MPI SERIES
IN PSYCHOLINGUISTICS

LEXICAL ACCESS IN THE PRODUCTION OF ELLIPSIS AND PRONOUNS

Bernadette M. Schmitt

LEXICAL ACCESS IN THE PRODUCTION OF ELLIPSIS AND PRONOUNS

MPI SERIES IN PSYCHOLINGUISTICS

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing
   Miranda van Turennout

2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography
   Niels O. Schiller

3. Lexical access in the production of ellipsis and pronouns
   Bernadette M. Schmitt

ISBN: 90-76203-03-2
Design: Linda van den Akker, Nijmegen; Inge Doehring, Nijmegen
Illustration: After © 1994, Rüdiger Gay, in L. Juchli (Ed.). *Pflege*. Georg Thieme Verlag, Stuttgart. Reprinted with permission.
Printed and bound by: Ponsen & Looijen bv, Wageningen

LEXICAL ACCESS IN THE PRODUCTION OF ELLIPSIS AND PRONOUNS

een wetenschappelijke proeve
op het gebied van de Sociale Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit Nijmegen,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op
maandag 1 december 1997
des namiddags om 1.30 uur precies

door

Bernadette Maria Schmitt

geboren op 24 juni 1965 te Wolfsburg, Duitsland

Promotor: Prof. dr. W. J. M. Levelt
Co-promotores: Dr. A. S. Meyer
Dr. R. Goebel (Max Planck Institut für Hirnforschung, Frankfurt)

Manuscriptcommissie:
Dr. A. Roelofs
Prof. dr. H. Schriefers
Prof. dr. P. Zwitserlood (Westfälische Wilhelms-Universität, Münster)

The research reported in this thesis was supported by a grant from the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany.

DANKE

Vielen habe ich vieles zu verdanken. Am Max Planck Institut für Psycholinguistik in Nijmegen haben vorallem die Mitglieder der Projektgruppe *Lexical Access in speech production* mit ihren lebendigen Ideen den Inhalt dieser Arbeit mitbestimmt. Besonderer Dank geht dabei an meine Begleiter, Pim Levelt und Antje Meyer. Obwohl das Betreuen meiner Arbeit nur eine von 1001 ihrer Aufgaben war, hatten sie immer Zeit, um zu hören, und Ruhe, um mitzudenken. Dabei spielte die Gesprächssituation kaum eine Rolle: Wir haben Ellipsen diskutiert bei einem von Hans Gerrits oder Theas leckeren Käsebrötchen, Pronomen beim Abfischen des MPI-Teiches, und Netzwerke zwischen Farbtöpfen in Pims Büro.

Wichtig und variantenreich waren auch die Treffen mit meinen Mit-PhD-Studenten. Gespräche in PhD-Luncheons, in abendlichen Leseclubs und beim Tanz durch die Nacht waren sowohl fachlich als auch freundschaftlich bedeutend für mich. Ich danke allen, ganz besonders jedoch Angelika Wittek, Arie van der Lugt, Jos van Berkum, Kay Behnke, meinen Paranimfen David van Kuijk (und seiner Band *Love kills*) und Niels Schiller, und Miranda van Turennout - für ihre Freundschaft.

Das Datensammeln erfolgte sowohl in Nijmegen als auch in Münster. Für alle Experimente wurden hunderte wunderschöner Bilder von Inge Doehring und Godelieve Hofstee vorbereitet und mit eigenen Zeichnungen verknüpft - vielen Dank für die gute Zusammenarbeit! Die akustischen ‘Störwörter’ wurden gesprochen durch Edith Sjoerdsma, Agnes Bolwiender und Niels Schiller. Nach wochenlangem Hören der Aufnahmen sind mir ihre Stimmen wahrscheinlich für ewig ins Gedächtnis geprägt. Ich hoffe jedoch, daß ich sie auch weiterhin hören darf.

Die deutschen Experimente wären nicht so einfach durchführbar gewesen ohne ein festes ‘Lab’ in Deutschland. Pienie Zwitserlood und Jens Bölte haben mir dafür freundlicherweise ein Jahr lang einen Raum in der Allgemeinen Psychologie an der Uni Münster bereit gestellt. Vielen Dank - auch für die guten Diskussionen und die vorzüglichen Abendessen.

Beim Sammeln der Daten haben mir in Nijmegen Godelieve Hofstee und Marianne Wissink geholfen. In Münster wurde ich durch Anne Wülkner, Maurus Bannert und Sandra Muckels unterstützt. Ich bin ihnen sehr für ihr Organisationstalent dankbar, und für die Geduld, die sie beim Durchführen einiger ziehmlich langer Experimente gezeigt haben. Ganz besonders danke ich meiner Schwester Marion und ihrer WG: sie haben mir in Münster ein Zuhause gegeben.

Tausend Dank (da ungefähr soviele Telefonate nötig waren) geht an Paul Verleg und Sylvia Aal, die mir geholfen haben, Versuchspersonen zu organisieren. Genauso viel Dank geht an alle Kolleginnen und Kollegen der Verwaltung.

Das Experimentieren in Nijmegen und in der ‘Außenstelle’ Münster wäre ohne die Hilfe der Technischen Gruppe kaum möglich gewesen. Großer Dank geht an alle TGler, besonders aber an Ad Verbunt, John Nagengast und Reiner Dirksmeyer. Sie haben gezeigt, daß Problemlösen - auch über größere Distanzen hinweg - im Zeitalter der elektronischen Kommunikations-Netzwerke kein Problem mehr ist.

Der Begriff 'Netzwerke' ist für mich automatisch verbunden mit Arbeitstreffen in Frankfurt bei Rainer Goebel. Rainer, als ein unglaublich guter Lehrer und Ideen-Generator, hat meine 'neuronalen' Strukturen so programmiert, daß ich sicher in einem Bild-Wort-Interferrenz Experiment einen phonologischen Effekt auf 'Netz' oder 'Werk' produziere, wenn ich ein Bild vom Frankfurter Funkturm benennen müßte.

Beim Schreiben dieser Arbeit hatte ich als *master of english agreement errors* die Hilfe von *native speakers* dringend nötig. Ich danke Dale Blue, Laura Walsh-Dickey, Marry Swift und Mike Dickey für das Korrekturlesen. Außerdem geht ein großer Dank an Astrid Sleiderink, die für mich die Zusammenfassung ins Niederländische übersetzt hat- und an Edith Sjoerdsma, für das Kontrollieren der Referenzen. James McQueen, Kay Behnke und Wolfgang Klein haben mir sehr geholfen durch ihre guten inhaltlichen Kommentare zu früheren Versionen dieser Arbeit. Als Fachmann für Methodik (und Single Malt) kann ich Sven Blankenberger nur weiterempfehlen. Für die graphische Gestaltung des Buches danke ich Linda van den Akker und Inge Doehring sehr.

Es war eine wichtige Zeit für mich hier in Nijmegen, und ich werde viele Menschen sehr vermissen, wenn ich gehe. Allerdings freue ich mich auch schon auf die Zukunft, deren 'wissenschaftliche' Seite durch Colin Brown und Marta Kutas mitgeplant wurde. Der Dank an sie folgt sicher später.

Keine Zukunft ohne Vergangenheit... und daher möchte ich hier auch Danke sagen an Werner Deutsch, der mir das MPI und 'Speaking' während meines Studiums gezeigt hat - und an Dirk Vorberg, für seine faszinierende Art zu vermitteln, wie wichtig 'gute' Experimente in der Kognitiven Psychologie sind.

Schließlich - und ganz besonders - möchte ich meinen Eltern und Geschwistern danken für ihre Unterstützung, die sie mir während dieses PHD-Projektes und überhaupt gegeben haben. Sie haben mich durch ihre gute Art dazu ermutigt, diese Arbeit zu beginnen und abzuschließen. Danken möchte ich auch Jari Nickel - der sicher ebenso froh ist wie ich, daß diese Arbeit nun geschrieben ist.

Eine Szene im Kolwigsgarten von Nordsteimke...
of in de Ooijpolder te Nijmegen...
or at the beach in San Diego.

MEINEN ELTERN

TABLE OF CONTENTS

CHAPTER 1

Speaking is the capacity to turn thoughts into a structured stream of sound. This transformation is a highly complex process which is determined by various factors. Among these are the inherent properties of the human speech processing system, such as the nature of the human brain, the nervous system, and the articulatory organs. Another important factor is the speaker's linguistic knowledge: The speaker must know the rules of a particular language and apply them in the production process in such a way that a listener can understand the meaning of the sound.

One important component of the thoughts-to-sound transformation is the "word". The word is an arbitrary match between the preverbal meaning and the sound. The word *butterfly* does not look like a butterfly or sound like one, but it gives access to the stored knowledge about this animal. The access can be seen as looking up available information about the butterfly in a mental lexicon. The information can be semantic: For example, the word could remind you of a good friend who suffered from a butterfly phobia. The stored knowledge can be syntactic: The word has the lexical entry of being a noun in some cases, and it can be used as a verb in another case, as in *She butterflied extremely fast* (to swim in a particular style). The lexical information can also be morphological in nature: For example, the information might become available that *butterfly* is a combination of two words, *butter* and *fly*.

A second important component of language processing is the rule-based combination of words, the grammar. Commonly, three types of grammatical knowledge are distinguished: Syntactic rules, morphological rules, and phonological rules. Syntactic rules determine the construction of phrases and sentences. Word order, case assignment, or subject-verb agreement are typical examples. The word order *The dog bites the man* means something different than *The man bites the dog*. The speaker must have knowledge about the different semantic roles a verb assigns to the subject and the object of the sentence. In the utterance *The dog bites him* the speaker knows that the verb syntactically assigns accusative case to the object and uses the pronoun *him* instead of *he*. Morphological rules specify how words and parts of words can be put together. For example, the speaker knows that the verb *to bite*, when it is to express present tense, third person, singular, must have an 's' as ending.

In addition to syntactic and morphological processing the speaker follows phonological rules. Phonology involves rules for the segmental and prosodic build-up of a word. Segmentation principles tell the speaker what is a possible

English word, such as *dog*, or an impossible word, such as *dgo*. Prosodic principles create metrical and intonational patterns. For example, the prosodic structure of the question *The man bites the dog?* is different from the statement *The man bites the dog*. The final step in the thoughts-to-sound transformation is the articulation of the message. The mentally planned words and sentences are converted into motor commands for speaking.
Linguists have made numerous proposals on how these rules of speaking should be characterised, either within a particular language or even across languages. They also discuss how the various components of speaking are interrelated to each other. But normally, linguists do not consider the processing side of these rules. They only state that processing components must somehow interact, and they try to characterise the principles according to which the components interact. Linguists would not consider, for example, at which point in time a certain lexical item is "activated", at what point in time a syntactical rule is applied to two lexical item, or whether morphological rules are applied before or after syntactic rules, and so on. The time course of the retrieval of various types of information and the application of various types of rules is a matter of psycholinguistic research. Some aspects of this process will be explored in the following chapters.

As a rule, speaking is a highly automatic process. Whereas language researchers are fascinated, and sometimes perplexed, by the complexity of the language system and the way in which it works in communication, the everyday language user normally is not. The speaker is not aware of the planning process, except perhaps when it runs into trouble or even breaks down. It is this automaticity which makes speaking fluent. A speaker, at a normal speech rate, produces about 150 words per minute (Maclay & Osgood, 1959) or 5 to 6 syllables in one second (Deese, 1984). Errors are surprisingly rare. Deese (1984) counted only 77 syntactic anomalies in a tape recorded corpus of nearly 15,000 utterances. This is roughly one error in every 200 utterances. Selecting the wrong lexical entry happens less than once per thousand words (Deese, 1984). This fluency also shows that speakers do not plan one utterance, then articulate it, then plan the next utterance, and so on. Speakers normally prepare the next utterance while talking. But how exactly does speaking work?

Most psycholinguistic theories of speech production (for example, Garrett, 1975, 1988; Dell, 1986; Levelt, 1989, 1992; Dell & O'Seaghdha, 1991, 1992; Bock & Levelt, 1994) assume that there are three main levels of speech production which precede the final process of articulation: Message encoding, grammatical encoding, and phonological encoding. In what follows, these three levels are outlined in some detail. The discussion follows Levelt (1989, 1993).

In addition to the skill of speaking we also have the skill of listening to other people (and also to ourselves). Therefore, a model of the speech production system must allow for a (parallel) process of listening. This component of the speaker's processing system will also be briefly discussed below. Figure 1.1 provides a graphic framework with the speaking components on the left side, and the comprehension components on the right.

1.1 The blueprint of speaking

1.1.1 Message encoding

In a message the speaker wants to refer to persons, objects, situations, events, emotions and so on. At the same time the speaker wants to make propositions (statements) about these referents. The encoding of this message is called conceptualisation (Levelt, 1989). The conceptualiser draws upon the speaker's world knowledge, the representation of the current communicative situation, and the memory of what has already been communicated. These issues are briefly discussed next.

World knowledge is stored in long-term memory (Baddeley, 1986; Ericsson & Kintsch, 1995). It consists of information about persons, things, events, actions, and many more. These semantic categories are conceptually structured. The nature of these structures has been analysed in divergent ways. Some research focused on the semantic representation of concepts and posit that a concept is represented as a whole (for example, the mother as MOTHER, see Collins & Loftus, 1975; Roelofs, 1992b, for a review). Others posit that a concept is combined of parts (for example, the mother as PARENT and FEMALE, see Dell, 1986; Dell & O'Sheaghdha, 1991, 1992; Goldman, 1975; Morton, 1969; Miller & Johnson-Laird, 1976). In these approaches, the semantic knowledge representation is often described in so-called network representations where nodes stand for concepts and arcs stand for relations between them. Another class of semantic research concentrates on the relation between semantic and syntactic functions.

Here, one alternative is an indexed tree (Jackendoff, 1983, 1987) where an event like *The dog bites the man* can be combined by using so-called function/argument relations. Here the action function BITE takes two arguments, ANIMAL and PERSON, where in this case the PERSON concept is *man* and the ANIMAL concept is *dog*, respectively. The semantic categories can be combined in order to form a proposition, such as *the dog bites the man*. This combination can be represented in bracket notation, such as $(_{\text{EVENT}}$ BITE$(_{\text{ANIMAL}}$ DOG, $_{\text{PERSON}}$ MAN)). For more details on semantic relations see, for example, Levelt (1989) and references therein.

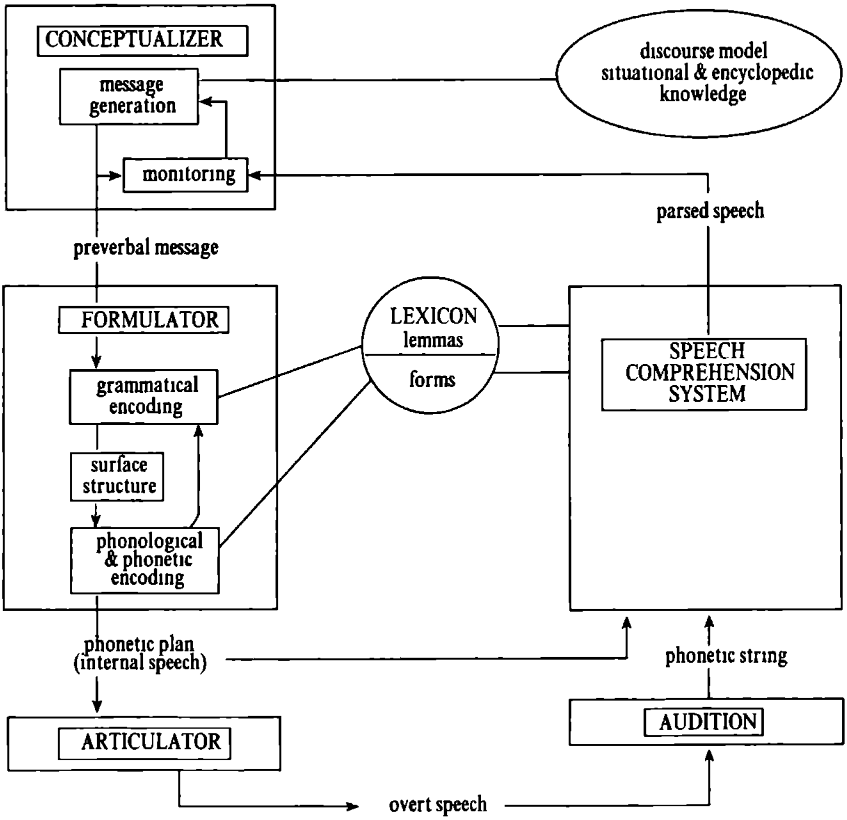


Figure 1.1 A blueprint of speech production (left) and comprehension (right), after Levelt (1993)

The speaker also takes into account the *current communicative situation.* In the present discourse situation the speaker is aware of the visual scene where the communication takes place. She or he also attends to the partner(s) in the conversation.

Finally, the speaker keeps track of what has been said already (Clark, 1994). This *discourse record* varies over time with each contribution made to the conversation, whether by the speaker or by another participant. Some of its content stays in working memory (Baddeley, 1986; Ericsson & Kintsch, 1995) only for a short period of time, and some content gets stored more deeply in long term memory.

Conceptualisation involves macro- and micro-planning (Levelt, 1989). During *macroplanning* the speaker's intention is encoded. He or she has to select and to linearize the information to be expressed. Selection refers to deciding on what to express at all. Principles that guide the selection process have been formulated by Grice (1975). They are basically conversational maxims, such as: be co-operative, be as informative as required, don't be redundant, and - important for researchers - do not say that for which you lack adequate evidence, and many more (for a review see Levelt, 1989).

