO was more elaborate in the minimal, than in the neutral instruction group.

Table B3. Effects of Instruction (I), Condition (C) and Array (A): ANOVA results for the total number of glances and the number of viewed objects before utterance onset

	Subjects		Items	
	df	*F1*	*df*	*F2*
N(tot)				
I	1,30	11.74***	1,107	33.34***
C	2,61	46.98***	2,107	29.49***
A	5,154	62.06***	5,107	40.90***
I x C	2,61	30.16***	2,107	138.27***
I x A	5,154	4.08**	5,107	2.93*
N(obj)				
I	1,30	11.64***	1,107	29.16***
C	2,61	24.94***	2,107	67.08***
A	5,154	49.63***	5,107	63.10***
I x C	2,61	49.14***	2,107	50.01***
I x A	5,154	1.87	5,107	2.97*

Note. * $p < .05$; ** $p < .01$, *** $p < .001$

The results of the ANOVA over N(TO), N(CF) and N(SF) revealed significant main effects for Instruction, Condition and Array and significant interactions between Instruction and Condition for all variables (cf. Table B4). For N(TO), this interaction was based on different relations among individual conditions for each instruction group: Whereas for the neutral instruction group more glances at the TO were registered under condition CO than under condition SO, the opposite occurred in the minimal instruction group.

Table B4. Effects of Instruction (I), Condition (C) and Array (A): ANOVA results for the number of glances at individual objects before utterance onset

	Subjects		Items	
	df	*F1*	*df*	*F2*
N(TO)				
I	1,30	6.09*	1,107	12.25**
C	2,61	4.49*	2,107	6.05**
A	5,154	10.04***	5,107	14.31***
I x C	2,61	29.23***	2,107	82.10***
I x A	5,154	4.95***	5,107	10.83***
N(CF)				
I	1,30	18.13***	1,107	22.62***
C	2,61	16.87***	2,107	22.24***
A	5,154	116.52***	5,107	210.60***
I x C	2,61	80.85***	2,107	149.22***
I x A	5,154	3.13*	5,107	8.06***
N(SF)				
I	1,30	9.45**	1,107	17.18***
C	2,61	94.34***	2,107	115.83***
A	5,154	84.77***	5,107	150.59***
I x A	2,61	20.60***	2,107	26.11***
I x P	5,154	5.90***	5,107	16.90***

Note. * $p < .05$; ** $p < .01$, *** $p < .001$

For N(CF) and N(SF), the interaction of Instruction and Condition was based on a significant increase of the number of glances the objects in Conditions SCO and SO in the minimal instruction group, which did not occur for condition CO. This corresponds to the prediction that because of the high saliency of the color discrepancy no additional effort had to be spent to produce a minimal specification in condition CO. The interaction of Instruction and Array was significant for N(TO) and N(SF) (cf. Table B4). This was due to the stronger effect of the middle position on the number of glances at the TO and at the SF in the minimal instruction group: As Figure B2 shows, the increase of the number of glances in array types 3 and 4 for N(TO) and array types 2 and 5 for N(SF) was more prominent in the minimal instruction group. I assume that this effect is a consequence of the more elaborate exploration of the display in the minimal instruction group: The more glances at the display are registered, the larger will be the relative increase of the number of fixations on the object in the middle position of the array.

9.2.2 Viewing Patterns During Articulation

9.2.2.1 Neutral Instruction Group

ANCOVAS of N(tot), N(obj) and N(TO) revealed neither significant effects for Condition, nor for Array and their interaction (cf. Figures B3a and B4a).