After the speaker selects the intended elements, they should be put into some order. The linearization may follow natural order principles, such as chronological order. These linearization principles can only be effective in the discourse if speaker and listener agree upon them. Linearization is therefore determined by the speaker's assumption about the mutual knowledge in the speech situation. In situations where there is no such common order principle the speaker must randomly start somewhere. He or she has to keep track of what has been said already, what the listener is still to be told, and whether the listener is still with him or her. This process requires special memory resources which in turn influence linearization. Ordering principles, such as spatial or mental connectivity, are used in order to keep memory load low (Levelt, 1982). The result of the macroplanning is the speech act intention (Levelt, 1989).

During *microplanning* the speech act intention becomes more fine-grained. A speaker will mark referents in a message for their accessibility as being new, or being part of the topic, or being the focus (Chafe, 1976). This marking guides the listener's attention to what is already given in the discourse or to signal that a new entity is being introduced. For example, when a particular referent has already been introduced, the speaker needs to mark its accessibility, for example, by using a pronoun in the second sentence of *The dog is big. It seems to be dangerous.*

The result of the conceptualisation is a preverbal message. This preverbal message serves as input to the so-called formulator, which takes care of grammatical and phonological encoding.

1.1.2 Grammatical encoding

During grammatical encoding message elements become associated with lexical information, with sentence structure, and with word forms. The process of grammatical encoding can be subdivided into *functional processing* and *positional processing* (Bock & Levelt, 1994; Bock, 1995)[1].

Functional processing involves *lexical selection* and *function assignment.* During *lexical selection*, the selected concepts activate the corresponding lexical entries that are needed to convey the speaker's meaning. These entries are called lemmas (after Kempen and Huijbers, 1983). In order to express a message, such *as The man frightens the dog*, the lemmas for *man*, *frightens*, and *dog* become selected. The selection of the lemma makes available the grammatical information that is associated with the word (for example, information about its number and tense, and whether it is a noun or verb). Evidence for a lexical selection mechanism comes from speech errors such as *The my frightens the dog*. In this case, the error *my* is assumed to be the result of the parallel selection of two closely related meanings (such as *man* and *guy*), which creates a so-called blend (Bolinger, 1961; MacKay, 1972). Another form of error is the semantic substitution as in *A man frightens the cat* when *dog* instead of *cat* was intended. The substitutions preserve features of the meaning of the intended word (Hotopf, 1980). In addition, they are nearly always members of the same grammatical class, for example, nouns are substituted by nouns, and verbs by verbs (Stemberger, 1985).

Another part of functional processing is *function assignment.* It determines what grammatical roles different phrases or words will play. For example, in the sentence *He kissed her,* the masculine referent is associated with the nominative noun phrase, the feminine referent with the accusative noun phrase. The verb *kissed* unites them in the intended way because the verb maps the agent/patient argument structure of the concept KISS (X,Y) onto the syntactic subject/object function in the sentence. Errors of function assignment arise when elements are assigned to the wrong functions, as in a hypothetical phrase exchange such as *She kissed him* when *He kissed her* was intended (Garrett, 1980). This exchange error involves constituents of the same type (both are noun phrases). It is not a simple exchange of word forms, because the error does not result in, for

---

[1] This subdivision generally corresponds to the *functional* and *positional* levels in Garrett's theory (1975, 1988), and to the top two levels of the lexical network with their associated *tactic frames* in Dell's (1986) spreading-activation model

example, *Her kissed he.* The error, therefore, involves a switch in the function assignments.

During functional processing the encoded elements send information to positional processing, which fixes the order of the elements in an utterance. The assumption that the process of ordering does not take place during functional processing came from a contrast in scope between different types of errors (Garrett, 1980). Garrett observed that whole word exchanges occurred within the same phrases only 19% of the time, whereas sound exchanges (as in *lemporal tobe*) originated in the same phrase 87% of the time. This difference in error proportions implies that the two kinds of errors, word and sound exchanges, originate from different processing stages.

Positional processing involves *lexical retrieval* and *constituent assembly. Lexical retrieval* involves the activation of grammatical morphemes (Bock, 1995[2]). Grammatical morphemes consist of inflections and closed-class words, such as *-ed* and *the* in *The guy called her.*

Constituent assembly creates a tree-like hierarchy of phrasal constituents and inflectional morphemes. This hierarchy manages the order of word production. Following Dell (1986, p. 290) the hierarchical structure consists of syntactic and morphological frames. The frames specify an ordered set of categorically labelled slots. For example, for the utterance *Some dogs bite* the slots in the syntactic frame consists of the grammatical functions 'noun phrase' and 'verb phrase'. They have to be filled with the selected lemma information. The slots in a morphological frame of *dogs* consist of stem and affix. They have to be filled by the stem *dog* and the plural ending 's'. Constituent assembly, therefore, combines information of functional and morphological encoding.

One type of error directly supports the idea of having inflection procedures during constituent assembly that are independent of the word's meaning. The error is known as 'stranding' (Garrett, 1982). For example, the sentence *The new building has marvellous resting rooms* could be erroneously uttered as *The new resting room has marvellous buildings.* In this error, the suffix 'plural -s' shows up in its proper location in the utterance but in combination with the wrong (exchanged) word stem. In addition, error data support the idea that grammatical morphemes are retrieved differently from phonological information. Garrett (1982) showed that grammatical morphemes are rarely involved in errors - in contrast to phonological segments which are vulnerable to sound exchanges and substitution (but see Dell, 1990).

As Bock (1995) pointed out, although the general distinction between functional and positional processing is widely accepted in the literature, there are still

---

[2] In Bock and Levelt (1994) lexical retrieval is called 'inflection'.

aspects under debate. One aspect concerns the nature of processing flow within the grammatical encoding stage and towards the adjacent level of phonological encoding. This will be addressed in more detail below (see section 1.1.4 regarding strict two-stage vs. cascaded processing views). In addition, the interplay between grammatical encoding and conceptual encoding is under debate. It is still unclear how message details serve to specify grammatical details in the course of speaking, and vice versa. It seems to be obvious that the grammatical structure has to express the speaker's intentions. However, it is still an open issue whether grammatical encoding can also influence current conceptual access processes. This issue addresses the question about the architecture of the message-to-syntax processing system. If the grammatical encoding involves the updating of the message, the system can be called *interactive*. In contrast, if grammatical encoding only imposes constraints on its own product, it can be seen as *modular*. A theory of language comprehension that postulates the interactive processing view has been proposed by Bates and MacWhinney (1989). In contrast, Frazier (1987) postulates autonomous grammatical encoding in a modular processing system that does not influence conceptual encoding. For a theory of grammatical encoding during language production I refer the reader to Kempen and Hoenkamp (1987, see also De Smedt, 1996).

1.1.3 Phonological encoding

During the stage of phonological encoding an articulatory phonetic shape of the utterance is generated. This process involves the retrieval of the words' morpheme(s), its metrical structure, and its segments. The process can be seen as filling categorially labelled slots of phonological frames (Dell, 1986; Garrett, 1975; Levelt, 1989; 1992; Shattuck-Hufnagel, 1979). For example, the lemma *dog* (marked for plural) is morphologically constructed by accessing its stem <*dog*> and the affix <*s*>. For <dog> the metrical structure says that it is monosyllabic and gives information about whether or not it is stressed. The segmental spell-out for <dog> is /d//ɔ//g/. In the next step, the selected segments are inserted into the current metrical template. They incrementally build a phonological syllable. The exact nature of the syllables depends on the speech context. For example, for *hound* as a single noun the syllable structure is 'hound' /haʊnd/, for *hound-dog* it is 'houn-dog' /haʊn.dɔg/. The segment /d/ of the morpheme <hound> takes coda position in the first case, but it gets dropped in the second case (Baumann, 1995; Levelt, Roelofs, & Meyer, accepted; for a review on syllables see Schiller, Meyer, Baayen, & Levelt, 1996). Once the phonetic form has been computed, articulation takes place.

Generally, it is assumed that sentence production is incremental (following Kempen & Hoenkamp, 1987). Incrementality means that the first activated concepts pass their information on to the lemma level. While the lemma level encodes the syntactic information, other concepts become active. At the next time step this conceptual activation is sent to the lemma level, which has just sent its first working output towards the phonological level, and so on. This cascading procedure means that the processing stages work in parallel. But each fragment of an utterance still has to go through all the stages in a sequential manner. The direction of the processing flow is strictly top-down (but see Dell, 1986, 1988; Dell & O'Seaghdha, 1992, for a different view). All bottom-up information is assumed to be the result of a monitoring process. According to Kempen and Hoenkamp, the monitor is an extra 'module', which has access to every speech level. Levelt (1989) postulated that the monitor is a feedback flow of phonological or semantic information within the comprehension system.

Evidence for the existence of these separate stages has come primarily from analyses of speech errors (cf. Garrett, 1975, 1988, 1992; Stemberger, 1985; Dell, 1986). In addition, systematic problems in word finding in aphasic patients has revealed insights into the speech encoding process (for a review see Caplan, 1994; Zurif & Swinney, 1994). A third important source of information is experimental work, mainly picture naming studies. Schriefers (1990) empirically separated conceptual and lexical encoding. The distinction at the grammatical encoding level between functional and positional assignments was shown empirically by Bock and Cutting (1992). Glaser and Düngelhoff (1984), Schriefers, Meyer, and Levelt (1990) and Levelt, Schriefers, Vorberg, Meyer, Pechmann, and Havinga (1991a) collected information about the time course of the semantic and phonological stages. Data indicate that semantic encoding precedes phonological encoding. This finding was replicated recently by means of neurophysiological studies (Van Turennout, Hagoort, & Brown, 1997). The fine-grained timing of phonological encoding has been investigated by Meyer (1990, 1991) and Roelofs (1996, in press).

1.1.4 Strict two-stage vs. cascaded processing view

Although there is reasonable agreement on the broad outline of the production process as sketched above, the exact characterisation of the time course of engagement of the levels is still a matter of debate (Levelt et al., 1991a, 1991b; Dell & O'Seaghdha, 1991, 1992). Some researchers (Levelt et al. 1991a; Schriefers et al., 1990) opt for a two-stage model, and posit that grammatical and phonological encoding are distinct stages, where each stage is influenced only by information represented at the level directly 'above' it. Others assume a global modular top-down process combined with bottom-up feedback that can influence higher level processing in what is described as a local interactive manner (Dell & O'Seaghdha, 1991, 1992). Reaction time data from picture-word interference studies seem to support the two-stage model. Mixed speech errors, such as saying *rat* instead of *cat*, support the interactive view, because they seem to involve both semantic and phonological information. These two sources of behavioural data have led to an as yet unmovable stalemate concerning the characterisation of the flow of information during speaking.

In this thesis the use of reduced speech, such as pronouns or ellipsis, should serve as a methodological means to distinguish between the two theoretical approaches.

1.2 The speech comprehension system

The comprehension of spoken language involves the extraction of information from the acoustic signal of the incoming speech. The listener integrates this information with his or her stored knowledge in the mental lexicon in order to understand the speaker's message.

The recognition of spoken language begins with the extraction of acoustic-phonetic information from the speech signal. This process is not trivial because the physical signal of acoustic speech shows no obvious information about where a word begins and ends (for a review see Lively, Pisoni, & Goldinger, 1994). One of the questions addressed by psycholinguistic theories of language comprehension is what information the listener uses to extract meaning. Are there units of perception, specific pattern of prosody, or top-down processes from the lexicon that help the listener to understand?

The *units of perception* which may become extracted from the acoustic signal are still under debate. Different types of units have been postulated, including acoustic-phonetic features (Marslen-Wilson, 1987; Marslen-Wilson & Warren, 1994), phonemes (McClelland & Elman, 1986), syllables (Segui, Frauenfelder & Mehler, 1981), and articulatory gestures (see McQueen and Cutler, 1997, for a review). There is empirical evidence for each of the proposed units of perception. However, as McQueen and Cutler have pointed out, there is still no consistent evidence about a hierarchy of these units. For example, Segui et al. (1981) compared the perception of phonemes and of syllables in a monitoring task. Participants had to detect syllables and phonemes in lists of acoustic stimuli. The authors demonstrated that syllable monitoring was faster than phoneme monitoring. This finding was interpreted as support for the syllable as a fundamental unit of perception. But Norris and Cutler (1988) argued that the results of Segui et al. were material and task specific because participants could detect the syllable (but not the phonemes) already after partial information analysis. Norris and Cutler tested their assumption by using materials where the participant had to analyze the stimuli completely before making a decision. Exactly the opposite pattern of results was observed as in the study by Segui et al.: Phoneme monitoring was faster than syllable monitoring, indicating that extraction of phonemes was faster than the extraction of syllables.

This example demonstrates that there is not only one basic unit that can be used during comprehension, but that the listener can choose units of perception, depending, for example, on the experimental task or the communicative situation (McQueen & Cutler, 1997).

A second source of information for language comprehension is *prosody*, such as lexical stress and the rhythm of the sentence. For a review of experiments concerning prosody I refer the reader to McQueen and Cutler (1997). In their conclusion, the authors pointed out that there are at least two ways in which prosody could be relevant in comprehension. The two ways are based on the distinction between lexical access and lexical retrieval. On the one hand, prosody could be part of the access code, i.e. it could help to make initial contact with entries in the mental lexicon. On the other hand, prosody could be part of the phonological code listed for a word in the lexicon and be consulted only during retrieval, i.e., after access has been achieved. McQueen and Cutler opt for the latter possibility. They argue that in order to know the stress pattern of a word, the listener's word recognition system must know how many syllables the word has. The system could, therefore, not begin the process of lexical access until the end of the word if it needed stress pattern information before access could take place. This would lead to a disadvantage of delayed initiation of access, which is not predicted if prosody is a lexical process.

A third source of information during comprehension is the *lexicon*. As for language production, it consists of a variety of information. It carries phonological and morphological information that should be accessed by incoming acoustic-phonetic information or by pre-lexically extracted perceptual units. But, in addition, in order to make an interpretation of spoken language possible, the lexicon consists of syntactic and semantic information. Empirical evidence exists that during word recognition several lexical candidates initially become activated (accessed), and only later on in the recognition process one candidate from this cohort becomes selected (Marslen-Wilson & Welsh, 1978; Marslen-Wilson, 1987, 1990; Zwitserlood, 1989). The members of the cohort are assumed to compete with each other at the lexical level (Colombo, 1986; Goldinger, Luce, Pisoni, & Marcario, 1992; McQueen, Norris, & Cutler, 1994; Slowiaczek & Hamburger, 1992).

One important question that has been addressed by theories of language comprehension is whether phonological, syntactic or semantic lexical

information is involved in pre-lexical processing. Two major classes of theories differ with regard to the answer. Interactive models hold that lexical information influences pre-lexical processing. In contrast, autonomous models assume that lexical information is not involved in pre-lexical processes.
One interactive model is TRACE (McClelland & Elman, 1986). It consists of three levels of processing units: Features, phonemes, and words. Units within a level compete with each other via lateral inhibition, realized by means of inhibitory connections between these units. Units at lower levels activate corresponding units at higher levels via facilitatory connections. This leads to a bottom-up spreading of activation during word recognition: Feature units activate phoneme units, which in turn activate word units. But higher level nodes also facilitate lower level units. That means that activated word units send activation down to their corresponding phonemes, indicating that lexical information influences pre-lexical processing.
An example for an autonomous model is the two-stage model of spoken-word recognition SHORTLIST (Norris, 1994). In the first stage pre-lexical segmental information is represented. In a second stage, a lexical 'shortlist' of candidate words is represented. The words compete with each other via a process of lateral inhibition, until one word dominates the activation pattern, and can then be recognised. Information flows only bottom-up. The model, therefore, claims that lexical information does not influence pre-lexical processing.
Both models can account for lexical effects in a variety of experimental tasks (for a detailed discussion see McQueen & Cutler, 1997).

As lexical items and metrical structure become successively available, the listener will immediately try to interpret these materials. 'On-line' with the incoming information syntactic and semantic processing take place together with discourse processing. Grammatical encoding and discourse processing in speech comprehension are subjects of extensive research. They will, therefore, not be addressed here (for a review see Levelt, 1993).

1.3 Differences and commonalties between the two language modes

Before I address the topic of this thesis, I would like to add some remarks on the crosstalk of production and comprehension. The nature of this crosstalk is especially important in experiments where cross-modal tasks are involved, as are described in this thesis. In these experiments, participants are usually in a picture naming mode, but occasionally they hear an acoustic stimulus. Depending on the paradigm, they have to ignore the stimulus, as in picture-word interference task in Chapter 2. Or they have to actively make a decision on the stimulus, as during lexical decision in the dual task paradigm in Chapter 3. The nature of any priming or interference effect has to be seen in the light of the intersection of the production and the comprehension systems.

Figure 1.1 suggests that word recognition and production share the same lexicon, which includes word forms and lemmas. Although the term 'lemma' originated in speech production research, its meaning can be transferred into the word recognition domain as well: 'lemma' means the syntactic and semantic information of a word. Whether the same lexicon is involved in both production and comprehension is still under discussion (Levelt, 1993; Cutler, 1995). Some theorists opt for different representations for the two cognitive tasks. Others think that the representation (the stored knowledge) is the same but that the processing (temporal information flow) differs.

With regard to phonological representations Zwitserlood (1994) posits that phonological effects in experimental cross-modal tasks are evidence for a crosstalk between the production and the comprehension systems. Zwitserlood argues that different directions of effects in production and comprehension strongly suggests different processes - but not necessarily different representations. Therefore, only one phonological level might be involved. This idea is in line with Starreveld and La Heij (1995). In contrast, Roelofs, Meyer, and Levelt (1996) assume two different phonological representations, one for acoustic/orthographic input and one for naming output. In their model, the contact between comprehension and production is established by direct connections between the two phonological levels.

With regard to syntactic representation, Roelofs et al. (1996) assume that lemmas are shared for both the acoustic words and to-be-named pictures (in line with Dell & O'Seaghdha, 1991, 1992; and Harley, 1993). The same holds for the conceptual representation. This topic will be discussed in more detail in the connectionist modelling part of this thesis (Chapter 4).

1.4 The research question in this thesis

The blueprint of the speaker is normally supposed to generate complete utterances, such as the statement *The dog bites the man.* However, speech output can also be reduced. Reductions include the use of pronouns and the omission of parts of the message (ellipsis)[3]. These reductions occur in situations where the speaker wants to refer to something which has just been mentioned before in the current discourse and which has already been in focus (Chafe, 1976; Marslen-Wilson, Levy, & Tyler, 1982).

---

[3] These reduced forms can also be called *anaphora* (see for example, Hankamer & Sag, 1976) They are linguistic expressions that refer to entities in the discourse The reference process can either be directed backward in time or forward. In this thesis only backward reference has been addressed.

Omissions, also called ellipses, are often generated in adjacent question-answer pairs. When someone asks *Who did the dog bite?* the answer can be *David.* What is the underlying message of this answer? It might be PAST(BITE(DOG, DAVID)) as in a complete utterance. It might also be DAVID, the mere PERSON concept. If it is the complex variant, the conceptualiser first generates the whole package of information. But later, most of the information is deleted somewhere between the semantic encoding and the overt articulation. The cancellation might take place during grammatical encoding or during phonological encoding. At a first glance, this seems to be a somewhat redundant procedure. But it may be the case. The alternative procedure of choosing only the PERSON concept DAVID seems to be more efficient. However, this solution is not without its problems either. Under certain circumstances the answer to the question *Who did the dog bite?* might not be *David* but *him.* Note, that the pronominalized answer in this context cannot be *he.* The pronoun has to take accusative case. But how does the grammatical encoding know about this? The semantic fact that the person here is the patient of the action is not sufficient to mark the pronoun accusative. There are also cases where the patient does not receive accusative case, for example, in the German utterance *Ich helfe ihm* (I help him, dative case) vs. *Ich unterstütze ihn* (I support him, accusative case). Somehow, the grammatical encoding has to have access to the information which verb is involved, because the verb carries the information about the case to be used for its grammatical object. This so-called ellipsis problem was first mentioned by Bühler (1934) and is still a serious issue in linguistic research (Klein, 1984, 1993).

This ellipsis problem will be addressed in more detail in the second chapter of this thesis. There, a psycholinguistic way is described to distinguish between complete and reduced conceptualisation. The underlying assumption is that the two suggested ways of building the elliptic form differ with respect to whether or not the verb's conceptual meaning becomes active. If semantic activation of the elided verb can be traced, that might be evidence for complex conceptual encoding. If no semantic activation is found this might be support for reduced conceptual encoding during the generation of ellipses. Given that semantic activation of the elided verb is found, another interesting question is where in the speech system the verb's activation is cancelled. I used picture-word-interference experiments to investigate whether the meaning and/or the phonological form of an elided verb is active during the generation of elliptic utterances.

Another form of reduction is the use of pronouns, such as in *He dropped a stick.* Here, the pronoun *he* refers to a person who has been introduced in the discourse before and has been in the focus, as in the adjacent sentence pair *David was not* *in perfect shape yesterday. He dropped a stick.* The activation of the concept DAVID can also activate information linked to this person, for example that David is the excellent drummer of a very famous punk band in Nijmegen. Because the pronoun refers to the PERSON concept DAVID, this semantic information should also be active when the speaker is planning to produce the reduced form *he*. According to the theory of lemma selection (Roelofs, 1992a, b) the concept DAVID automatically activates the lemma *David* which in turn activates the appropriate syntactic gender *masculine*. The ACTION concept DROP automatically activates the lemma *drop* which assigns nominative case to its grammatical subject[4]. In isolated sentences this situation would lead to the overt naming of the name *David*. But within a coherent speech situation this selection process becomes influenced by the local discourse structure (Chafe, 1976; Gordon & Scearce, 1995): The preceding sentence mentioned DAVID already. DAVID, therefore, is in the present focus of the discourse record. The discourse somehow influences lexical access in such a way that the nominative masculine pronoun *he* becomes selected instead of the noun *David*. However, how exactly the discourse record influences the lexical access process is still an open issue. Does discourse affect access at the conceptual level or at the lemma level? It may influence syntactic processes at the word or at the sentence level. It may also affect the phonological surface structure of an utterance.

One first step towards addressing this question is to look whether the activation of overt noun naming differs from the activation during pronoun generation. Let us assume first that the conceptual activation of *David* in *David is the drummer* is the same as for *he* in *David won't play the guitar tonight. He is the drummer.* In this particular adjacent sentence pair the speaker is obviously referring to the PERSON concept by using the pronoun. The question then is whether lemma access and/or phonological form retrieval differ for nouns and pronouns. In this thesis the phonological activation for a noun word during both its overt noun generation and its pronoun generation are compared. According to theories of language production (Dell, 1986; Levelt, 1989; Roelofs, 1992a, b) the overt noun naming should lead to phonological activation of its phonemes, because they are needed for articulation. However, the situation is not clear for pronoun generation. As discussed above, the lemma has to be activated during pronoun generation, otherwise the correct access of its gender or case information would not be possible. What are the phonological consequences of having the noun's lemma active? According to a two-stage theory of lexical access (Roelofs,

---

[4] Note that the gender in this particular example might also be activated by the biological gender at the conceptual level. However, in languages that randomly assigns gender to objects, such as Dutch (neuter, common gender) or German (masculine, feminine, neuter gender) conceptual gender access is not possible. Here the gender access is assumed to be located at the syntactic level (for a review on gender access see Van Berkum, 1996, in press; Jescheniak & Levelt, 1994).

1992a, b) a distinction is made between activated and selected lemmas. It is possible to have several lemmas active at the same time, but only one (the winner) can be selected. A selected lemma automatically spreads activation towards the phonological form. Because the noun's lemma is active - but not selected - for articulation in the pronoun case, the noun's phonemes should not be activated. An alternative way of processing may be that the noun's lemma becomes active, selects the appropriate pronoun, and in addition, spreads activation towards the form level. Under these circumstances the phonemes of the not overtly spoken noun may be activated. These hypotheses were investigated empirically by means of dual task experiments, the so-called 'lexical decision during picture naming' paradigm (see Chapter 3). In addition, a neural network approach was used to simulate the dual task situation and to explain the results (Chapter 4). Finally, the results for ellipses and pronouns are compared in the light of the fact that they are different forms of reductions (Chapter 5).

In this chapter, I will outline the so-called ellipsis problem, following Klein (1993). Two possible ways of producing ellipsis are discussed: Reduction and completion. In addition, a psycholinguistic model of speech production is described (Roelofs, 1992a, b) and a view of how to create partial corrections is given, following De Smedt and Kempen (1987). The theoretical assumptions are then empirically tested in a series of picture-word interference experiments using the adjacency ellipsis "partial correction". Before I describe the experiments, the background of the picture-word interference paradigm is outlined.

2.1 The ellipsis problem

As mentioned in the introduction, the way in which a speaker creates ellipsis is still an open issue in current linguistic and psycholinguistic research. It is not clear, for example, to what extent the various kinds of linguistic phenomena that are traditionally discussed under the label 'ellipsis' are uniform in nature, and whether they are produced in the same way. Klein (1993) gives a comprehensive survey of the various types of ellipsis. One such type is, for example, inscriptions on signs. The message "No smoking" usually tells us that you must not smoke in the area around this sign. This knowledge is not explicitly mentioned on the sign, but the reader can infer it from world knowledge. The same holds for newspaper headlines, such as "Simpson no killer" where world knowledge might tell the reader who Simpson is and who got killed and that Simpson might not be responsible for it. Fixed expressions like "after you" are quite a different type of ellipsis. The message behind this might be something like "Please, go first, because you are female and I am male and well-educated." Furthermore there are ellipses which are specific to speakers who have just started to learn a language. Children, for example, initially use single word sentences to communicate with others. The same holds for some second language learners. There are also processing dependent ellipses, for example, if people are under time pressure they might use a kind of telegram style in order to keep the amount of cognitive load low. This might also be the case for aphasic patients during agrammatic talk (see Heeschen and Kolk, 1988, for details). In all of these cases, additional information beyond what is said in the elliptic construction is supplied from context, for example, from world knowledge or from visual information which speaker and listener have access to in the situation. It is also possible that this contextual information is explicitly mentioned in the linguistic context, for example, in a preceding clause or sentence. The best-studied case of this type of ellipsis is elliptical coordination. The utterance *I feed the dog and the cat* could be a reduced form of *I feed the dog and I feed the cat.* However, the sentence might also be a phrasal coordination without involving deletions or reductions. Ellipsis within coordinations has been extensively studied in the linguistic literature over the last thirty years (for different grammatical approaches to the construction of coordination ellipsis see Klein, 1993). This is much less the case for ellipsis within 'adjacency pairs', that will be addressed in this thesis. Here, the full and the elliptical expression form a closely related pair of utterances. The clearest and best-known examples are question-answer sequences. When someone asks *Who played where?* the answer can be *Jari in Amsterdam.* Here, the verb *played* is somehow understood in the answer but it is not expressed there - it is elided. Another case of adjacency ellipsis are partial confirmations, such as *It was a good concert. Yes, the best one this year.* Yet another form of adjacency ellipsis are partial corrections, such as *Jari played the piano. (No), the guitar.* How does the speaker create such a partial correction? Having the processing stages of the blueprint of the speaker (see introduction) in mind, the processing at each single level of the speaking process might be looked at: Is there conceptual activation of the elided element? What does the syntactic form of ellipsis look like? Is a phonological representation for the elided element present? These questions will be addressed separately next.

2.1.1 What happens to the concepts of elided elements?

What is the underlying message of the correction *(No), the guitar,* when uttered after the preceding statement *Jari played the piano*? There are two possibilities. The message might be PAST(PLAY(JARI, GUITAR)), as in the complete utterance *Jari played the piano.* Alternatively, it might just be GUITAR, the mere OBJECT concept. If the underlying message is the full variant, the conceptualizer first generates the whole package of information. But later, most of the information is deleted somewhere between the semantic encoding and the overt articulation. If this is the case, the concepts of JARI and PLAY should be active. Under the alternative assumption, the conceptualizer only selects the OBJECT concept GUITAR. The concepts of JARI and PLAY might not be activated at all.

These two different approaches to the message generation of ellipsis can be tested. The basic idea is to look whether JARI and PLAY are semantically active if speakers generate self-corrections. The way of measuring semantic activation is described below in the section about the experimental paradigm "picture-word interference". If we do find semantic activation for the elided units, this might be evidence that the conceptual encoding for ellipsis is the same as for complete utterances. It then has to be investigated where the activation of elided units become cancelled in the speech generation process.

2.1.2 What is the syntactic form of ellipsis?

As discussed in the introduction, there is evidence that at least a part of the syntactic information provided by the elided verb is available in the elliptic construction. This is less clear in English with its limited case marking than in German. Consider, for example, *Ich begegne ihm. Nein, ihr.* (I meet him. No, her.), where the verb requires the dative, in contrast to *Ich treffe ihn. Nein, sie.* (I meet him. No, her.), where the more or less synonymous verb requires the accusative. There is no source where this varying case-requirement could come from except from the elided verb.

Klein (1993) distinguishes two different approaches to cover this fact: *Reduction* and *completion*. Under the completion approach, it is assumed that there are specific syntactic rules just for elliptical constructions. Thus, the complete syntactic information is initially not available. The 'initial structure' must somehow be completed. Reduction mean that the entire information is available right from the beginning, and it is then successively reduced so as to obtain the final elliptical utterance. Two types of reduction, syntactic reduction and phonological reduction will be discussed below.

Klein gives an example for a partial correction in order to show the differences underlying the completion and the reduction assumption:

Deine Uhr geht vor. (Your clock is fast.)
Deine nach. (Yours slow.)

The partial correction *Deine nach* means in this context the same as the non-elliptic construction *Deine Uhr geht nach*, where *nachgehen* is a lexical verb with a separable particle.

A. The completion approach

Under the *completion approach*, it is assumed that there are not only syntactic rules which generate the full sentence but also specific syntactic rules for the elliptical structure. For example, a full sentence (S) might consist of a noun phrase (NP) *Deine Uhr* and a verb phrase (VP) *geht vor*. The NP in turn consists of a possessive pronoun (POSS) and a simple noun (N), and the VP consists of a particle (Vpart) and the verb (V). Such a structure could be generated, for example, by the rules S -> NP VP, NP -> POSS N, VP -> Vpart V (we ignore morphological variation here). For the construction of ellipsis, however, the syntactic processor might need special ellipsis rules that directly generate the elliptical form *Deine nach*. The differences can best be seen in the outline of syntactic tree structures, as depicted in Figure 2.1.

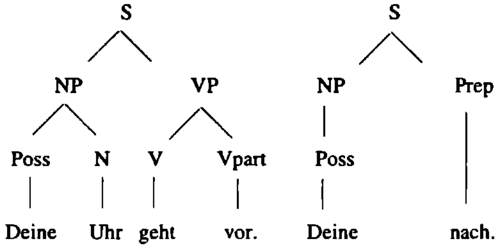


Figure 2.1
The completion approach
(after Klein, 1993)

The core problem of this approach is the fact that it requires a lot of additional and often quite arbitrary rules. In addition to the rule NP -> POSS N, as needed for the complete structure, one would also need a rule such as NP -> POSS. Such rules virtually violate all normal constraints on syntactic rules. In view of the huge amount of different forms of ellipses, it is hard to find a general solution for their construction. A general solution here means to have a finite set of syntactic rules which are specific for ellipsis and can cover all cases. Klein (1993) pointed out that at the moment such a general solution does not exist.

B. Two reduction approaches

Under the *reduction approach* it is assumed that the rules for the complete and the elliptical syntactic form are the same, and hence they have the same underlying structure. In the case of ellipsis, this structure has to be reduced. Klein (1993) distinguishes between two ways of reduction. The first one concerns a potential *syntactic reduction*. The second approach is the *phonological reduction* (p-reduction).

The syntactic reduction is based on Chomsky's (1965) notion of deletion transformations, which basically says "delete identical constituents (if they are recoverable)". For the above mentioned example the syntactic structures might be as depicted in Figure 2.2.

As Klein pointed out for coordination ellipsis this deletion rule might hold for examples like *The parents and the kids are sleeping*. Here, the complete form is *The parents are sleeping and the kids are sleeping.*

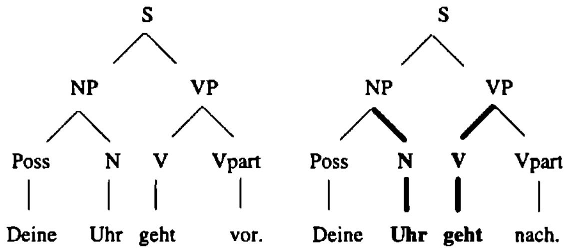


Figure 2.2
Syntactic reduction:
'Delete identical constituents'
(depicted in bold)

However, there are other coordinations where the deletion does not work, as in *Peter and Colin are a funny team.* This structure cannot be based on a complete form like *Peter are a funny team and Colin are a funny team.* More recent approaches of generative grammar (for example, Government and Binding) do not use deletions anymore. Instead they try to find specific rules for a phrasal coordination (see Klein, 1993, for a review). But, according to Klein, here the same problems hold as for the completion approach. A general solution on how to handle ellipsis does not exist yet.

The second reduction approach, the so-called phonological reduction process, assumes identical semantic and syntactic structures for both the complete and the elliptical utterance. Only the phonological form of parts of the messages is reduced under certain conditions (see Figure 2.3).

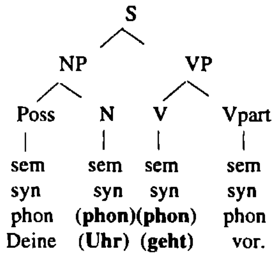


Figure 2.3
Phonological reduction (after Klein, 1993)

What are these conditions? They vary somewhat with the particular type of ellipsis. In the cases discussed here, there are two requirements: First, the information must be maintained from a previous utterance, and, second, it must be 'topic' information, rather than 'focus' information. In principle, these conditions are independent, but in practice, they often go hand in hand. This is particularly clear in the case of question-answer adjacency pairs, as in *Who bought the tickets for the concert?* Assume the complete answer here is *Niels bought the tickets for the concert.* The question *who* introduced all possible persons who might have bought the tickets, for example, all the PhD students at the Max Planck Institute. In addition, the question asks for a specification of which one among those is the one for which the assertion is true. The group of candidates is called the *topic*. The topic is specified by the topic expression *bought the tickets for the concert*. The answer specifies one of these persons as the one who in this particular context did the action. This person is called the *focus* of the utterance. According to Klein (1993, p. 791) every lexical unit which expresses the topic can be phonologically reduced. In the above mentioned example this is exactly the topic expression *bought the tickets for the concert*. Therefore, the answer to the question can be just *Niels*. For the focus a phonological reduction is not allowed. *Niels* has to be generated.