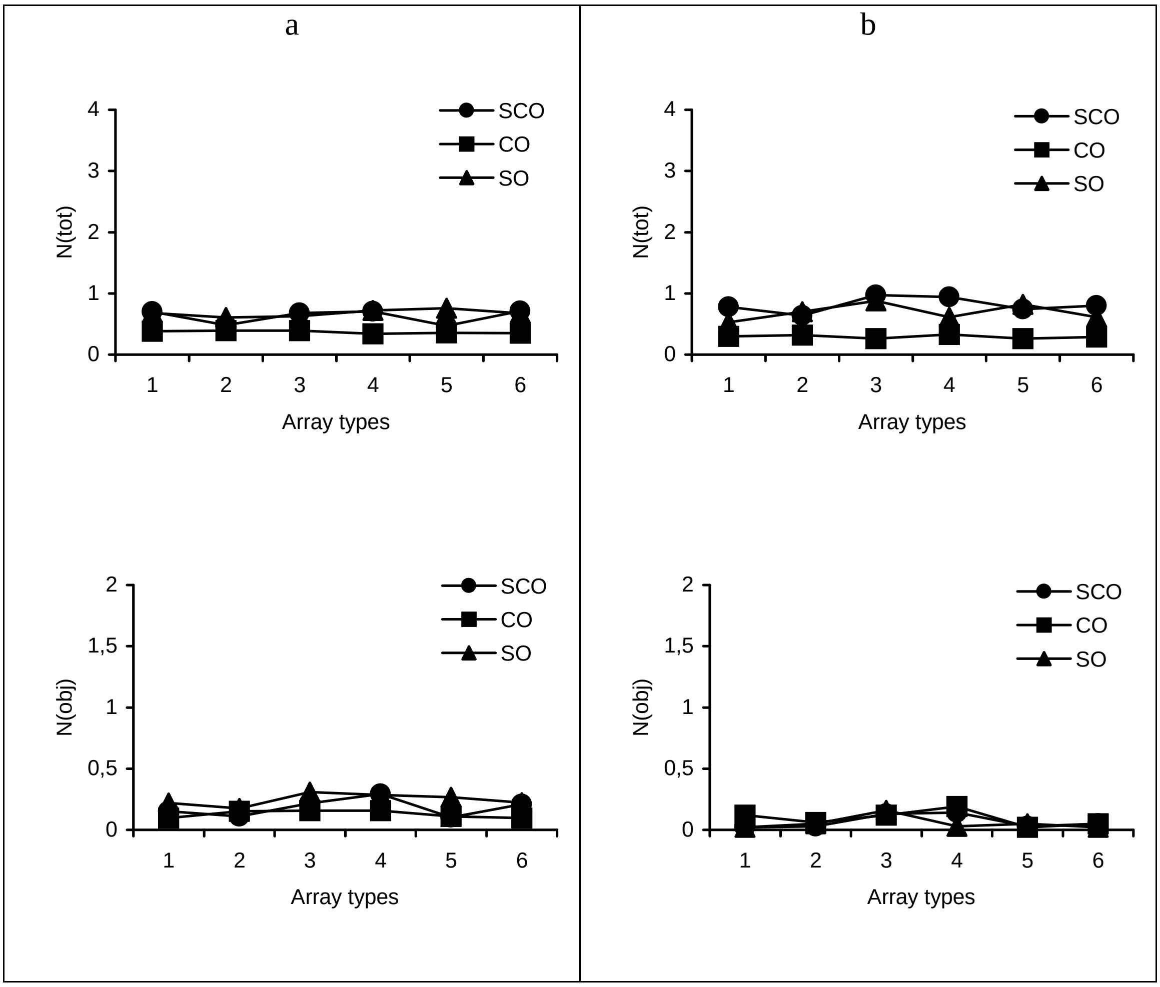


Figure B3. Mean overall number of glances and mean number of objects viewed during articulation by conditions and array types for the neutral instruction group (a) and the minimal instruction group (b).

Array was significant for N(CF) ($F1(5,79) = 5.97$, $p < .001$; $F2(5,107) = 10.63$, $p < .001$), and N(SF) ($F1(5,79) = 4.09$, $p < .005$; $F2(5,107) = 4.54$, $p < .001$). Paired comparisons between individual array types for N(CF) revealed significantly more glances at the CF under array types 1 and 6 than under array types 2, 3 and 5 (see Figure B4a and Appendix C, Table C10). These differences were based on the effects of the middle position of the CF in conditions 1 and 6.

The findings for the SF were similar (cf. Figure B4a): There was no significant effect of the middle position of the SF in array type 2, but in array type 5, leading to significantly more glances at the SF in this array type than in array types 1, 3, 4 and 6 (cf. Appendix C Table C10).

An additional ANCOVA over N(CF) and N(SF), including the factors Object and Condition, revealed a significant effect of Object ($F1(1,15) = 6.68$, $p < .05$; $F2(1,140) = 9.41$, $p < .005$) and its interaction with Condition ($F1(1,31) = 3.34$, $p < .05$; $F2(2,140) = 4.68$, $p < .05$). The main effect of Condition was not significant. Paired comparisons between the objects within individual conditions revealed significant differences between the CF and the SF in conditions SCO and SO with N(SF) being significantly larger than N(CF) (SCO: $t1(16) = 2.33$, $p < .05$; $t2(47) = 2.21$, $p < .05$; SO: $t1(16) = 2.12$; $p < .05$; $t2(47) = 2.04$, $p < .05$). There were no significant differences between N(CF) and N(SF) for condition CO (cf. Figure B4a).