Klein pointed out that what matters is that the topic itself is identical. This does not necessarily mean that the syntactic constructions which express it in question and answer are identical in form. It is, therefore, not crucial to have the identical topic expression in question and answer. Klein mentioned the case of deictic expressions. For example, if there is a change of speaker the same topic has to be realised in a different topic expression, as in *your ticket* vs. *my ticket*. The topic is the same, the ticket, but the topic expression changes depending on the perspective of the speakers.

In the particular case of adjacent question-answer pairs the topic-focus structure of the answer is relatively fixed. The focus expression of the answer is that part which corresponds to the specific wh-phrase (who, when, what). The topic expression is the rest and can be p-reduced. In other constructions it is not that easy to define the topic expression. The definition is dependent on contextual factors, such as which elements are old and which are new. The topic expression is also dependent on the internal topic-focus structure of the utterance, for example, the marking of topic and focus by using a particular word order or intonation. However, as Klein pointed out, it is possible to define the topic expression, and the p-reduction rule can be applied to a broad range of different types of ellipses (for details and constraints, see Klein, 1993; for a short introduction of the terms 'topic' and 'focus' see also Chapter 3). The phonological reduction rule can now serve as a hypothesis for an empirical investigation of the third question of interest:

2.1.3 Is there phonological activation of the elided element?

As outlined above, Klein's p-reduction would predict that there is no phonological activation of the elided elements, whereas there is activation for the syntactic and the semantic part of the lexical information. Because Klein's theory is a formal linguistic one that makes no specific claims about psychological processing, his idea should be applied to current psycholinguistic notions on how the speaker generates speech. I will introduce a model of lexical access (Roelofs, 1992a, b), because this model makes detailed assumptions on how a lemma, and by that, syntactic information is accessed. This background is important, because Klein's p-reduction assumed a complete semantic and syntactic representation.

2.2 A network model of lexical access

According to Roelofs' (1992a, b) model of lexical access the processing of a particular message involves three levels of representation: The conceptual level, the lemma level, and the word form level. The conceptual level represents meaning. The lemma level gives access to syntactic knowledge. The word form level represents phonological information of a word. Part of the network is illustrated in Figure 2.4.

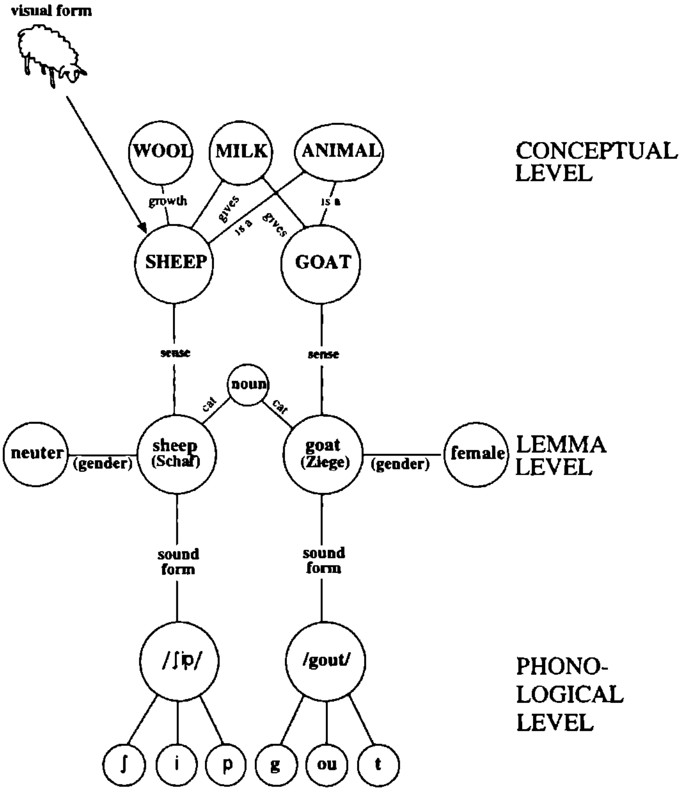


Figure 2.4 A network model of lexical access (after Roelofs, 1992b)

It gives an example about the knowledge we have about the words *sheep* and *goat*. At the conceptual level, the nodes represent concepts. They are linked by labelled arcs that represent the nature of relationships. That a sheep is an animal is represented by an *IS-A* link from the node SHEEP to the node ANIMAL. There is also a connection *GIVES* to the concept node MILK, because some sheep might give milk. A word's meaning as a whole is stored by such a network of connections. These conceptual nodes have connections to nodes at the second, the lemma level. Here, syntactic properties are stored. The lemma *sheep* has a category link to the noun node. Depending on the language, the syntactic structure might differ. This is illustrated for syntactic gender access. In French the lemma *mouton* (sheep) has a gender link to the masculine node, whereas in German it has connection to the neuter gender node. The lemma node has direct connections to the corresponding word form node at the word-form level, which in turn is connected to its phonological segments (for a detailed description of the network with regard to phonological encoding, see Roelofs, 1996, in press).

In this model, activation spreads from concepts to lemmas to word forms. A concept node becomes active in very different ways. For example, it can be activated by an idea that has to be expressed, or by a visual scene which the speaker wants to describe. A very simple form of a visual scene is a line drawing that depicts an object, for example that of a sheep. If a concept gets activated, it spreads its activation to all connected concept nodes. For example, if the SHEEP concept is active, the GOAT node will receive some activation as well. This can happen either directly, or along mediated pathways, for example from SHEEP to ANIMAL to GOAT. In addition, the concept will spread activation towards its corresponding lemma. In the above mentioned example, SHEEP will activate the lemma *sheep*, but GOAT will also activate its lemma *goat*. The probability that any given lemma will be selected during a specified time interval is the ratio of its activation to the total activation of all lemmas. This ratio follows Luce's rule (1959). Hence, Roelofs distinguishes between activated and selected lemmas. There might be several lemmas active at a time. But only one lemma wins the competition and becomes selected. This target lemma automatically activates its form which then may become articulated. This way of processing can be interpreted as a two-stage model. One stage is the conceptual activation and the lemma selection, another stage is the phonological encoding.

If we apply this model to the elliptical case, we could assume the following: In the partial correction *Jari played the piano. No, the guitar* during the complete utterance the concepts JARI, PLAY and PIANO become active, their lemmas selected, and the phonological units articulated. In the reduced part of the utterance only the concept of GUITAR goes through the activation procedure. What happens to the elided elements *Jari* and *play*? As discussed above, their concepts should be active and activate in turn the appropriate lemmas. At the lemma level, the syntactic information is accessed (such as case and gender). The further processing from lemma to form can now go on in different ways, depending on the amount of activation the lemma may have. According to Roelofs the state of activation of a lemma can have different qualities.[1] It can be slightly activated, and it can be highly activated in comparison to other lemmas. The latter case leads to the selections of that lemma. Lemma *selection* automatically leads to phonological activation of its phonemes. Mere lemma *activation* does not necessarily lead to phonological activation, but it might be the case. This means, that whereas Klein's theoretical approach posits no phonological activation of elided elements, this is still open for psycholinguistic models of speech production. The question is open, because the activational state a lemma must have in order to give access to syntactic functions is unknown yet: Activation of the lemma (not its selection) may be sufficient. Therefore, the mere fact that syntactic access takes place need not tell much about the exact activational state of the lemma.

Because Roelofs' model focuses on single word production and not on sentence production I will introduce a model which handles aspects of complex sentence production. In addition, the model deals with the generation of self-corrections. An introduction to the nature of corrections is important here because the utterance format used in the present experiments is so-called partial corrections, one form of adjacency ellipsis.

2.3 Incremental sentence production and self-corrections

De Smedt and Kempen (1987) developed a model of sentence production including a mechanism for the production of self-corrections. Basically they assume the same levels of processing that have been already introduced before: The conceptual stage, the lemma level, and the phonological level. But now the processing content is not a single word but a multiple-word utterance, a sentence. De Smedt et al. posit that a sentence cannot be processed as one whole unit from level to level. They infer this from the fact that if a sentence is produced as a whole unit, a correction could only occur at the end of a sentence. This is clearly not the case. Correction takes place during sentence production as well as at the end of it. De Smedt et al., therefore, assume that sentence

---

[1] It should be mentioned here that Roelofs (1992a, b) modelled single word production. The assumption of how the model would perform on ellipsis is purely theoretical.

production is incremental (following Kempen & Hoenkamp, 1987; see Chapter 1).

This cascade-like procedure means that the processing stages work in parallel. But each fragment of an utterance still has to go through all the stages in a sequential manner. The direction of the processing flow is strictly top-down. All bottom-up information is assumed to be the result of a monitoring process. This monitor is an extra 'module' which has access to every speech level.

As causes for corrections De Smedt et al. listed three kinds of conceptual modification: Deletion, replacement and addition.

Deletion:	Miranda bought a new Jaguar for...uh...a Jaguar for her vacation.
Replace:	Jos..uh sorry...José got the money for buying this marvellous house.
Addition:	Arie...and Ardi couldn't stop playing laser quest.

The cause of corrections might be located at the conceptual level. The meaning of the to be expressed utterance has to be changed. This change can be due to a change of a visual scene a speaker is trying to describe, or by a change in the speaker's ideas to be expressed. This conceptual change can effect the speech system in different ways. It can lead to a simple exchange of lemmas, as in *I met two..no.. three friends of mine yesterday.* But a conceptual change might also involve a change in the syntactic structure. For example, after a conceptual addition it might not always be possible to just continue with the sentence in an incremental fashion. An addition might lead to a syntactic dead end, as in the following English example: *Angelika comes...Angelika would like to come.* In order to express to concept LIKE in this context the whole verb phrase has to be re-generated. This is not necessary in Dutch for this particular case: *Angelika komt...graag.* The language specific differences indicate that the rules governing repair are not necessarily conceptual in nature, but often syntactic (for a syntactic well-formedness rule of repairs see Levelt, 1983; for rules of reformulation and lemma substitution see Van Wijk and Kempen, 1987).

The monitor's responsibility is the detection of problems and the initiation of a self-correction. In principle, monitoring holds for problems at all speech levels: concept, lemma, form. The access of the monitor to all speech levels makes the interpretation of the cause of a correction sometimes ambivalent. The above mentioned example of replacement might be a conceptual correction: The concept of the man *Jos* had to be replaced by the concept of his girlfriend *José.* But in this particular case the phonological encoding might have caused a mistake by forming /Jos/ instead of the phonologically closely related alternative, although the concept JOS was at no moment active.

The description of the De Smedt et al. model should give an impression on how the speech system creates complex utterances and how it deals with corrections. It should serve as a demonstration that corrections are of different kinds. Corrections can be caused by malfunctions within the speech system. Corrections can also be due to speech-external changes of the speaker's intentions. Of course, intentional changes may lead to problems in the speech system, but not necessarily so. The elliptical correction that is used in the following experiments are corrections of the type that involves no malfunctions. In our correction elicitation paradigm described below, the cause of the elliptical correction is a perceived pictorial change. This change in the perceptual world is external to the language production system. It is assumed that the visually driven change enters the speech system at the topmost level, the conceptual encoding stage. The assumption that the input for the conceptual stage is the same for both the complete and the elliptical description is crucial. Only if the input is the same, for example, visually driven, a speech specific comparison of semantic and phonological activation during overt and elided naming makes sense.

A method to measure semantic and phonological activation is the picture-word interference paradigm. Its underlying theoretical assumptions are described next.

2.4 The picture-word interference paradigm

At the beginning of this chapter different ways for building ellipsis were outlined. They are summarised here by formulating two working questions. First, is the concept of an elided element active? If semantic activation of the elided element can be traced, that might be evidence for a complex conceptual encoding. If no semantic activation is found this might be support for reduced conceptual encoding during the generation of an elliptical expression. Given that semantic activation of the elided element is found, another interesting question is where in the speech system the element's activation is cancelled. This leads us to the second working question: Is the phonological form of an elided element active? According to Klein's p-reduction rule the concept of an elided element should be active, but not its phonological form. Semantic and phonological activation can be measured by means of picture-word-interference experiments.

The picture-word interference paradigm is a generalisation of the Stroop task. Stroop (1935) asked participants to either name the colours of coloured squares or to name the colours of colour words in a list. The colour of the ink could either be congruent to the meaning of the colour word (for example, the word *red* written in red) or it could be incongruent (for example, the word *red* written in blue). In comparison to naming the colours of squares, the naming of incongruent colours in the colour word condition is hampered and naming latencies are increased. This increase in time is known as the Stroop interference effect (see MacLeod, 1991, for a review). The basic idea is that during naming of the colour of words, two cognitive encoding mechanism are at work in parallel: The encoding of the visual colour information, and the encoding of the word's meaning. In some cases the encoding of the meaning might be faster and the participant tends to name the meaning instead of the ink colour. This tendency has to be suppressed in the incongruent condition in order to make the right response. As a result of this response competition the naming gets hampered (Morton, 1969; Posner & Snyder, 1975). How the race between the two encoding processes might look like is still under debate (for a detailed discussion on deterministic vs. stochastic race models see Vorberg, 1985).

In a picture-word task the participant has to name a picture and should ignore the word which is presented with the picture (either acoustically or written, we focus on acoustic stimuli here). The picture naming here is an analogue to the colour naming in the original Stroop task. The acoustic word can be seen as the (to be ignored) meaning of the colour word (Glaser & Düngelhoff, 1984). In the picture-word task two general encoding processes are involved: The naming of the picture and the recognition of the word. According to Glaser and Glaser (1989) the process of picture naming involves four steps, which are comparable to the already introduced levels of speech production: perception, encoding of conceptual meaning (semantic memory), encoding of linguistic knowledge (words, morphemes, phonemes), and phonemic execution. According to Glaser and Glaser, word recognition is assumed to involve phonemic decoding, phonological decoding, and semantic decoding, respectively. Both processing routines share the same representation at the semantic and word form levels.

Two effects are quite robust within the literature of the picture-word task: Phonological facilitation and semantic interference. The *phonological facilitation effect* means that the naming of a picture is speeded up when an acoustic distractor word is presented that is phonologically related to the picture's name - in comparison to a presentation of an unrelated distractor word. For example, the naming of the picture 'dog' is faster while hearing the phonologically related 'doll' in comparison to the presentation of an unrelated word 'table'. The effect is assumed to be generated during word-form retrieval. At this level the distractor word will activate its phonemes. For example, the acoustic distractor 'doll' activates the phonemes /d//ɔ/. Therefore, a distractor word that is phonologically related to the word of the depicted object will also to some extent activate the phonemes of the to be named picture word 'dog'. The phoneme retrieval of this word will be faster in comparison to the situation in which an unrelated distractor word is presented (Schriefers, Meyer, & Levelt, 1990). This explanation is in line with cohort theories in the spoken word

recognition literature (Marslen-Wilson, 1987; Marslen-Wilson & Tyler, 1980; Slowiaczek & Hamburger, 1992)
Very important to mention here is that the effect only occurs when the activation of the two stimuli overlap in time at the phonological level. When is this overlap going to happen? According to the assumed naming process it takes some time to reach phonological encoding (namely after the successful encoding of the visual and semantic information). In contrast, the decoding of the acoustic word enters the phonological stage relatively early, because this stage is the beginning of the word recognition process. In order to obtain overlap of phonological activation the picture naming process will need a head start. This temporal advantage can be given by presenting the picture first, and after some time the acoustic stimuli. This onset manipulation is called the SOA variation (stimulus onset asynchrony). Given the time course assumptions for naming and word recognition, a phonological effect is expected at some positive SOA (picture first, acoustic probe second). In contrast, no phonological facilitation is expected at SOA = 0 (presenting the two stimuli at the same time) or at an negative SOA (presenting the acoustic word first, and the picture second). This is exactly what Schriefers et al. (1990) found in their time course study.

The *semantic interference effect* means that the naming of a picture is slowed down when a distractor word is presented that is semantically related to the picture word - in comparison to the presentation of an unrelated distractor word. For example, the naming of 'dog' takes longer if the participant hears 'cat', which is of the same semantic category than the picture name - than if she or he hears the unrelated word 'table'. The locus of the semantic effect is still under debate. Some theorists assume the *conceptual level* to be responsible (Glaser & Düngelhoff, 1984; Lupker & Katz, 1981): The decision about which concept to express verbally takes more time when a semantically similar item is active simultaneously with the target concept. Some theorists explain the interference by *competition at the lemma level* (Roelofs, 1992a, b). The semantically related concepts of picture word and distractor word activates each other, and in turn their lemmas. In Roelofs' network the path from picture to distractor lemma node (DOG -> CAT -> cat; upper case represents concepts, lower case lemmas) is shorter than the path from distractor word to target lemma node (cat -> CAT -> DOG -> dog). Therefore, the picture will prime the distractor lemma more than the distractor will prime the target lemma. As a result, the distractor lemma is highly activated. The higher alternative lemmas are activated the longer it takes to make a lemma selection for the target word. Empirical evidence to favour the lemma instead of the concept as cause of inhibition came from a picture-word interference study by Schriefers et al. (1990). In a standard picture-word interference study they found a typical semantic inhibition effect. But in addition, they showed that this effect disappeared when the participant's task was not naming but picture recognition. The pictures and distractors were the same ones as in the naming experiment, but now a push button response had to be carried out. A *yes* or *no* response decision was made according to whether a picture had appeared in the presession or not. The lack of semantic inhibition in a non-language task implies that the semantic effect is located within the lemma level and not at the conceptual level.

A third potential locus of semantic interference is the *word form level*, as suggested by La Heij (1988) and Starreveld and La Heij (1995). In their view, there exists no lemma level. The target picture will activate its concept and by spreading of activation all nodes of its semantic neighbours. All activated concepts will automatically activate their word forms. Therefore, a semantically related acoustic distractor will get activation from both its own acoustic input and from the picture encoding process. It will, therefore, be more active than an unrelated distractor word. In any case, the word form of the picture word has to be retrieved. The retrieval process takes longer the higher the activation of alternative - among them semantically related distractors. Empirical evidence here came from picture-word interference studies involving priming of orthographically similar semantic distractors (Starreveld & La Heij, 1995). The authors found the usual semantic interference effect by semantic distractors that were not orthographically related to the picture's name. But the semantic interference effect was significantly reduced if the distractor words were orthographically related. Starreveld et al. argued that the observed modification of the semantic effect by orthographic similarity can only take place at the phonological level, during word form retrieval. Here, an interaction takes place between phonological facilitation and semantic inhibition, leading to a zero semantic effect (for a different interpretation of the results see Roelofs et al., 1996).

A completely different account of the semantic interference effect comes from the field of attention research. Here, theorists favour locating the inhibition at the level of *response execution* - rather than in semantic memory. It is suggested that as perceptual representations of two stimuli are produced, they prime their associated responses automatically. These responses are held in check by an inhibitory mechanism until one response reaches an activation threshold. The response that would be given to the ignored prime competes with that required by the attended probe. The competition has to be resolved, which takes time. In addition, the inhibited responses remain activated and will impair subsequent performance of the same stimuli (Eriksen & Eriksen, 1974; Posner & Snyder, 1975). This effect can be called negative priming (Tipper, 1985). Both alternative accounts for interference, a language specific semantic effect and a language unspecific response effect, were tested by Tipper and Driver (1988) and Tipper, MacQueen, and Brehaut (1988). According to response competition a negative priming effect should not be present in cross-modal responses (for example, naming the prime, keypress reaction to the target, or vice versa). In contrast, if the interference effect is of a semantic nature, this effect should be independent of response modality variations. The data of Tipper et al. (1988) clearly favoured the semantic account, showing that negative priming was observed between response modalities. Tipper et al. made no claims about the precise locus of the semantic effect. Their work is mentioned here only for ruling out response mode explanations of interference effects.

Although there is at the moment no general consensus about the location of the semantic effect within the language system (lemma vs. word form), it is linguistic in both cases. Therefore, for this study it was assumed that the semantic effect is a real lexical access effect, not an effect of non-linguistic specific cognitive processes. A semantic interference effect is expected only if the concept of the elided element is active with its lemma, or its word form, respectively. The hypothesis for the phonological effect is obvious: Phonological facilitation is expected if the phonological form of the elided element is available during the generation of ellipsis.

In the experiments to be reported in this section the picture-word interference paradigm has been used. In the first part of this section two pretests are described. These pretests included single picture naming. They were carried out for testing the materials and SOA that should be used in the main experiments. In the second part I present a series of experiments that investigated the activation of elided verbs.

In order to entice participants to produce elliptical utterances, the standard picture-word paradigm had to be changed. Instead of presenting one picture, a sequence of two pictures was presented. For the description of the pictures an imperative mood was chosen, such as *Kiss Pien!*. The imperative guarantees the target, in this case the verb, to appear in the utterance-initial position. This initial position of the target element facilitates the estimation of the optimal timing between picture and distractor onset. Usually, the first position in naming has the advantage of being independent of speaking rate and length of the other items in the sentence. The varying length of other elements does not change the position of the target, and therefore, does not change the timing between distractor and target.

Although grammatical theories might assume that in imperatives the verb is 'moved' from a later position in the deep structure towards the initial position in the surface structure (Pollock, 1989; Visser, 1992), details of this assumption are not relevant here. Whatever the reality of this movement is, it will be constant for complete and elliptical utterance generation.

As shown in Figure 2.5, the pictures depicted two kids being involved in an action. The actor was always the same kid that wears sunglasses. The patient was one of four potential candidates, two girls and two boys. They were introduced as Pien, Tess, Paul, and Toon. In an instruction-like fashion the participant should order the actor of the scene to do the particular action to the other kid. The description of the first picture was, for example, *Kiss Paul!*, if the kid with the sunglasses kissed the boy named Paul. On the second picture, the patient Paul changed to Tess. The participants were asked to interpret this change as a kind of correction of the visual scene and to name it aloud. This led to a naming, such as *Kiss Paul.....Tess!'*, which created a gap for the elided verb 'kiss' in front of 'Paul'. This elliptical naming condition will be compared with two other naming formats. They are described in the main section.

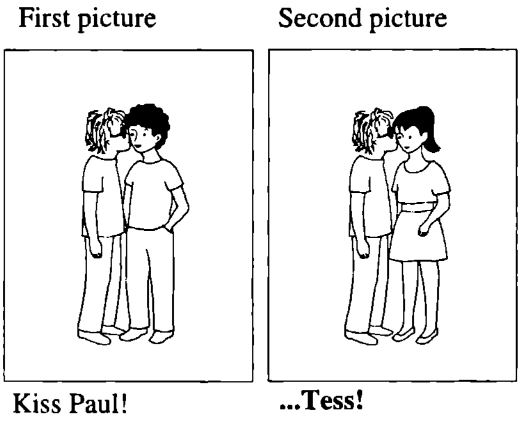


Figure 2.5
Example of a sequence of two pictures that elicits an elliptical description The target verb is the elided verb 'Kiss' during the second picture description.

2.5 Pretests of pictures and interfering stimuli

The pictures and interfering stimuli used in the main experiments were selected in pretests. These experiments were done in order to get an idea of how long it takes the participants to name a picture in an imperative manner. The naming latencies were needed to find an appropriate picture-word SOA in the main experiment. The pretests should also provide information on whether the chosen acoustic material was suitable to induce phonological and semantic effects during the main picture naming task.

2.5.1 Method

Participants: The participants were drawn from the Max Planck Institute subject pool. Their native language was Dutch. They were paid for their participation. They were 18 to 36 years old. The naming pretest with an SOA of 0 ms was carried out with 25 participants. In the naming pretest with an SOA of 200 ms 14 participants were tested.

The description task: In the main experiment the generation of an imperative was the target of investigation. Therefore, in the pretests the same utterance format was tested. Participants had to describe pictures in an imperative manner. A kid wearing sunglasses 'did something' to one of four other children (Toon, Paul, Tess, Pien). Participants were asked to tell the boy with the sunglasses what he should do to the other child. For example, participants should say *Kiss Paul!* if the child Paul was kissed by the boy wearing glasses. During the pretests only the complete utterance (verb plus patient) was tested.

Materials:

Pictures: 23 different actions were depicted, such as to kiss, to wash, to turn around, and so on. (see Appendix B for the used Dutch materials). In each picture two kids were involved in some action. The actor was always the kid wearing sunglasses. The patient was one of four potential candidates. Each action was combined with each of the four candidates, resulting in 92 different pictures. The pictures were black-on-white line drawings of approximately equal size, created by the author. The pictures of the children were available in the Max Planck picture pool.

Acoustic distractors: Each target verb presented as an action on the picture was combined with four prime categories. The prime was of the same syntactic category as the target (verb). It was either phonologically related (for example, *kill* to the target *kiss*) or unrelated (*wash*). In the related condition the depicted target verb and the interfering stimuli shared the same phonemes, as onset and nucleus. In addition, there was a semantically related verb (*hug* to the target *kiss*) and an unrelated verb (*find*). The semantic relation was defined according to Miller and Fellbaum (1991), using same category members. In the unrelated condition the same words were used as in the related conditions, but new combinations of distractors and targets were created - with the constraint not to relate semantically or phonologically to the target verb. This rotation guaranteed the same word frequency and recognition time of the acoustic distractors for both the related and unrelated condition. The acoustic distractor words were spoken by a female speaker and recorded using a Sony 59ES DAT recorder. They were digitised with a sampling frequency of 16kHz and stored on the hard disk of the computer.

Design: There were two separate distractor conditions: Semantic (related, unrelated) and phonological (related, unrelated) distractors. These factors were tested in a complete within-subject design. Each participant saw each picture eight times (the four priming conditions were combined with two different patients). The picture-prime pairs were randomly presented per participant, with some constraints: First, identical target verbs did not follow each other. At least one different depicted action had to be in between. Second, a sequence of members of the same semantic category was excluded (for example, *kiss* and *hug*).

Apparatus: The experiment was run on a Hermac 386 SX computer. The pictures were presented on a Nec MultiSync 4FG screen. The participants heard the acoustic distractors on a Sennheiser HMD 224 Headphone-Microphone combination. The participants' speech was recorded using a Sennheiser HMD 224 Headphone-Microphone combination and a SONY TCD D3 DAT recorder. Reaction times were measured by a voice key.

Procedure: The participants were tested individually or in groups of two. At the beginning of the experiment, they were given a booklet including the instruction and one example for each action. The target verb was printed below the picture. After the participant learned the names of the kids and the target verbs, a second set of pictures was presented. Now the action scene were printed without the target verb. The participants were asked to describe each scene as trained before. They were asked to do this as fast and as accurately as possible. This practice part guaranteed the use of the right verb, the correct names of the children, and the correct naming in an imperative way. This procedure was followed by a practice session on the computer, where again one version per target-verb was presented. The practice procedure lasted 20 minutes and was followed by four experimental blocks, with short pauses between them. After each pause, the next block started with 3 warming up trials, which were excluded from the analysis. The entire experiment lasted approximately 50 minutes.

The SOAs: According to theories about time course of activation (Schriefers et al., 1990) a short SOA was expected to result in semantic interference. A longer SOA should result in phonological facilitation. Therefore, two versions of the picture-word-interference tests were carried out to find the appropriate SOAs per condition. The versions only differed with respect to the timing of prime and picture onset. In one experiment the SOA was 0 ms: The onset of the acoustic prime was at the onset of the picture. In the second experiment an SOA of 200 ms was tested: The onset of the acoustic prime was 200 ms after picture onset.

Trial structure: First, for 500 ms a black fixation cross was presented in the centre of the white screen. Then, for 250 ms the screen turned white again. After this interval the picture appeared for 1500 ms. The onset of the acoustic prime was 0 or 200 ms after picture onset. The time-out for naming was 2000 ms.

After this interval the screen remained white for another 1000 ms. Then, the fixation cross started the next trial.

2.5.2 Results

Statistical analysis: The statistical analyses were based on mean reaction times for correct naming responses. Paired sample t-tests were carried out separately for the semantic and the phonological condition.[2] Errors were excluded from the analysis, defined as wrong naming, time-outs (no reaction after 2 seconds), and voice key triggers due to non-speech signals. This resulted in excluding 2-3% errors per naming condition.

Results for SOA = 0: The mean naming latencies for the SOA = 0 ms pretest showed a 23 ms effect of semantic inhibition. Naming latencies were increased for the semantically related condition in comparison to the unrelated condition: Subjects, $t(24)=4.06$, $p<0.01$, items: $t(22)=1.99$, $p=0.059$.[3] No phonological effect was found at SOA = 0 ms.

Results for SOA = 200: The mean naming latencies for the SOA 200 ms pretest showed a 50 ms effect of phonological facilitation. Naming was faster in the phonologically related condition than in the unrelated condition: Subjects, $t(13)=-3.63$, $p<0.01$; items, $t(22)=-4.68$, $p<0.01$. No significant semantic effect was observed at SOA = 200.

To summarise, the results of the pretests showed that timing and materials were appropriate to obtain semantic and phonological activation of the target verb in complex picture naming. In addition, the results supported the theory about the time course of semantic and phonological activation during speech production (Schriefers et al., 1990; Levelt et al., 1991a; Van Turennout et al., 1997). Semantic inhibition was found at short SOA, phonological facilitation was found a long SOA, indicating that semantic activation precedes phonological activation during naming. I decided to use an SOA of 200 ms in the main experiments for the phonological priming condition. The SOA of 0 ms was thought to be a suitable timing for the semantic condition. Separate SOAs per condition were chosen because I was not primarily interested in the time course of activation, but in whether or not an effect could be observed.

---

[2] Separate analyses were chosen, because I was not primarily interested in the comparison between semantic and phonological effects, but in whether or not these effects could be found.

[3] Note that the result for the by-item-analysis in the SOA = 0 ms semantic condition was significant if a one-tailed testing is assumed The one-tailed testing is possible because the direction of the semantic effect was expected.

2.6 The first ellipsis experiment: A respond-respond variant

Is it possible to show activation of an elided verb during picture naming? The present experiment tested the activation of the target verb on the semantic and on the phonological level in utterance production. As described above, in order to elicit a more complex utterance a sequence of two pictures were presented. Depending on the content of the two pictures, different naming formats can be obtained. In this study, three different utterance conditions were investigated: The elliptical condition, a complete control condition, and an identical condition (see also Appendix A for an illustration).

The elliptical condition: The elliptical condition was triggered by presenting a sequence of two pictures where the action remained the same, but where the patient changed from the first to the second picture. This sequence elicited an elliptical partial correction, such as 'Kiss Paul .....Tess!'. The elliptical condition should test whether the generation of ellipsis could be influenced by the priming of the elided element, in this case the elided verb 'kiss' at the gap position in front of 'Tess'.

The complete control condition: By presenting a sequence where the action and the patient changed in the second picture, a complete utterance was elicited, for example, 'Find Paul....Kiss Tess!'. The complete condition was included for two reasons: First, it served as control for getting interference effects with the present material for overt naming. Given the results of the pretest, during complete encoding of the second picture semantic interference and phonological facilitation should be observed. If no effects are obtained, the results for the elliptical condition cannot be interpreted. The second reason for including the complete control condition was to minimise strategic effects. If the actions were always the same for the two pictures, the participant might tend to ignore the verb and might not even activate its concept - at least for the second picture. By introducing the possibility of having different actions in one sequence, this strategy should be less probable.

The identical condition: As just mentioned, depending on the sequence of the two pictures complete or elliptical utterance was required. On first view the activation of the overtly spoken verb and of the elided verb might be directly comparable. But one should be careful. Eberhard, Bock and Griffin (1994) showed that word naming can influence the naming of a target picture on the following trial, if the two naming processes are related. Eberhard et al. used phonologically related or unrelated prime-picture pairs. The prime word was presented first and had to be read aloud. Then, the picture was presented and had to be named. Naming latencies were increased when a related prime had to be processed before. The results indicated phonological inhibition of the picture naming by the related prime. According to Eberhard et al., the inhibition may be located at the lemma level or at the phonological level due to competition.

In our experiment, during elliptical response the overt articulation of an identical verb immediately precedes the target utterance (this is in a sense the nature of most ellipses). How would the content of the first picture influence the priming and generation of the second picture description? Because we had no idea about that in advance, we included a so-called identical condition. Here, the visual input is exactly the same as in the elliptical condition. But the participants are instructed not to use a reduction but to name the second picture completely, such as 'Kiss Paul....Kiss Tess!'. In this case identical priming of the second verb by the first mentioned verb might take place. This priming might effect the semantic encoding and/or the phonological encoding of the second picture (Wheeldon & Monsell, 1994). This information of potential identical priming might also be crucial for elliptical processing, because in the present experiment the elided element has just been overtly uttered in the preceding picture description.

Note that in the three naming conditions the target verb, referring to the action in the second picture, stayed the same, 'kiss'. The conditions differed with respect to whether or not this verb was overtly spoken (ellipsis vs. complete and identical condition). And the conditions differed with respect to whether or not the action stayed the same or differed (ellipsis and identical vs. complete condition).

Following the logic of the picture-word interference paradigm the description of the second picture should be systematically influenced by the interfering stimuli. The focus of this experiment was whether this target verb 'kiss' is still active in the second part of the utterance. Is the target verb semantically or phonologically active? To test this, an acoustic distractor stimulus was presented time-locked to the onset of the second picture. This stimulus was either semantically or phonologically related, or unrelated to the target verb 'kiss'.

2.6.1 Method

Participants: 29 participants were tested. They were between 18 and 36 years old and were recruited from the Max Planck participant pool. They were paid for participation.

Materials and utterance conditions: The same pictures and interfering stimuli were used as in the pretests. The phonological primes were given 200 ms after picture onset, the semantic primes at SOA = 0. These SOAs were chosen because of the results of the pretests. Three utterance conditions, called 'elliptical', 'complete' and 'identical', were created. Table 2.1 shows an example for every condition.

Table 2.1 Example for the utterance conditions of the ellipsis experiment.

Picture 1	Picture 2	Naming condition	Naming
Kiss Paul	Kiss Tess	elliptical	Kiss Paul.... Tess
Find Paul	Kiss Tess	complete	Find Paul.... Kiss Tess
Kiss Paul	Kiss Tess	identical	Kiss Paul.... Kiss Tess

Each utterance condition was paired with four different distractor types: A phonologically related and unrelated distractor, a semantically related and unrelated distractor (see Appendix A and B).

Table 2.2 Block design

Block A: Elliptical block	**Block B: Identical block**
Elliptical utterance	Identical utterance
Complete utterance	Complete utterance

Design: Blocked presentation (see Table 2.2). It was not possible to run both the elliptical and the identical control condition within the same session, because different instructions had to be given. So the decision was made to present one experimental block where the elliptical and the complete conditions were combined (block A). In a second block, the identical and the complete conditions were presented (block B). This blocking led to testing the complete condition twice (once per block). These two complete conditions were seen as control conditions. They were analysed separately in a 2 (blocks: A and B) x 2 (related and unrelated distractor) within-design. The experimental elliptical and identical naming conditions were analysed in the same way, as a 2 x 2 within-design. All analyses were carried out separately for phonological and semantic distractor conditions.