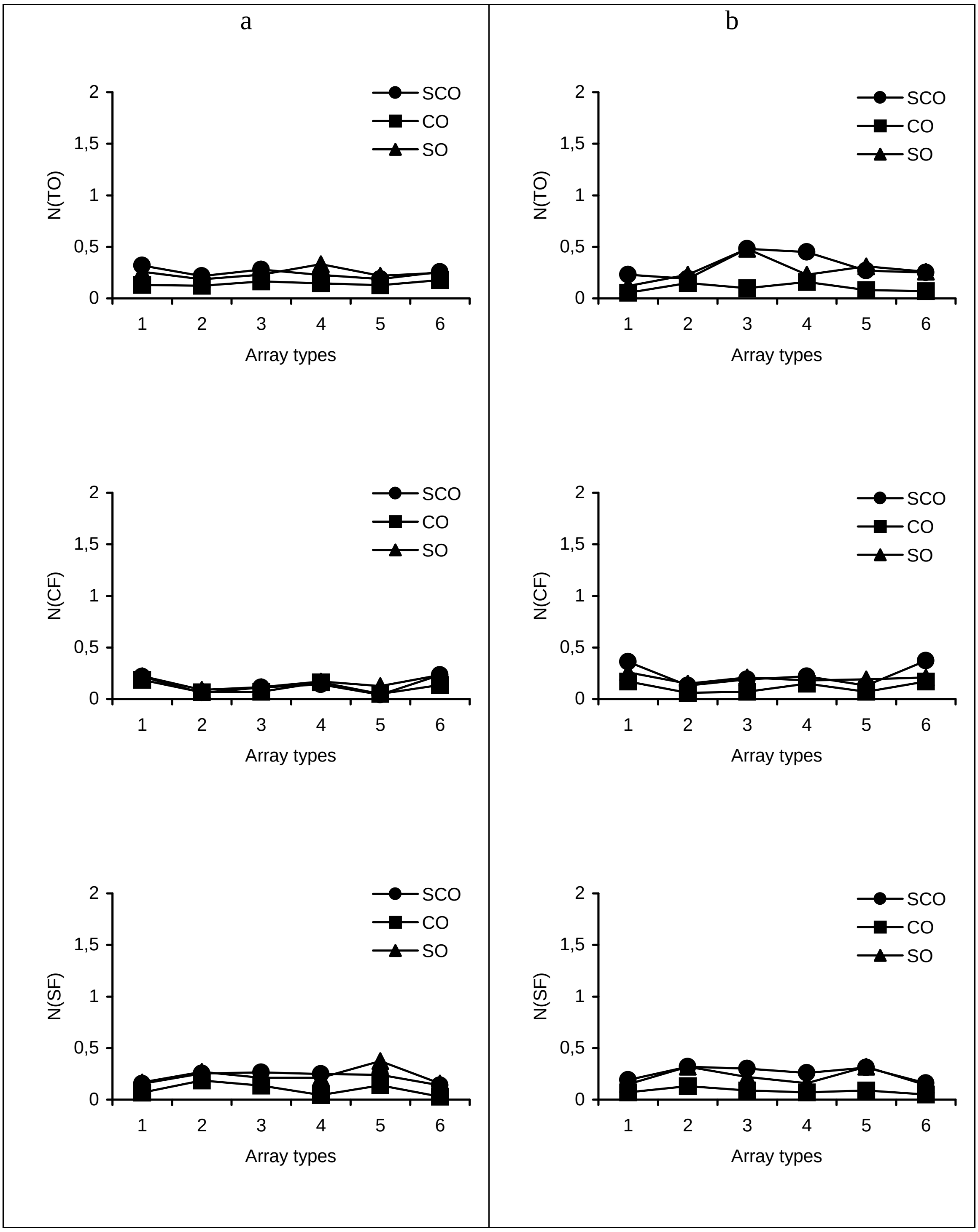


Figure B4. Mean number of glances at individual objects during articulation by conditions and array types for the neutral instruction group (a) and the minimal instruction group (b).

9.2.2.2 Minimal Instruction Group

The results of the ANOVA over the metavariables defined to analyse differences between viewing patterns are summarized in Table B5. Except for N(obj), all variables were significantly affected by Condition. Array had significant effects on all variables, except for N(tot). The interaction of Array and Condition was significant for N(TO) and N(obj).

As Figure B3b shows, the overall number of glances at the object display was significantly smaller for condition CO than for conditions SCO ($t1(15) = 4.8$, $p < .001$; $t2(94) = 12.02$, $p < .001$) and SO ($t1(15) = 6.14$, $p < .001$; $t2(94) = 8.25$, $p < .001$).

The number of objects fixated during articulation did not differ significantly between conditions (see Figure B3b). Paired comparisons between array types revealed a significant decrease of N(obj) in array types 3 and 4, when the target object was in the middle of the object display (cf. Appendix C, Table C11). In order to assess the interaction between condition and Array, separate analyses of the effects of array types were conducted for each condition (see Appendix C, Table C12): For condition SCO, N(obj) was significantly smaller in array types 3 and 4 than in all other array types (cf. Figure B3b). For condition SO, this effect is significant for array type 3 only and in condition CO the effect nearly vanishes completely.

The number of glances at the TO was significantly smaller in condition CO than in conditions SCO ($t1(15) = 3.73$, $p < .001$; $t2(94) = 8.03$, $p < .001$) and SO ($t1(15) = 5.54$, $p < .001$; $t2(94) = 6.15$, $p < .001$). The effect of Array was grounded in significantly less glances at the TO in array types 1 and 2 than in all other array types (cf. Figure B4b and Appendix C, Table C11). Neither this effect nor the results of the more detailed assessment of the interaction of Condition and Array (see Appendix C, Table C12) can be explained on the basis of the systematic construction of array types.

Table B5. Effects of Condition (C) and Array (A) for the minimal instruction group: ANOVA results for the number of glances at the objects and the number of viewed objects during articulation.

	Subjects		Items	
	df	*F1*	*df*	*F2*
N(tot)				
C	2,30	17.83***	1,141	69.31***
A	5,75	1.35	2,141	1.83
C x A	10,150	1.50	2,141	1.99*
N(obj)				
C	2,30	2.35	2,108	4.15*
A	5,75	7.85***	5,108	11.45***
C x A	10,150	2.95**	10,108	3.32***
N(TO)				
C	2,30	11.32***	1,141	41.27***
A	5,75	7.59***	2,141	8.95***
C x A	10,150	3.01**	2,141	3.23***
N(CF)				
C	2,30	11.06***	1,141	14.23***
A	5,75	4.04**	2,141	6.50***
C x A	10,150	1.35	2,141	1.49
N(SF)				
C	2,30	14.48***	1,141	42.65***
A	5,75	2.63*	2,141	7.78***
C x A	10,150	.80	2,141	1.43

Note. * $p < .05$; ** $p < .01$, *** $p < .001$

In contrast, N(CF) still displayed the effect of the middle position of the CF in array types 1 and 6 leading to a significant increase of N(CF) in these array types (Appendix C, Table C11). The CF was looked at significantly less often in condition CO than in conditions SCO ($t1(15) = 4.36$, $p < .001$; $t2(94) = 4.50$, $p < .001$) and SO ($t1(15) = 3.09$, $p < .01$; $t2(94) = 3.97$, $p < .001$). The same holds for the SF (SCO-CO: ($t1(15) = 4.76$, $p < .001$; $t2(94) = 9.37$, $p < .001$; SO-CO: $t1(15) = 4.9$, $p < .001$; $t2(94) = 5.74$, $p < .001$). There was no significant effect of the middle position of the

SF in array types 2 and 5; the significant effect of Array was rather based on unsystematic differences between some array types (see Appendix C, Table C11).

In an ANOVA over N(CF) and N(SF), the factor Object and its interaction with Condition was non-significant. The main effect of Condition, however, turned out to be significant ($F1(2,30) = 16.35$, $p < .001$; $F2(2,141) = 39.37$, $p < .001$), as in both variables, CO was associated with significantly less glances than SCO and SO.