Combination of actions and patients: The actions depicted in the first picture in the complete condition were chosen with the constraints not to share phonological or semantic relations with the target action in the second picture.

The patient in the first picture was only followed by a patient in the second picture that did not have a name with the same onset. This was done in order to exclude phonological effects of the "p"- and "t"-onsets of the children's names. For example, Toon at picture 1 could be followed by Pien or Paul, but not by Tess at the second presentation. The sequence of patients were kept constant within a distractor condition. For example, a change from Toon to Pien was presented for both the related and unrelated distractors. By doing so, variance due to different visual encoding within a related-unrelated distractor pair was excluded.

Hence, 8 (4 distractor words x 2 patients) trials per target action were presented per utterance condition. There were 23 different target verbs. This resulted in a total number of 184 trials per condition per subject. Therefore, each block consisted of 2 x 184 = 368 trials, the whole experiment consisted of 736 trials per subject, plus warming-up trials and practice.

Sequence of stimuli: The picture-prime pairs were randomly presented with some constraints within this randomisation: Identical target verbs on the second picture did not directly follow each other. At least one different picture-word-pair was inserted to make sure that a kind of reset of the verb was carried out between trials. For each experimental session two randomised sequences existed for every participant: One for the first session, one for the second session. In double subject sessions (given the same sequence of stimuli) one subject received the elliptical utterance instruction, and the other participant got the identical utterance instruction first, and vice versa in the second session. This rotation procedure counterbalanced potential block-sequence effects.

Procedure: The subjects were tested individually or in groups of two. Practice involved the same steps already described in the pretest section. The practice procedure lasted 20 minutes and was followed by the first experimental block. This session consisted of 4 experimental units with short pauses in between. After each break the first 3 trials were warm-ups - not included in the analysis. At this first experimental block 15 of the subjects started with the elliptical/complete condition, 14 started with the identical control/complete condition. After a coffee break the instructions were switched between participants. The second block started with 23 practice trials. Then, again 4 experimental units followed. The entire experiment lasted approximately 120 minutes.

Trial structure (see also Appendix A): For 500 ms a fixation cross was shown in the centre of the screen. Then the screen turned white for 250 ms before the first picture was presented with a duration of 2400 ms. At offset of picture 1, the second picture was presented for 1500 ms. The onset of the acoustic prime was time-locked to the onset of the second picture by an SOA = 0 for semantic primes and by an SOA = +200 ms for phonological primes. The time-out for responding was 2000 ms after onset of the second picture. After this interval, the screen turned white again for 1000 ms before the next trial started.

2.6.2 Hypothesis

Crucial to this study is the pairwise comparison of related and unrelated prime-target pairs within each condition. If priming effects were observed they could be interpreted as evidence for a phonological or semantic representation of the target verb on that specific level of processing.

Complete condition: Because of the pretests for the control condition phonological facilitation and semantic inhibition were expected.

Identical condition: If the naming of an identical first picture did not influence the priming and naming of the second picture the same results as in the pretests were expected: Semantical inhibition and phonological facilitation. However, given the above mentioned effects of repetition priming (Eberhard et al., 1994; Wheeldon & Monsell, 1994) the identical first picture might systematically influence the naming of the second picture. In addition, the first picture response might also influence the processing of the prime, or of both, the prime and the second picture response. Therefore, the results could be different to those of the pretest data. The quality of this difference was unknown.

Elliptical condition: If the concept and lemma becomes activated during elliptical naming, semantic interference was expected. If its form becomes activated, phonological facilitation should be observed.

Complete vs. Elliptical/Identical condition: In the complete condition both patient and action changed between the first and the second picture. Therefore, the second picture in the complete condition is expected to be harder to encode visually than a picture where just a patient had been changed. The latter was the case during the elliptical and identical control picture sequence. Hence, reaction times for the complete condition were expected to be longer than for the elliptical and identical condition. Besides, the latencies for the two complete utterances (one per session) should not differ.

2.6.3 Results

Statistical analysis: The results are shown in Table 2.3 in Figure 2.6 for the semantic condition. Table 2.4 and Figure 2.7 display the results for the phonological condition. The tables display the mean response latencies for the participants for every distractor condition per utterance type.

Errors were defined as a wrong naming of the verb or the patient, and the use of a wrong utterance frame. In addition, errors were time outs (reaction times longer than 2000 ms), and wrong voice key triggers (because of still describing the first picture while the voice key for the second picture description was already active). In addition 0.6% extreme outliers were detected (following Tukey, 1977), and set to missing. The errors were counted per condition. Percentages per condition are displayed in the result tables (100% were 1334 data points per condition). All error data were excluded from further reaction time analyses.

The error proportions were arc sin transformed (following Winer, 1971). A 2x2x2 repeated measures ANOVA on this arcsin transformed error proportions over subjects was carried out separately for the two distractor conditions. The factors were defined as follows: 2 blocks, 2 relatedness conditions (related, unrelated), 2 utterance conditions per block (elliptical vs. complete, identical vs. complete).

For the *semantic condition* no significant differences of error proportions for main factors and interactions were observed.

In the *phonological condition* only the interaction 'block x relatedness' was significant. $F_1[1,28]=5.4$, $p=0.03$. However, a speed-accuracy trade-off can be excluded, if we look at the reaction times. Here we observed means for the elliptical block and related primes of 910 ms (6,4% errors), for unrelated primes 927 ms (7,8% errors), for the identical block and related primes 912 ms (7,9% errors) and unrelated primes 917 ms (7,5% errors). These data showed that for one fast condition (identical block, related) the error rate was the highest. However, in the fastest condition (elliptical block, related) the error rate was the lowest.

Reaction times. The subject and item analyses of variance (ANOVAs) were performed on the mean response latencies. All analyses reported here were complete within-designs for both subjects and items. The complete conditions were seen as a separate control and were analysed separately in a 2x2 within-design. The factors were 'block' and 'relatedness'. The comparison of elliptical and identical utterances was the most relevant one for the research question. The two utterances, therefore, were directly compared in a 2x2 within-design with the factors 'utterance' and 'relatedness'. The analyses were carried out separately for the semantic and phonological distractor condition.

Results for the semantic condition

The two *complete utterance* conditions showed a significant main effect 'relatedness'. $F_1(1,28) = 9.7$, $Ms_e = 1589$, $p < 0.01$; $F_2(1,22) = 7.0$, $Ms_e = 1587$, $p = 0.01$. This main effect indicated a significant semantic inhibition. The main effect 'block' was not significant over subjects, but over items. $F_1(1,28) < 1$, $Ms_e = 5645$; $F_2(1,22) = 4.4$, $Ms_e = 781$, $p = 0.04$. This result showed that items were named faster in the elliptical block (975 ms) than in the identical block (988 ms). The interaction was not significant. $F_1(1,28) < 1$, $Ms_e = 825$; $F_2(1,22) < 1$, $Ms_e = 664$.

For the comparison of the *elliptical and the identical utterances* the main effect 'relatedness' was significant. $F_1(1,28) = 9.6$, $Ms_e = 792$, $p< 0.01$; $F_2(1,22) = 7.07$, $Ms_e = 816$, $p = 0.01$.

Table 2.3 Mean response latencies (in ms), standard deviation (SD) and percentage of errors for the semantic condition in the first ellipsis experiment

	Sequence of identical actions		Sequence of different actions	
Distractor	Block A **elliptical utterance** Kiss Pien...Toon.	Block B **identical utterance** Kiss Pien. Kiss Toon.	Control block A **complete utterance** Find Pien. Kiss Toon	Control block B **complete utterance** Find Pien. Kiss Toon
semantic. related	827 (124)	793 (90)	985 (127)	1002 (120)
unrelated	808 (117)	780 (91)	966 (116)	975 (109)
Diff. (unr-rel) (-) = inhibition	- 19	- 13	- 19	- 27
% Error rel.	8	8	7	9
% Error unr.	6	8	5	7

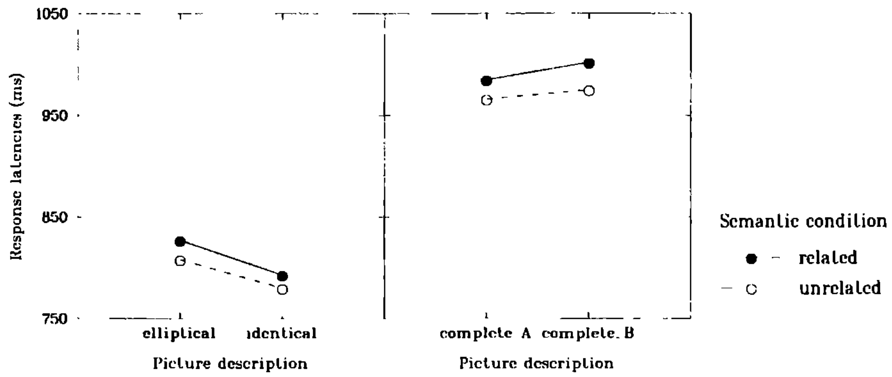


Figure 2.6 Mean response latencies for elliptical and identical responses (left) and for the complete responses (right) in the semantic distractor condition.

The main effect 'utterance' was not significantly different for subjects. But across items 'utterance' differed. $F_1(1,28) = 3.2$, $Ms_e = 8604$, $p = 0.08$; $F_2(1,22) = 33.8$, $Ms_e = 695$, $p < 0.01$. This result indicated that across items naming was faster in the identical condition (785 ms) than in the elliptical condition (816 ms). The interaction 'utterance x relatedness' was not significant. $F_1(1,28) < 1$, $Ms_e = 658$; $F_2(1,22) < 1$, $Ms_e = 432$. This lack of interaction and the significant main effect for 'relatedness' clearly showed a semantic inhibition effect for both the elliptical utterance and the identical utterance. Statistically this semantic inhibition was of equal size for the two utterance conditions.

Results for the phonological condition

The comparison of the two *complete utterance conditions* showed a significant main effect 'relatedness' (28 ms). $F_1(1,28) = 7.9$, $Ms_e = 2867$, $p < 0.01$; $F_2(1,22) = 7.8$, $Ms_e = 2188$, $p = 0.01$.

Table 2.4 Mean response latencies (in ms), standard deviation (SD) and percentage of errors for the phonological condition in the first ellipsis experiment

	Sequence of identical actions		Sequence of different actions	
Distractor	Block A **elliptical utterance** Kiss Pien...Toon.	Block B **identical utterance** Kiss Pien. Kiss Toon.	Control block A **complete utterance** Find Pien. Kiss Toon	Control block B **complete utterance** Find Pien. Kiss Toon
phon. related	836 (123)	822 (116)	985 (146)	1002 (136)
unrelated	842 (130)	805 (105)	1013 (118)	1030 (118)
Diff. (unr-rel) (+) = facilitation	+ 6	- 17	+ 28	+ 28
% Error rel.	6	7	7	9
% Error unr.	8	7	8	8

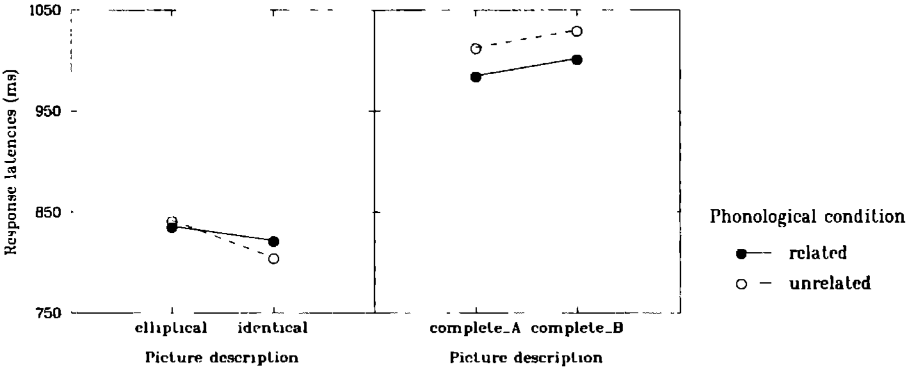


Figure 2.7 Mean response latencies for elliptical and identical responses (left) and for the complete responses (right) in the phonological distractor condition.

This main effect indicated a significant phonological facilitation. The main effect 'block' was not significant over subjects, but over items. $F_1(1,28) = 1.5$, $Ms_e = 5527$, $p = 0.22$; $F_2(1,22) = 12.7$, $Ms_e = 546$, $p < 0.01$. This result showed that items were performed faster in the elliptical block (997 ms) than in the identical block (1015 ms), although exactly the same items were used in the two blocks in the complete utterance conditions. The interaction was not significant. $F_1(1,28) < 1$, $Ms_e = 1076$; $F_2(1,22) < 1$, $Ms_e = 849$.

The comparison of the *elliptical and the identical utterances* revealed no significant main effect 'utterance' for subjects. But 'utterance' was significantly different for items. $F_1(1,28) = 2$, $Ms_e = 9714$, $p = 0.16$; $F_2(1,22) = 11.7$, $Ms_e = 1465$, $p < 0.01$. This indicates that across items responding was faster in the identical condition (812 ms) than in the elliptical condition (840 ms). The main effect 'relatedness' was not significant. $F_1(1,28) = 1.4$, $Ms_e = 714$; $F_2(1,22) = 1.7$, $Ms_e = 780$. The interaction 'utterance x relatedness' was significant.

$F_1(1,28) = 4.5$, $Ms_e = 833$, $p = 0.04$; $F_2(1,22) = 10.9$, $Ms_e = 326$, $p < 0.01$. This interaction showed that phonological inhibition of 17 ms was found for the identical condition, whereas there was a 6 ms effect in the opposite directions in the elliptical condition.

To summarise, the complete control condition showed the expected phonological facilitation and semantic inhibition. For elliptical responses no phonological effect, but a clear semantic interference effect was found. The identical condition revealed inhibition in both the phonological and semantic distractor condition.

2.6.4 Discussion

The results will be discussed separately for every utterance condition. The *complete utterance condition, Find Toon...Kiss Paul!*, was included in order to control for the appropriateness of the method. As expected, significant semantic inhibition and phonological facilitation were found. As discussed in the introduction of the picture-word interference paradigm, the locus of the semantic interference can either be at the lemma level (Roelofs, 1992a, b), or at the phonological level (La Heij, 1988; Glaser & Glaser, 1989). The finding of the phonological effect is in line with the above described facilitation during word form retrieval. The sizes of the effects are directly comparable with the results of the pretest. Fortunately, the clear results of this control naming proved the method of the two-pictures-description task to be suitable for tapping into semantic activation of overtly spoken verbs in this particular experimental set-up. The method now can be used to look into semantic activation in elliptical utterances.

For *elliptical utterance production* semantic inhibition was observed. The naming of the second picture, *...Paul!* in the naming *Kiss Toon....Paul!*, is slowed down if a distractor is presented that is related to the elided verb in comparison to presenting an unrelated distractor. This result indicates that the concept and the lemma of the elided verb are active in the gap position of the reduced partial correction. This finding puts some light on the discussion of the ellipsis problem, mentioned at the beginning of this chapter. There, the question was whether or not the concept of an elided element becomes activated. The results reported here indicate that the concept is indeed active. Where does the conceptual activation come from? In a picture naming study, it seems to be obvious that the concept gets activated by the visual input of the picture. Evidence that the meaning of a picture becomes encoded relatively early during visual perception came from semantic prime-picture naming studies (McCauley, Parmelee, Sperber, & Carr, 1980). McCauley et al. showed that a preceding picture facilitated the naming of a target picture if the pictures were semantically related in comparison to that they were not. The facilitation even occurred if the exposure time for the preceding picture was below identification threshold. This finding indicated a very early semantic/conceptual processing of the preceding picture. This result implies for the present study that the mere fact of having an ACTION concept active does not say anything about the production process involved. More important for the nature of processing ellipsis is the interpretation of the semantic effect, more specifically, the assumption about its location. As mentioned in the introduction to the picture word paradigm, this effect should be located at the lemma level (following Roelofs, 1992a, b; Schriefers et al., 1990) or at the phonological level (following Starreveld and La Heij, 1995). Both alternatives imply that a semantic interference is evidence for having the elided element active in the speech system. The plausibility of locating the semantic effect at the form level, however, becomes unlikely in the present case of elliptical speech. This thought is supported by the phonological null effect discussed next.

The phonological null effect during elliptical naming cannot be interpreted in terms of lack of power (Cohen, 1988). This conclusion can be drawn because the complete condition, which served as a control for power in this experiment, revealed a phonological effect. The observed phonological null effect of the present study, therefore, indicated that the form of an elided verb is not present in the speech system. This result plus the semantic effect are in line with the phonological reduction assumption of Klein (1993). Klein assumed complete semantic and syntactic information of the elided element. This assumption was supported by the semantic interference effect, if the effect is located at the lemma level. This is relevant, because the lemma is assumed to carry syntactic information. Klein expected no phonological activation, which was supported by the phonological null effect. The phonological null effect is also in line with the assumption of speech production theories that favor a strict two-stage lexical access (Levelt et al., 1991a, b; Roelofs, 1992). A lemma that is not selected for articulation should not activate its phonological form. In contrast, the phonological null effect contradicts cascaded processing theories (Dell & O'Seaghdha, 1991, 1992). The cascaded processing view predicts that an activated lemma automatically spreads activation toward its form, even if it is not selected for articulation. According to this view a phonological effect should be observed because the lemma is active (as shown by the semantic inhibition), and activates its phonemes, which should lead to an interaction with phonemes of a phonologically related prime. However, this phonological interaction was not observed.

The question is now where the cancelling of activation takes place. The semantic inhibition effect indicates that the lemma or the word form of the elided element is active (depending on the theory). Again, the word form approach assumes competition during form retrieval (Starreveld et al, 1995). To get this competing situation during elliptical production, the elided element has to become phonologically active. However, the present finding of a phonological null effect clearly rule out this assumption: The word form is not active. In contrast, the lemma approach (Roelofs, 1992a, b; see also the description of the network model in this chapter) assume that the *activated* target lemma gives access to syntactic information. But in contrast to single overt word naming, the lemma should not become *selected* in the sense of the Luce ratio of lemma selection. This selection cannot be the case because a selected lemma would spread activation to the form. According to the present finding of a phonological null effect, this selection assumption does not hold for elliptical utterance production. How an activation criterion for lemmas might be conceived is an interesting question for future research. The interpretation of the present finding is that the activation of the elided verb stops somewhere between the lemma and the phonological level. How this cancelling of information works and where it takes place is still an open question.

The *identical utterance condition, Kiss Toon...Kiss Paul!,* was included in order to control for identical priming effects from responding to the first picture onto responding to the second picture. As in the complete and the elliptical conditions semantic interference was observed. Interestingly, the pattern of result looks different for the phonological condition. Whereas in the complete condition a significant phonological facilitation was observed, and a null effect for ellipsis, in the identical naming significant phonological inhibition was found. This result is in line with findings by Eberhard et al. (1994). Their experiments employed phonologically primed picture naming. In one condition the participant had to name a visually (or acoustically) presented prime. Immediately or after a short delay a picture appeared that had to be named. The pictures were phonologically related to the preceding prime or unrelated. Eberhard et al. found a significant inhibition of the naming if prime and picture were phonologically related in comparison to the unrelated condition. But, Eberhard et al. also investigated the phonological effect if the prime word was not read aloud but only silently. Here, no phonological effect was observed for the subsequent picture naming. Because the modality of responding to the first picture in the present study will be changed in the next experiment, I will postpone the discussion about phonological effects during identical naming to the comparison of the two experiments.

To summarise, the data of this experiment showed that the concept of an elided element becomes active during the production of partial corrections. It has been discussed that the conceptual information enters the speech system as far as at least the lemma level. This finding is in line with Klein's (1993) assumption of building ellipsis by phonological reduction while keeping semantic and syntactic information constant. In his article Klein also posits that the p-reduction rule is context dependent, but that it should be obtained both within one speaker's utterance and when the speaker changes, as in a question-answer sequence. This speaker-independency will be addressed in the next experiment.

2.7 The second ellipsis experiment: A listen-respond variant

Context has been defined by Klein (1993) as the topic-focus relation in the ongoing discourse. In the introduction I discussed the example of a question-answer adjacency pair, *Who bought the tickets for the concert?* The complete answer was *Niels bought the tickets for the concert.* The question *who* introduced all possible persons who might have bought the tickets. The topic is specified by the topic expression *bought the tickets for the concert.* The answer specifies one of these persons as the one who in this particular context did the action, in this case it was *Niels.* This person is called the focus of the utterance. According to the phonological reduction rule every lexical unit which expresses the topic can be phonologically reduced. In the above mentioned example this is exactly the topic expression *bought the tickets for the concert.* Klein pointed out that it is not crucial to have the identical topic expression in question and answer. He mentioned a situation where there is a change of speaker. In this case the same topic has to be realised in a different topic expression. One example is deictic expressions, such as *your ticket* vs. *my ticket.* The topic is the same, *the ticket,* but the topic expression changes depending on the speaker's perspective. Speakers have to keep in mind what the topic is, and might simply delete the rest (the topic expression) in an elliptical utterance. The information about what the topic is might be stored in the discourse record, in short- or long-term memory (see Levelt, 1989, for a review). Following Klein a change of speakers should not matter for defining the topic. Regardless of whether the preceding question was asked by another speaker or by the speaker who creates the ellipsis, the same p-reduction rule should hold for generating it. This assumption of speaker-independency will be addressed next.

In the preceding experiment I investigated the nature of the generation of ellipsis in a situation where the speaker produces the context him- or herself. The participant had to respond to the first picture (the context) and had to create a reduction in responding to the second picture. The generation of the elliptical utterance was context dependent. It could only take place if the sequence of the pictures involved the same action. The results of this picture responding variant showed a particular pattern which was discussed above.

The question to be addressed next is whether the same data pattern can be observed if the context is produced by another speaker. According to Klein, the p-reduction rule depends on the context, not on the speaker. Therefore, the same results should be obtained if the first picture in our experimental sequence is named by someone else. The participant simply has to listen to the description of the first picture, and then respond to the second picture. However, if the pattern of result of this listen-respond variant looks different from the first experiment, the respond-respond variant, this might have at least two reasons: First, the p-reduction rule might not be independent of who produced the preceding context. Second, the incoming contextual information might be exactly the same. But the internal activation of the speech system is different for the two variants at the moment the context comes in. The latter case would give some information about the intersection of the discourse information and the speech production system.

2.7.1 Method

The listen-respond variant was carried out in the same way as the respond-respond variant, except for presenting an acoustically given description of the first picture via headphone. In order to contrast the results of the present listen-respond variant with the respond-respond variant with respect to the variable 'change of speaker', only this variable should be manipulated. The rest should be equal.

Participants: 27 participants were tested. They were between 18 and 34 years old and were recruited from the Max Planck participant pool. They were paid for participation.

Materials, timing, and utterance conditions: The same pictures, interfering stimuli, and SOA were used as in the preceding experiment. The description of the first picture was spoken by a female speaker and recorded using a Sony 59ES DAT recorder. They were digitised with a sampling frequency of 16kHz and stored on the hard disk of the computer.

Procedure: The practice and experimental procedure was the same as in the first experiment, except that instead of naming the first picture the participants listen to its description via headphone. The entire experiment lasted approximately 120 minutes.

Trial structure: For 500 ms a fixation cross was shown in the centre of the screen. Then the screen turned white for 250 ms before the first picture was presented with a duration of 2400 ms. The acoustic description was presented 900 ms after the onset of the first picture. This SOA was chosen because the grand mean of the naming latencies in the first experiment was 904 ms. At offset of picture 1 the second picture was presented for 1500 ms. The onset of the acoustic prime was time-locked to the onset of the second picture by an SOA = 0 for semantic primes and by an SOA = +200 ms for phonological primes. The time-out for responding was 2000 ms after onset of the second picture. After this interval, the screen turned white again for 1000 ms before the next trial started.

2.7.2 Hypothesis

The goal for this experiment is twofold. On the one hand it concerns the outcome of the experiment as independent investigation of generating ellipsis, given the context is produced by another speaker. On the other hand the results should be directly compared with the preceding experiment, where no change of speaker took place. Again, relevant in this study was the pairwise comparison of related and unrelated prime-target pairs within each distractor and utterance condition. If priming effects were observed, they could be interpreted as evidence for a phonological or semantic representation of the target verb on that specific level of processing.

Complete condition: The complete rendering of a different action should be independent of the source of the first picture description. Therefore, the same results as for the pretests and the respond-respond variant were expected: Phonological facilitation and semantic inhibition.

Identical condition: During the discussion of the first experiment it was assumed that responding to a second picture that is identical to the first picture could show an effect of the response to the first picture. The first actively executed response was assumed to leave a trace in the speech system. The nature of this trace might differ in a change-of-speaker situation. Therefore, the source of the first picture description may play a role in preparing the response to the second picture.

Elliptical condition: Klein assumed that p-reduction is speaker independent because it is the context that matters. That would mean that a change of speaker should not change the elicitation process, given the context remains constant. The same results should be observed as in the preceding study: An activated concept and lemma should be indicated by semantic interference. No phonological facilitation should be observed.

2.7.3 Results

Statistical analysis: The results for the semantic condition are shown in Table 2.5 and in Figure 2.8. Table 2.6 and Figure 2.9 display the results for the phonological condition.

Errors were defined as in the preceding experiment. 0.9% extreme outliers were detected (following Tukey, 1977) and set to missing. The errors were counted per condition. Percentages per condition are displayed in the result tables (100% were 1242 data points per condition). All error data were excluded from further reaction time analysis.

The error proportions were arc sin transformed (following Winer, 1971). A 2x2x2 repeated measures ANOVA on this arcsin transformed error proportions over subjects was carried out separately for the two distractor conditions. The factors were defined as follows: 2 blocks, 2 relatedness conditions: related, unrelated, 2 utterance conditions per block: elliptical vs. complete, identical vs. complete.

For the *semantic condition* no significant differences of error proportions for main factors and interactions were observed.

In the *phonological condition* only the interaction 'block x utterance' reached significance, $F_1[1,26]=3.9$, $p=0.057$. A speed-accuracy trade-off could be excluded. This becomes clear in the reaction time-error comparison. The reaction time means for the elliptical block and elliptical utterance were 845 ms (6,5% errors), for the complete utterance 953 ms (4% errors). The means for the identical block were: Identical utterance 754 ms (5% errors), complete utterance 942 ms (7,5% errors). This statistical trend indicated that in the elliptical condition more errors were performed than in the other utterances. As can be seen, the elliptical condition was not the fastest one.

The higher proportion of errors was due to the fact that in 1% of all elliptical utterance cases the participants used a complete utterance instead of producing the reduction (26 cases in 2484 trials). The error the other way around, that is reducing where a complete response was desired, almost never took place (0.01%, 12 cases in 7452 trials).

Reaction times. The analysis procedure is the same as in the preceding experiment. The subject and item analyses of variance (ANOVAs) were performed on the mean response latencies. All analyses reported here were complete within-designs for both subjects and items. The complete utterance was seen as a separate control condition and, therefore, analysed separately in a 2x2 within-design. The factors were 'block' and 'relatedness'. A direct comparison of identical and elliptical utterances was of interest, because according to the hypothesis the two utterance conditions might have a different outcome. They were compared in a 2x2 within-design with the factors 'utterance' and 'relatedness'. As usual, the analyses were carried out separately for the semantic and phonological distractor condition.

Results for the semantic condition

The two complete conditions showed a significant main effect of 'relatedness'. $F_1(1,26) = 11.0$, $Ms_e = 1203$, $p < 0.01$; $F_2(1,22) = 5.8$, $Ms_e = 2155$, $p = 0.02$. This main effect indicated significant semantic inhibition, as in the first experiment. The main effect 'block' was not significant. $F_1(1,26) < 1$, $Ms_e = 10897$; $F_2(1,22) < 1$, $Ms_e = 608$. The interaction was not significant either. $F_1(1,26) = 1.9$, $Ms_e = 395$; $F_2(1,22) = 1.4$, $Ms_e = 453$.

For the comparison of the elliptical and the identical conditions the main effect 'relatedness' did not reach significance. $F_1(1,26) = 3.2$, $Ms_e = 867$, $p = 0.087$; $F_2(1,22) = 1.2$, $Ms_e = 986$, $p = 0.28$. Across subjects there was a trend towards semantic inhibition. The main effect 'utterance' was significant. $F_1(1,26) = 17.6$, $Ms_e = 15041$, $p < 0.01$; $F_2(1,22) = 257$, $Ms_e = 986$, $p < 0.01$. The identical condition was faster than the elliptical one (754 vs. 845 ms). The interaction 'utterance x relatedness' was not significant. $F_1(1,26) = 1.3$, $Ms_e = 647$; $F_2(1,22) < 1$, $Ms_e = 940$. This lack of interaction would lead to the interpretation of the main factor 'relatedness' as being equal for elliptical and identical conditions. But it was obvious that the -2 ms difference in the identical condition should be seen as a null effect. Therefore, a simple contrast was carried out for the elliptical condition separately. It should directly investigate the statistical nature of the observed -13 ms semantic interference. The subject analysis revealed a significant difference. $F_1(1,26) = 3.3$, $Ms_e = 1428$, $p = 0.04$ (one-tailed[4]). The item analysis showed a trend for inhibition. $F_2(1,22) = 1.9$, $Ms_e = 1631$, $p = 0.085$ (one-tailed).

Results for the phonological condition

The comparison of the two *complete utterance conditions* showed a significant main effect 'relatedness'. $F_1(1,26) = 26.1$, $Ms_e = 1286$, $p < 0.01$; $F_2(1,22) = 20.7$, $Ms_e = 1427$, $p< 0.01$. This result indicated significant phonological facilitation. This finding replicated the results of the first experiment. The main effect 'block' was not significant. $F_1(1,26) < 1$, $Ms_e = 12125$; $F_2(1,22) = 4.0$, $Ms_e = 615$, $p = 0.055$. The statistical trend in the item analyses showed that items were named more slowly in the elliptical block (953 ms) than in the identical block (942 ms).

---

[4] One-tailed testing seemed to be appropriate here, because in the hypothesis the direction of the effect, interference, was given

Table 2.5 Mean response latencies (in ms), standard deviation (SD) and percentage of errors for the semantic condition in the listen-respond variant

Distractor	Sequence of identical actions Block A **elliptical utterance** Kiss Pien...Toon.	 Block B **identical utterance** Kiss Pien. Kiss Toon.	Sequence of different actions Control block A **complete utterance** Find Pien. Kiss Toon	 Control block B **complete utterance** Find Pien. Kiss Toon
semantic. related	829 (152)	724 (131)	927 (104)	927 (102)
unrelated	816 (131)	722 (124)	910 (98)	899 (101)
Diff. (unr-rel) (-) = inhibition	- 13	- 2	- 17	- 28
% Error rel.	7	5	4	5
% Error unr.	6	5	4	5

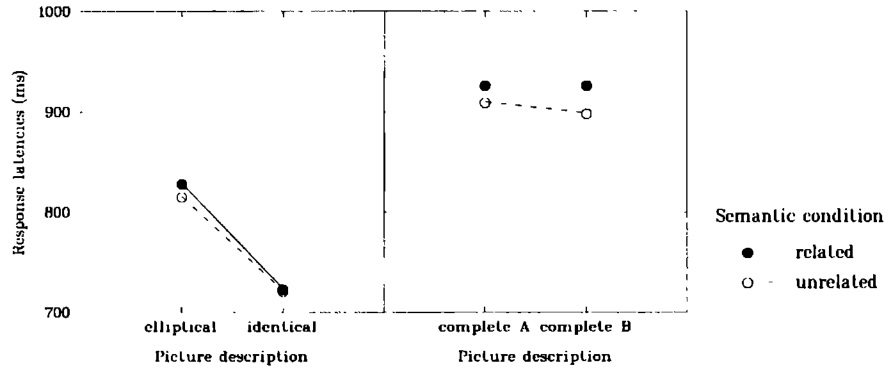


Figure 2.8 Mean response latencies for elliptical and identical responses (left) and for the complete responses (right) in the semantic distractor condition (listen-respond variant).

The interaction just failed to reach significance across subjects, but was significant across items. $F_1(1,26) = 4.08$, $Ms_e = 1006$, $p = 0.054$; $F_2(1,22) = 8.5$, $Ms_e = 399$, $p < 0.01$. This trend indicates that the observed phonological facilitation was greater if the complete utterance was produced in the identical block (+ 47 ms) than when it was produced in the elliptical block (+23 ms).

For the comparison of the *elliptical and the identical utterance conditions,* the main effect ‘relatedness’ did not reach significance across subjects and items. $F_1(1,26) = 3.0$, $Ms_e = 867$, $p = 0.09$; $F_2(1,22) = 4.2$, $Ms_e = 664$, $p = 0.051$. However, there was a trend which indicates a phonological facilitation. The main effect ‘utterance’ was significant. $F_1(1,26) = 13.1$, $Ms_e = 16893$, $p < 0.01$; $F_2(1,22) = 102$, $Ms_e = 1683$, $p < 0.01$. The identical condition was faster than the elliptical condition (754 vs. 845 ms). The interaction ‘utterance x relatedness’ was not significant. $F_1(1,26) = 1.6$, $Ms_e = 857$, $p = 0.20$; $F_2(1,22) = 3.1$, $Ms_e = 381$, $p = 0.09$.

Table 2.6 Mean response latencies (in ms), standard deviation (SD) and percentage of errors for the phonological condition in the listen-respond variant

Distractor	Sequence of identical actions Block A **elliptical utterance** Kiss Pien...Toon.	 Block B **identical utterance** Kiss Pien. Kiss Toon.	Sequence of different actions Control block A **complete utterance** Find Pien. Kiss Toon	 Control block B **complete utterance** Find Pien. Kiss Toon
phon. related	844 (150)	746 (147)	942 (133)	919 (121)
unrelated	847 (151)	763 (155)	965 (107)	966 (109)
Diff. (unr-rel) (+) = facilitation	+ 3	+ 17	+ 23	+ 47
% Error rel.	6	3	4	4
% Error unr.	6	4	4	4

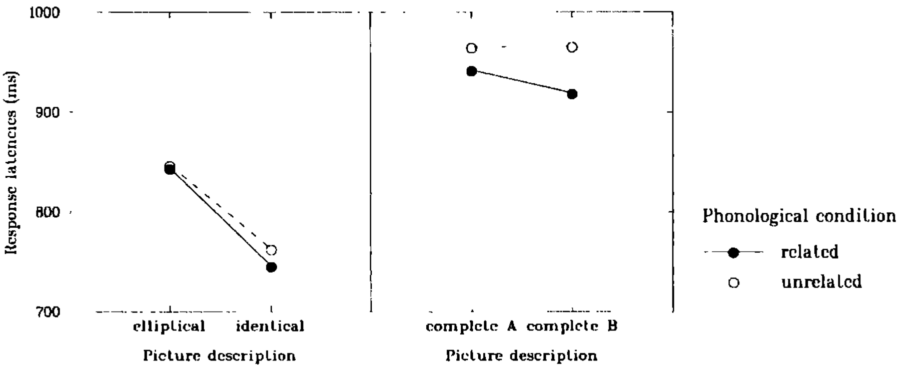


Figure 2.9 Mean response latencies for elliptical and identical responses (left) and for the complete responses (right) in the phonological distractor condition (listen-respond variant).