9.2.2.3 Neutral vs. Minimal Instruction Group

The ANCOVA over both instruction groups including the between-subjects factor Instruction revealed a significant main effect of Condition for the total number of glances at the object display during articulation ($F1(2,61) = 12.44$, $p < .001$; $F2(2,107) = 32.13$, $p < .001$). For N(obj), the interaction of Instruction and Condition was significant ($F1(2,61) = 4.22$, $p < .05$; $F2(2,107) = 9.49$, $p < .001$): The increase of the mean number of viewed objects from the neutral to the minimal instruction group apparently occurred for conditions SCO and SO only, but not for condition CO (cf. Figure B3). Although Condition was significant in the separate analysis of the minimal instruction group only (see above), the analysis of the number of glances at individual objects revealed significant overall effects of Condition for N(TO) ($F1(2,61) = 7.57$, $p < .001$; $F2(2,107) = 20.21$, $p < .001$), N(CF) ($F1(2,61) = 3.94$, $p < .05$; $F2(2,107) = 4.95$, $p < .01$), and N(SF) ($F1(2,61) = 9.83$, $p < .001$; $F2(2,107) = 21.73$, $p < .001$). Array was significant for all these variables, too (N(TO): $F1(5,154) = 5.12$, $p < .001$; $F2(5,107) = 6.6$, $p < .001$; N(CF): $F1(5,154) = 9.51$, $p < .001$; $F2(5,107) = 15.25$, $p < .001$; N(SF): $F1(5,154) = 6.3$, $p < .001$; $F2(5,107) = 14.11$, $p < .001$). The interaction of Instruction and Condition turned out to be significant for N(CF) ($F1(2,61) = 3.21$, $p < .05$; $F2(2,107) = 5.69$, $p < .01$), which is due to the selective increase of N(CF) from the neutral to the minimal instruction group in conditions SCO and SO only (cf. Figure B4). In addition, N(TO) displayed a significant interaction of Instruction and Array ($F1(2,61) = 3.71$, $p < .05$; $F2(2,107) = 4.12$, $p < .01$), which is obviously caused by the differences between the neutral and the minimal instruction group in array types 3 and 4.

9.3 Part C – Experiment 3: Effects of Array Type

Table C1. Neutral instruction group: Results of paired t-tests between array types for the viewing times of each object before utterance onset

Table C2. Neutral instruction group: Detailed analyses of array types within experimental conditions for VT(TO) and VT(SF) before utterance onset

Table C3. Minimal instruction group: Results of paired t-tests between array types the viewing times of each object before utterance onset

Table C4. Neutral instruction group: Results of paired t-tests between array types for the number of glances at CO1 and CO2 and the associated viewing times during articulation

Table C5. Minimal instruction group: Results of paired t-tests between array types for the viewing times of the TO during articulation

Table C6. Neutral instruction group: Results of paired t-tests between array types for the variables defined to analyze viewing patterns observed before utterance onset

Table C7. Neutral instruction group: Detailed analyses of array types within experimental conditions for N(obj) and N(SF) before utterance onset

Table C8. Minimal instruction group: Results of paired t-tests between array types for the variables defined to analyze viewing patterns observed before utterance onset

Table C9. Minimal instruction group: Detailed analyses of array types within experimental conditions for N(Obj) before utterance onset

Table C10. Neutral instruction group: Results of paired t-tests between array types for the number of glances at CF and SF during articulation

Table C11. Minimal instruction group: Results of paired t-tests between array types for the variables defined to analyze viewing patterns observed during articulation and for the viewing times of the TO during articulation

Table C12. Minimal instruction group: Detailed analyses of array types within experimental conditions for N(Obj) and N(TO) during articulation

9.3.1 Analysis of Viewing Times

9.3.1.1 Viewing Times before Utterance Onset

9.3.1.1.1 Neutral Instruction Group

Table *C1*. Neutral instruction group: Results of paired t-tests between array types for the viewing times of each object before utterance onset

		VT(TO)[a]	VT(CF)[a]	VT(SF)[a]
1-2	*t1*[a]	1.08	6.74***	2.27**
	t2[b]	.42	12.36***	5.74***
1-3	*t1*	4.79***	7.34***	2.65***
	t2	6.37***	13.49***	2.81***
1-4	*t1*	2.31**	6.16***	1.37
	t2	5.58***	11.23***	1.36
1-5	*t1*	.95	7.67***	3.94***
	t2	.13	12.89***	8.17***
1-6	*t1*	1.35	.73	.32
	t2	.04	1.51	.26
2-3	*t1*	5.86***	.58	4.91***
	t2	5.95***	1.14	8.55***
2-4	*t1*	3.44***	.44	3.67***
	t2	5.17***	1.12	9.92***
2-5	*t1*	.13	.91	1.65
	t2	.28	.54	1.43
2-6	*t1*	.27	6.00***	2.59**
	t2	.45	9.84***	6.01***
3-4	*t1*	2.41***	1.03	1.24
	t2	1.71*	1.26	1.26
3-5	*t1*	5.72***	.32	6.57***
	t2	6.25***	.60	10.99***
3-6	*t1*	6.13***	6.58***	2.32*
	t2	6.41***	10.98***	2.55**
4-5	*t1*	3.31***	1.36	5.33***
	t2	5.45***	1.64	12.33***
4-6	*t1*	3.72***	5.55***	1.07
	t2	3.27***	9.29***	1.35
5-6	*t1*	.41	6.91***	4.25***
	t2	.17	10.39***	9.45***

Note.
t-values were computed on the basis of adjusted means (cf. Winer, 1971)
[a] $df = 79$
[b] $df = 137$
* $p < .05$;
** $p < .01$;
*** $p < .005$

Table C2. Neutral instruction group: Detailed analyses of array types within experimental conditions for VT(TO) and VT(SF) before utterance onset

		VT(TO)[a]			VT(SF)[a]		
		SCO	CO	SO	SCO	CO	SO
1-2	*t1*[a]	1.36	.63	1.08	1.61	5.77***	2.27*
	t2[b]	1.10	1.21	1.88*	1.73*	5.19***	3.22***
1-3	*t1*	3.80***	3.43***	4.79***	1.71*	.65	2.65***
	t2	3.23***	3.27***	5.02***	2.21**	.23	3.67***
1-4	*t1*	5.35***	3.61***	2.31*	3.08***	1.22	1.37
	t2	3.95***	3.45***	2.34*	3.66***	1.93*	1.56
1-5	*t1*	.75	.93	.95	2.79***	5.71***	3.94***
	t2	.68	1.12	1.64	3.19***	6.06***	5.66***
1-6	*t1*	.54	.43	1.35	.36	.31	.32
	t2	.60	1.51	1.34	.31	.01	.19
2-3	*t1*	2.43**	2.81***	5.86***	3.33***	5.11***	4.91***
	t2	2.12*	2.06*	6.91***	3.94***	5.42***	6.89***
2-4	*t1*	3.98***	2.97***	3.44***	4.71***	6.99***	3.67***
	t2	2.85***	2.24*	4.23***	5.39***	7.12***	5.88***
2-5	*t1*	.61	.29	.13	1.17	.05	1.65
	t2	.42	.09	.21	1.46	.87	2.41***
2-6	*t1*	.82	.20	.27	1.98*	5.45***	2.59**
	t2	.49	.30	.45	2.04*	5.19***	3.41***
3-4	*t1*	1.54	.17	2.41**	1.37	1.87*	1.24
	t2	.72	.18	2.66***	1.45	1.70	1.01
3-5	*t1*	3.04***	2.51**	5.72***	4.51***	5.06***	6.57***
	t2	2.55**	2.15*	6.68***	5.41***	6.29***	9.33***
3-6	*t1*	3.25***	3.00***	6.13***	1.34	.33	2.32*
	t2	2.62**	1.75*	7.36***	1.91*	.23	3.47***
4-5	*t1*	4.59***	2.67***	3.31***	5.88***	6.94***	5.33***
	t2	3.27***	2.33**	4.01***	6.86***	7.99***	8.32***
4-6	*t1*	4.81***	3.17***	3.72***	2.72***	1.54	1.07
	t2	3.38***	3.30***	3.87***	3.51***	1.07	1.46
5-6	*t1*	.21	.49	.41	3.16***	5.39***	4.25***
	t2	.07	.39	.67	3.68***	6.06***	5.86***

Note. t-values were computed on the basis of adjusted means (cf. Winer, 1971)
[a] $df = 79$
[b] $df = 41$
* $p < .05$; ** $p < .01$; *** $p < .005$

9.3.1.1.2 Minimal Instruction Group

Table C3. Minimal instruction group: Results of paired t-tests between array types the viewing times of each object before utterance onset

		VT(TO)	VT(CF)	VT(SF)
1-2	*t1*[a]	1.42	5.56***	3.72***
	t2[b]	1.14	4.98***	4.28***
1-3	*t1*	5.32***	5.95***	4.24***
	t2	4.81***	6.78***	2.76***
1-4	*t1*	5.85***	5.95***	2.87*
	t2	5.18***	5.35***	2.65*
1-5	*t1*	.96	4.59***	3.14**
	t2	1.06	5.31***	3.10**
1-6	*t1*	.57	.48	1.37
	t2	.53	.22	1.03
2-3	*t1*	5.87***	2.10	7.84***
	t2	6.38***	1.94	7.48***
2-4	*t1*	6.55***	.86	6.68***
	t2	7.13***	.35	7.55***
2-5	*t1*	.10	.33	2.00
	t2	.03	.37	1.68
2-6	*t1*	.76	7.42***	5.55***
	t2	.69	6.62***	5.76***
3-4	*t1*	.06	1.38	.24
	t2	.17	1.62	.31
3-5	*t1*	8.17***	3.85**	6.72***
	t2	5.90***	2.56**	6.84***
3-6	*t1*	7.71***	10.36***	2.26*
	t2	5.74***	9.88***	2.07*
4-5	*t1*	7.13***	.15	7.37***
	t2	6.42***	.02	6.97***
4-6	*t1*	7.67***	6.48***	4.42***
	t2	6.36***	7.27***	2.12*
5-6	*t1*	.84	8.28***	5.99***
	t2	.64	7.13***	4.77***

Note. [a] $df = 15$
[b] $df = 46$
* $p < .05$; ** $p < .01$; *** $p < .005$

9.3.1.2 Viewing Times during Articulation

9.3.1.2.1 Neutral Instruction Group

Table C4. Neutral instruction group: Results of paired t-tests between array types for VT(CF) and VT(SF) during articulation

		VT(CF)	VT(SF)
1-2	*t1*[a]	1.85*	1.22
	t2[b]	4.73***	1.97*
1-3	*t1*	2.00*	.12
	t2	4.51***	.78
1-4	*t1*	.75	.49
	t2	1.51	.63
1-5	*t1*	1.58	2.72***
	t2	2.92***	2.66***
1-6	*t1*	.18	.31
	t2	.45	.35
2-3	*t1*	.15	1.12
	t2	.21	1.18
2-4	*t1*	1.08	.74
	t2	1.91*	1.33
2-5	*t1*	.26	1.46
	t2	.18	.68
2-6	*t1*	2.03*	1.57
	t2	4.27***	2.32*
3-4	*t1*	1.23	.37
	t2	2.64**	.15
3-5	*t1*	.41	2.59**
	t2	.40	1.87*
3-6	*t1*	2.18**	.44
	t2	4.05***	1.14
4-5	*t1*	.81	2.21*
	t2	.41	2.02*
4-6	*t1*	.94	.82
	t2	1.11	2.01*
5-6	*t1*	1.76*	3.03***
	t2	3.46***	3.01***

Note.
t-values were computed on the basis of adjusted means (cf. Winer, 1971)
[a] $df = 79$
[b] $df = 137$
* $p < .05$;
** $p < .01$;
*** $p < .005$

9.3.1.2.2 Minimal Instruction Group

Table C5. Minimal instruction group: Results of paired t-tests between array types for the viewing times of the TO during articulation

		VT(TO)
1-2	*t1*[a]	1.41
	t2[b]	1.19
1-3	*t1*	3.04**
	t2	3.43**
1-4	*t1*	2.91*
	t2	2.89*
1-5	*t1*	2.71*
	t2	2.38*
1-6	*t1*	1.41
	t2	1.76
2-3	*t1*	2.61*
	t2	2.75**
2-4	*t1*	2.10
	t2	1.95
2-5	*t1*	2.00
	t2	1.66
2-6	*t1*	.97
	t2	.81
3-4	*t1*	1.18
	t2	1.19
3-5	*t1*	1.05
	t2	.98
3-6	*t1*	1.93
	t2	2.01
4-5	*t1*	.06
	t2	.10
4-6	*t1*	1.06
	t2	1.02
5-6	*t1*	.85
	t2	.95

Note. [a] $df = 15$
[b] $df = 46$
* $p < .05$; ** $p < .01$; *** $p < .005$

9.3.2 Analysis of Viewing Patterns

9.3.2.1 Viewing Patterns before Utterance Onset

9.3.2.1.1 Neutral Instruction Group

Table C6. Neutral instruction group: Results of paired t-tests between array types for the variables defined to analyze viewing patterns observed before utterance onset

		N(tot)	N(obj)	N(TO)	N(CF)	N(SF)
1-2	*t1*[a]	3.74***	3.41***	4.84***	7.79***	2.44***
	t2[b]	4.29***	4.56***	7.34***	14.36***	6.47***
1-3	*t1*	7.41***	7.68***	3.05***	8.12***	3.61***
	t2	8.76***	10.22***	4.88***	14.93***	3.03***
1-4	*t1*	5.20***	5.79***	2.59**	7.41***	1.75*
	t2	9.17***	10.68***	5.45***	13.76***	4.88***
1-5	*t1*	3.61***	4.10***	3.76***	8.35***	2.55***
	t2	4.26***	4.75***	6.67***	14.91***	6.91***
1-6	*t1*	1.80*	1.43	1.12	1.33	1.30
	t2	2.09*	1.61	1.02	1.31	.77
2-3	*t1*	3.64***	4.25***	1.80*	.31	6.03***
	t2	4.51***	5.65***	2.45**	.57	9.51***
2-4	*t1*	1.57	2.51**	2.19*	.22	4.23***
	t2	1.41	6.11***	1.89*	.59	11.36***
2-5	*t1*	.13	.69	1.09	.53	.11
	t2	.04	.18	.66	.54	.44
2-6	*t1*	1.93*	1.96*	3.72***	6.44***	3.74***
	t2	2.16*	2.65**	6.31***	11.26***	7.24***
3-4	*t1*	2.07*	1.74*	.39	.53	1.79*
	t2	3.96***	1.94*	.56	1.16	1.85*
3-5	*t1*	3.78***	3.56***	.71	.22	6.14***
	t2	4.46***	5.47***	1.71*	.02	9.95***
3-6	*t1*	5.58***	6.22***	1.91*	6.75***	2.29*
	t2	6.66***	8.31***	3.86***	11.82***	2.26*
4-5	*t1*	1.71*	1.81*	1.10	.76	4.34***
	t2	2.20*	2.84***	1.22	1.14	11.81***
4-6	*t1*	3.51***	4.47***	1.52	6.22***	.49
	t2	4.88***	5.93***	1.64	11.80***	.96
5-6	*t1*	1.80*	2.65***	2.62**	6.98***	3.85***
	t2	2.13*	2.83***	5.65**	11.81***	7.68***

Note. *t*-values were computed on the basis of adjusted means (cf. Winer, 1971)
[a] $df = 79$; [b] $df = 137$; * $p < .05$; ** $p < .01$; *** $p < .005$

Table C7. Neutral instruction group: Detailed analyses of array types within experimental conditions for N(obj) and N(SF) before utterance onset

		N(obj)			N(SF)		
		SCO	CO	SO	SCO	CO	SO
1-2	*t1*[a]	4.56***	.58	3.41***	1.51	6.31***	2.44***
	t2[b]	3.91***	1.29	3.27***	1.46	7.27	4.76***
1-3	*t1*	6.71***	5.07***	7.68***	1.94*	.34	3.61***
	t2	5.84*	4.53***	9.08***	2.06*	.23	5.11***
1-4	*t1*	8.19***	6.13***	5.79***	4.29***	1.48	1.75
	t2	6.86***	5.62***	6.96***	4.27***	2.17*	1.84
1-5	*t1*	5.12***	.85	4.11***	1.43	6.22***	2.55**
	t2	4.48***	.75	4.09***	1.12	8.21***	5.15***
1-6	*t1*	1.65	1.15	1.43	.45	.05	1.30
	t2	1.30	1.65	.35	.48	.06	.87
2-3	*t1*	2.14*	4.49***	4.25***	3.45***	5.96***	6.03***
	t2	1.93*	3.24***	5.81***	3.53***	7.04***	9.88***
2-4	*t1*	3.62***	5.56***	2.51**	5.81***	7.79***	4.23***
	t2	2.95***	4.33***	3.69***	5.74***	9.45***	7.61***
2-5	*t1*	.55	.27	.69	.07	.08	.11
	t2	.58	.53	.81	.34	.92	.38
2-6	*t1*	2.91***	.57	1.96*	1.96*	6.25***	3.74***
	t2	2.60***	.36	2.91***	1.94*	7.33***	5.64***
3-4	*t1*	1.47	1.07	1.74*	2.35*	1.83*	1.79
	t2	1.02	1.09	2.12*	2.21*	2.41**	2.27*
3-5	*t1*	1.59	4.22***	3.56***	3.37***	5.87***	6.14***
	t2	1.35	3.77***	4.99***	3.19***	7.97***	9.27***
3-6	*t1*	5.06***	3.91***	6.22***	1.48	.29	2.29*
	t2	4.53***	2.87***	8.72	1.59	.29	4.23***
4-5	*t1*	3.07***	5.29***	1.81*	5.72***	7.71***	4.34***
	t2	2.37***	4.87***	2.87***	5.39***	9.38***	8.00***
4-6	*t1*	6.54***	4.98***	4.47***	3.84***	1.53	.49
	t2	4.36***	3.69***	4.83***	3.21***	1.35	1.23
5-6	*t1*	3.47***	.30	2.65***	1.89*	6.17***	3.85***
	t2	3.18***	.89	3.73***	1.67*	8.26***	6.23***

Note. t-values were computed on the basis of adjusted means (cf. Winer, 1971)

[a] $df = 79$

[b] $df = 14$

* $p < .05$; ** $p < .01$; *** $p < .005$

9.3.2.1.2 Minimal Instruction Group

Table C8. Minimal instruction group: Results of paired t-tests between array types for the variables defined to analyze viewing patterns observed before utterance onset

		N(tot)	N(obj)	N(TO)	N(CF)	N(SF)
1-2	*t1*[a]	1.93	.22	2.60*	11.68***	8.91***
	t2[b]	.63	.08	1.72	6.78***	6.19***
1-3	*t1*	6.78***	8.05***	2.33*	11.77***	5.41***
	t2	3.15**	3.51**	2.08*	7.35***	2.41*
1-4	*t1*	5.89***	6.35***	1.24	10.46***	4.12***
	t2	2.89**	3.00**	1.08	6.87***	2.03*
1-5	*t1*	1.52	.47	.57	10.42***	8.65***
	t2	.60	.17	.68	6.57***	5.50***
1-6	*t1*	.56	.06	1.42	.70	.96
	t2	.49	.24	1.05	.65	.24
2-3	*t1*	5.49***	7.67***	3.92***	.60	10.15***
	t2	3.09**	4.05***	4.14***	.34	8.81***
2-4	*t1*	5.74***	8.03***	3.30**	.04	12.96***
	t2	2.76**	3.49***	2.63*	.19	9.03***
2-5	*t1*	.43	.57	1.74	.72	1.71
	t2	.04	.12	1.09	.33	.82
2-6	*t1*	2.87*	.54	3.81**	12.87***	12.81***
	t2	2.35*	.21	3.22**	8.80***	6.37***
3-4	*t1*	.75	1.80	.76	.44	.76
	t2	.21	.54	.68	.13	.60
3-5	*t1*	7.68***	9.89***	3.16**	1.83	11.27***
	t2	3.12**	4.11***	2.95**	.70	8.15***
3-6	*t1*	6.70***	7.71***	1.37	14.33***	2.79*
	t2	4.26***	3.92***	1.30	9.69***	2.15*
4-5	*t1*	6.49***	6.83***	2.68*	.68	10.96***
	t2	2.78***	3.55***	1.71	.52	8.32***
4-6	*t1*	6.95***	7.89***	.46	9.98***	4.23***
	t2	3.90***	3.39**	.30	8.87***	3.75***
5-6	*t1*	1.89	.34	1.91	11.63***	11.35***
	t2	1.31	.10	1.93	8.58***	5.69***

Note. [a] $df = 15$
[b] $df = 46$
* $p < .05$; ** $p < .01$; *** $p < .005$

Table C9. Minimal instruction group: Detailed analyses of array types within experimental conditions for N(Obj) before utterance onset

		N(Obj)		
		SCO	CO	SO
1-2	*t1*[a]	.63	2.53*	2.40*
	t2[b]	.64	2.44*	4.41***
1-3	*t1*	4.45***	6.73***	4.65***
	t2	6.84***	5.16***	7.21***
1-4	*t1*	6.53***	7.17***	3.21**
	t2	6.62***	2.83*	5.41***
1-5	*t1*	.29	2.98***	2.56*
	t2	.68	2.53*	2.42*
1-6	*t1*	.05	1.32	.43
	t2	.05	.72	.48
2-3	*t1*	4.85***	7.74***	2.98**
	t2	6.68***	9.28***	4.17***
2-4	*t1*	6.42***	8.45***	2.68***
	t2	6.47***	4.55***	3.24***
2-5	*t1*	.25	.14	.04
	t2	.13	.22	.33
2-6	*t1*	.72	.97	2.40*
	t2	.69	.75	3.20**
3-4	*t1*	.73	.70	.27
	t2	1.15	1.06	.02
3-5	*t1*	4.24***	7.43***	3.45**
	t2	6.11***	8.44***	3.71***
3-6	*t1*	5.11***	5.32***	4.53***
	t2	6.68***	6.99***	6.36***
4-5	*t1*	4.15***	7.56***	2.48*
	t2	5.72***	4.50***	3.08**
4-6	*t1*	5.35***	7.00***	3.12**
	t2	6.65***	3.70***	4.94***
5-6	*t1*	.25	1.18	2.40*
	t2	.72	.89	2.13*

Note. [a] $df = 15$
[b] $df = 14$
* $p < .05$; ** $p < .01$; *** $p < .005$

9.3.2.2 Viewing Patterns during Articulation

9.3.2.2.1 Neutral Instruction Group

Table C10. Neutral instruction group: Results of paired t-tests between array types for the number of glances at CF and SF during articulation

		N(CF)	N(SF)
1-2	*t1*[a]	2.09*	1.50
	t2[b]	3.24***	2.28*
1-3	*t1*	1.68*	.65
	t2	2.11*	1.51
1-4	*t1*	.81	.65
	t2	1.08	.75
1-5	*t1*	1.51	3.07***
	t2	1.14	3.09***
1-6	*t1*	.17	.13
	t2	.21	1.09
2-3	*t1*	.41	.84
	t2	1.13	.97
2-4	*t1*	1.27	.83
	t2	1.15	1.72*
2-5	*t1*	.58	1.56
	t2	.11	.61
2-6	*t1*	2.26*	1.63
	t2	2.04*	1.58
3-4	*t1*	.85	.01
	t2	1.02	.75
3-5	*t1*	.17	2.41**
	t2	.02	2.58**
3-6	*t1*	1.85*	.78
	t2	1.90*	1.61
4-5	*t1*	.68	2.39**
	t2	1.05	2.34*
4-6	*t1*	.99	.80
	t2	1.39*	1.19
5-6	*t1*	1.67*	3.20***
	t2	1.93*	4.20***

Note. [a] $df = 15$
[b] $df = 137$
* $p < .05$; ** $p < .01$; *** $p < .005$

9.3.2.2.2 Minimal Instruction Group

Table C11. Minimal instruction group: Results of paired t-tests between array types for the variables defined to analyze viewing patterns observed during articulation and for the viewing times of the TO during articulation

		N(Obj)	N(TO)	N(CF)	N(SF)
1-2	*t1*[a]	.28	2.01	3.77**	2.26*
	t2[b]	.43	1.66	3.81***	3.30***
1-3	*t1*	2.85*	4.28***	2.81*	2.37*
	t2	4.04***	3.85***	2.85**	2.68*
1-4	*t1*	2.68*	4.02***	2.63*	1.39
	t2	2.67*	3.50***	2.16*	1.43
1-5	*t1*	.79	3.17**	4.08***	1.98
	t2	.76	2.04*	3.75***	1.79
1-6	*t1*	.93	1.45	.11	.82
	t2	1.05	1.26	.01	.42
2-3	*t1*	3.43**	2.80*	.97	.98
	t2	4.63***	2.65*	1.11	.86
2-4	*t1*	2.80*	2.58*	2.39*	1.63
	t2	3.06**	2.11*	2.06*	2.00
2-5	*t1*	.62	1.04	.28	.81
	t2	.33	.53	.11	.42
2-6	*t1*	.67	.09	2.19*	2.63*
	t2	.64	.41	3.08**	3.57***
3-4	*t1*	.43	.74	.88	.91
	t2	.54	.73	.90	1.23
3-5	*t1*	3.93***	2.18*	1.20	.45
	t2	5.14***	2.12*	1.16	.41
3-6	*t1*	5.97***	3.31**	2.53*	3.04**
	t2	5.43***	2.96**	2.31*	2.99**
4-5	*t1*	3.20**	1.77	1.90	1.22
	t2	3.37***	1.52	2.05*	1.52
4-6	*t1*	3.74***	2.60*	1.27	2.39*
	t2	3.58***	2.47*	1.74	2.77*
5-6	*t1*	.25	.75	3.09**	2.64*
	t2	.33	.92	3.08***	3.07**

Note. [a] $df = 15$
[b] $df = 46$
* $p < .05$; ** $p < .01$; *** $p < .005$

Table C12. Minimal instruction group: Detailed analyses of array types within experimental conditions for N(Obj) and N(TO) during articulation

		N(Obj)			N(TO)		
		SCO	CO	SO	SCO	CO	SO
1-2	*t1*[a]	.24	1.65	1.63	.46	1.60	2.57*
	t2[b]	.31	1.83	1.66	.50	1.56	2.16*
1-3	*t1*	2.09*	.14	3.87**	2.55*	.69	4.14***
	t2	3.31***	.08	4.41***	3.73***	.28	5.05***
1-4	*t1*	2.36*	1.28	1.05	3.90***	1.60	2.57*
	t2	2.58*	1.84	.97	3.30***	1.95	2.31*
1-5	*t1*	.23	2.20*	1.38	1.11	.31	3.04***
	t2	.57	3.12**	1.41	1.09	.13	2.59*
1-6	*t1*	.50	2.04	.45	.51	.21	4.29***
	t2	.90	2.08	.43	.94	.22	3.86***
2-3	*t1*	2.38*	1.33	2.68*	2.74*	1.17	2.54*
	t2	3.53***	1.86	2.79*	4.03***	1.32	3.01**
2-4	*t1*	2.90*	2.61*	1.11	3.98***	.01	.40
	t2	2.74*	3.36***	.64	3.58***	.12	.48
2-5	*t1*	.42	1.08	.36	1.87	1.17	1.67
	t2	.85	1.06	.30	1.50	1.52	.75
2-6	*t1*	.66	.61	2.55*	.86	1.58	.27
	t2	1.20	.28	2.17*	1.47	1.70	.23
3-4	*t1*	.45	.18	3.12**	.28	1.24	2.59*
	t2	.28	1.97	3.39***	.06	1.67	2.34*
3-5	*t1*	2.49*	.75	3.07**	1.83	.33	1.73
	t2	2.65*	2.12	3.11***	2.57**	.17	2.26*
3-6	*t1*	2.48*	2.35*	4.02***	1.70	.52	2.74*
	t2	2.62*	3.28**	4.66***	3.55**	.49	3.58***
4-5	*t1*	2.63*	2.90*	.55	2.50*	1.38	.68
	t2	2.17*	4.44***	.36	2.32*	1.95	.22
4-6	*t1*	2.22*	2.46*	1.67	2.31*	1.59	.28
	t2	2.21*	3.55***	1.29	3.04**	2.09	.74
5-6	*t1*	.42	.47	2.01	.18	.11	1.09
	t2	.24	.71	1.74	.48	.36	1.08

Note. [a] $df = 15$
[b] $df = 14$
* $p < .05$; ** $p < .01$; *** $p < .005$