A simple pairwise contrast of related and unrelated distractors in the identical condition was carried out. The analysis revealed a significant phonological facilitation for identical utterances (17 ms). $F_1(1,26) = 4.1$, $Ms_e = 1898$, $p = 0.05$; $F_2(1,22) = 13$, $Ms_e = 587$, $p < 0.01$. The difference between related and unrelated conditions in the elliptical case was only +3 ms, which obviously is negligible.

To summarise, the complete control condition, *Find Pien...Kiss Toon*, showed the expected phonological facilitation and semantic inhibition. The elliptical condition, *Kiss Pien...Toon*, revealed no phonological effect and a weak trend towards semantic inhibition. The identical condition, *Kiss Pien...Kiss Toon*, revealed phonological facilitation and no semantic effect at all. Not expected, but observed, was a strong difference in facilitation between the two complete utterance conditions in the phonological condition (+47 ms in the identical block vs. +23 in the elliptical block). Also unexpected from the perspective of the first study was the observation that responses in the identical condition were faster than in the elliptical condition (about a 100 ms).

2.7.4 Discussion

The results of the *complete condition* replicated the findings of the preceding ellipsis experiment. They indicate that the method is sensitive to measure semantic and phonological effects in complex picture descriptions. The *elliptical condition* also replicated the results of the preceding study: A weak trend towards semantic inhibition indicates that the lemma of the referent noun is active during the generation of ellipsis. The observed phonological null effect showed that phonemes of the referent noun are not activated. This result supports Klein's (1993) assumption of p-reduction. In addition, the phonological null effect supports a strict two-stage theory of lexical access (Levelt et al., 1991a, b; Roelofs, 1992). A lemma that is not selected for articulation should not activate its phonological form. In contrast, the phonological null effect contradicts cascaded processing theories (Dell & O'Seaghdha, 1991, 1992). The cascaded processing view predicts that an activated lemma automatically spreads activation toward its form, even if it is not selected for articulation. The co-activation of picture name and prime should lead to a phonological effect due to competition at the phonological level. However, this phonological effect was not observed. The *identical condition* revealed facilitation in the listen-respond variant, whereas it revealed inhibition in the preceding respond-respond variant. This outcome will be discussed in detail in the next section, the comparison of the two experiments.

2.8 Comparison of the two experiments

The comparison of the two experiments was carried out in order to test Klein's assumption of speaker-independency for the rule of phonological reduction in the generation of ellipsis. This assumption was tested by varying the speaker of the first picture description between the experiments, all else being equal. The comparison was made separately for every utterance and distractor condition.

2.8.1 Comparison for the elliptical utterance conditions

The results were analysed in a 2x2 ANOVA involving the factor 'experiment' (respond-respond vs. listen-respond) as between-factor and the factor 'relatedness' (unrelated vs. related distractor) as within-factor for participants and items.

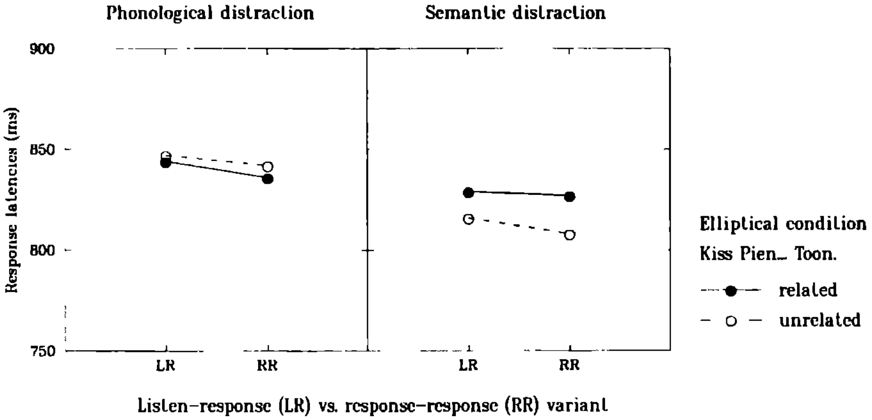


Figure 2.10 Mean response latencies (ms) for responding elliptically to the second picture of the listen-respond and the respond-respond variant.

The mean response latencies are shown in Figure 2.10, separately for the phonological and semantic distractor conditions. As can be seen from the figure, the results differ for semantic and phonological distractors. The analysis for the *semantic condition* during elliptical utterance generation showed a main effect 'relatedness' (13 ms and 19 ms). $F_1(1,54) = 11.3$, $Ms_e = 653$, $p < 0.01$; $F_2(1,44) = 9.3$, $Ms_e = 594$, $p < 0.01$. The main effect 'experiment' was not significant. $F_1(1,54) < 1$, $Ms_e = 34085$; $F_2(1,44) < 1$, $Ms_e = 2498$. There was also no interaction. $F_1(1,54) < 1$, $Ms_e = 652$; $F_2(1,44) < 1$, $Ms_e = 594$. The results showed that the experiments did not differ with regard to elliptical performance. The data revealed a clear semantic inhibition during the generation of ellipsis in both experimental contexts. A change of speaker did not change the elliptical performance.

The statistical analysis for the *phonological condition* during elliptical responses revealed no significant effect at all. The comparison showed that two experiments did not differ and that in both experiments no phonological effect was observed during the production of an elliptical utterance.

The results for *elliptical responses* can be interpreted as complete support for Klein's assumption of p-reduction. The semantic interference is evidence for having the concept and the lemma active during the generation of ellipsis, whereas the phonological zero effect showed that the form of the elided verb did not become activated. In addition, the phonological reduction was independent of speaker alternation. This means that Klein's assumption of speaker-independency for the creation of the phonological reduction was supported. In the two experiments, the context stayed constant (the content of first picture description), which should lead to the same reduction rule: Constant semantic activation, and constant phonological reduction across the two experiments. This is what was observed.

2.8.2 Comparison for the identical utterance conditions

The analyses were carried out in the same way as for the elliptical conditions involving the factors 'experiment' and 'relatedness'. The mean naming latencies are shown in Figure 2.11, separately for the phonological and semantic distractor condition. As can be seen from the figure, the results differ for semantic and phonological distractors. In addition, the reaction times are faster in the listen-respond variant.

The analysis for the *semantic condition during identical responses* showed a significant main effect of 'experiment'. $F_1(1,54) = 4.7$, $Ms_e = 23693$, $p = 0.03$; $F_2(1,44) = 36.4$, $Ms_e = 2313$, $p < 0.01$. The main effect 'relatedness' was not significant. $F_1(1,54) = 2.4$, $Ms_e = 643$; $F_2(1,44) = 1.7$, $Ms_e = 768$. There was also no interaction. $F_1(1,54) = 1.4$, $Ms_e = 643$; $F_2(1,44) < 1$, $Ms_e = 768$. The results showed that the experiments did differ with regard to latencies of responding to the second picture. The identical condition in the listen-respond variant was about 60 ms faster than in the respond-respond variant (723 vs. 786 ms).

The statistical analysis for the *phonological condition during identical responses* revealed a trend towards a difference between the experiments. The main effect 'experiment' was not significant across subject, but it was significant across items. $F_1(1,54) = 2.8$, $Ms_e = 33893$, $p = 0.098$; $F_2(1,44) = 52.6$, $Ms_e = 1397$, $p < 0.01$.

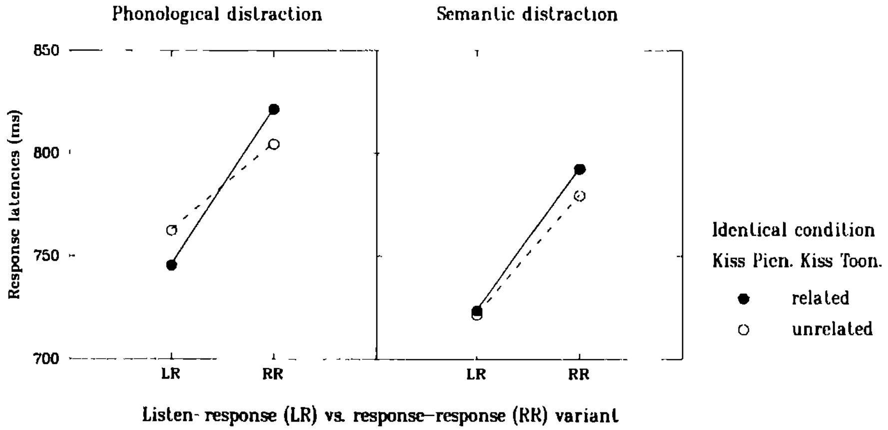


Figure 2.11 Mean response latencies (ms) of the identical utterance condition for responding to the second picture of the listen-respond and the respond-respond variant.

The main effect 'relatedness' did not show significant differences. $F_1(1,54) < 1$, $Ms_e = 959$; $F_2(1,44) < 1$, $Ms_e = 350$. But the interaction was significant. $F_1(1,54) = 8.6$, $Ms_e = 959$, $p< 0.01$; $F_2(1,44) = 24.4$, $Ms_e = 350$, $p < 0.01$. The interaction indicated that the two experiments differ with respect to the obtained phonological effect in the identical condition (17 ms facilitation in the listen-respond variant and 17 ms inhibition in the respond-respond variant). The results also showed the identical responses in the listen-respond variant to be approximately 60 ms faster than in the respond-respond variant (754 vs. 813 ms).

Thus, the *results for the identical condition* showed that the speaker source of the first picture description matters. This source effect revealed phonological facilitation in the listen-respond variant and inhibition in the respond-respond variant. This modality-specific difference was observed by Eberhard et al. (1994) as well. In their prime-picture study they varied the modality of the prime and whether or not it was phonologically related to the picture. If the prime had to be read aloud, phonological inhibition of picture naming was found. If the prime had to be read silently, no phonological inhibition was observed. Why?

Discussion of the phonological inhibition effect

For single word production the phonological inhibition effect might be explained in terms of three different processing accounts. The first account proposed a postselective inhibition (MacKay, 1987). The elements that are generated in the preceding utterance will be inhibited for a while. The gain of such a mechanism might be that the inhibition prevents items from keeping residual activation, which might lead to erroneously producing them again. In an experimental situation, such as the identical condition, the identical verb at the first picture got inhibited after naming. This inhibition affects the processing of the acoustically related prime and the response to the identical verb during the presentation of the second picture. Both suffer from the inhibition of the related first picture verb. As a result, phonological inhibition occurs.

The second account locates the effect at the level of phonological segment retrieval (Peterson, Dell, & O'Seaghdha, 1989; O'Seaghdha, Dell, Peterson, & Juliano, 1992). Interference results from the activity of a residual trace of the prime. The residual activation of the phonological segments of the prime (for example, the 'g' in *pig*) can create competition with the target's segments (like the 'n' in *pin*). According to Peterson et al., interference arises when two segments compete to fill the same slot in a word frame. The critical feature of the model is its prediction that a high-frequency target word is more likely to be encoded before the residual activation from the prime decays. From that it follows that high frequency targets should be more sensitive to interference from related primes than low-frequency target words.

The third account assumes an interaction of phonological facilitation and word (lemma) competition (Colombo, 1986; Slowiaczek & Hamburger, 1992). According to Colombo, phonological facilitation is, as in cohort theories, due to residual activation of phonemes of the prime, which helps the access of the target phonemes (‘p’and ‘i’ of the prime *pig* would speed up the phoneme access of *pin*). But, according to Colombo, the initial phonemes of a prime also activate word candidates that begin with these phonemes (like *pin, pig, plug*) at the lexical (lemma) level. These similar lexical units compete with each other. As in the second mentioned account, the inhibition effect should be frequency dependent. Only high frequency targets become inhibited by phonologically related low frequency prime words. The reason for assuming this frequency effect is an empirical finding (for example, Segui & Grainger, 1990). The authors used an orthographic priming paradigm in which lexical decision targets were preceded by related or unrelated prime words. It was observed that related primes that are lower in frequency than the target inhibited the lexical decision on the target. In contrast, when the prime was higher in frequency than the target, no effect or facilitation was found. The assumption of locating the inhibitory effect at a word (lemma) level was supported by results of Slowiaczek and Hamburger (1992). They used an auditory single-word shadowing task. Participants heard a word and had to repeat it as fast as possible. The acoustic target was preceded by either an acoustically or visually presented prime. The prime could either be phonologically related or unrelated to the target. Important here is, that the prime could also be a nonword, which is assumed not to have a (lemma) representation in the lexicon. Hence, nonwords should not interfer with targets at this level. This is what was found: Nonwords did not produce a phonological interference effect.

In addition to the just discussed assumptions about the nature of phonological inhibition in word-picture relations, the effect should be discussed from the perspective of having two pictures and a prime involved in the present experiments.

Location of phonological inhibition in a picture-word-picture context

The nature of the phonological inhibition during the present identical utterance condition may not have to do primarily with the distractor-target relation, but with the combination of three elements: The response to the first picture, the distractor, and the response to the second picture. In order to get an idea about what goes on in the speech system, the identical condition should be compared with the elliptical condition. In both cases, the overt production of the verb during responding to the first picture might leave a trace in the speech system. Wherever this trace might be, at the lemma level or the word form, it should look the same for identical and elliptical responses. The overt production of the first verb should also influence the prime processing in the same way in both cases, given the first verb influences the prime at all. So far, the activational state of the speech system might be the same for identical and elliptical responses. Now we have to consider the production of the second verb. Let us assume that during both utterances the lemma of the verb becomes active. From this it would follow that the phonological inhibition during identical production cannot be located at the lemma level. Otherwise we should find phonological inhibition in the elliptical utterance as well. This was not the case. Therefore, the inhibition should be located at the phonological level. This account is in line with the Peterson et al. account of phonological inhibition. Their frequency assumption could be translated in such a way that generating a just produced verb for a second time is fast (as is the production of high-frequency targets). The faster produced second naming then gets hampered by the phonologically related prime. Of course, this line of argument is speculative at the moment and has to be tested in future research.

Modality specific differences for the identical utterance conditions

As just discussed above, the phonological inhibition in the *respond-respond variant* may be explained by the idea that the production of the first picture leaves a trace in the speech system. According to Wheeldon and Monsell (1992) this trace may be facilitatory in nature. The authors argue that the locus of repetition priming (facilitation) is the semantic/conceptual processing, not the form. The form level was excluded as potential locus of repetition priming because the effect was not present if homophones preceded the target. The trace could speed up the processing of the second - identical - verb production, maybe during lemma access. The prime will also leave a trace, according to Peterson et al. (1989), at the form level. Here, the activation of the prime element competes with the to be produced target. The amount of competition should be frequency or speed dependent. The competition only takes place if the target is fast and/or higher in frequency than the prime. Given a constant trace and a constant decay of activation of the prime, a fast target will reach the form level earlier than a slow target. A fast target, therefore, will be confronted with a more highly activated inhibitory trace of the prime. This leads to competition. As argued above, in the identical condition, the previously produced - identical - verb speeds up the second processing of that verb, which then competes with the related prime. The result is phonological inhibition.

The pattern of results is different in the *listen-respond variant:* Instead of inhibition we observed phonological facilitation. Why? If we follow the just outlined argumentation the explanation might be as follows: When a participant has to listen to the first picture description this also leads to some sort of trace within the speech system. This 'listen' trace might look different than the 'respond' trace. It might be less strong than the 'respond' trace. If it is less strong, the target will not become speeded up so much by the acoustic input of its identical partner. A slower target will reach the form retrieval process at a moment where the prime's inhibitory trace has already decayed. Therefore, a situation for phonological competition cannot come up. Instead, the prime pre-activates the shared phonemes of the target. As a result naming gets facilitated.

Difference between the experiments regardless of priming effects

In addition to different phonological effects between the experiments, a second salient finding for the identical condition was that the experiments differed in mean naming latencies. The listen-respond variant is about 60 ms faster than the respond-respond variant. This effect cannot be due to two different subject groups, because it was not found for the ellipsis condition. The difference also can not be related to specific 'identical utterance' effects, because it is also observed in the complete utterance (see below). It cannot be a strategic effect of participants, such as 'simply ignore the first picture in the identical/complete block', because the effect is also found in the complete condition when it was intermixed with elliptical responses. Here, in order to use the right response format, the first picture had to be attended to. The only explanation I have at hand has to do with task specific cognitive load. In the listen-respond variant participants could fully concentrate on the onset of the second picture, while passively listening to the first picture description, whereas in the respond-respond variant they were still busy with monitoring their own speech when the second picture came in. In general, informal feedback from participants in the listen-respond variant was more pleasant, indicating that they had no problems with the task. Participants in the respond-respond variant complained about subjective impressions of time pressure. The cognitive load difference was not present during the elliptical productions. In all elliptical cases a context match had to be carried out between the just presented and the preceding picture, involving the same amount of cognitive load.

2.8.3 Comparison for the complete utterance conditions

Because the complete utterances served as a control for the elliptical and identical response conditions, they will be briefly described. The analysis were carried out in a 2x2x2 ANOVA with 'experiment' as between factor and 'block' and 'relatedness' as within factors. The mean response latencies are shown in Figure 2.12, separately for the phonological and semantic distractor conditions, and separately for each block.

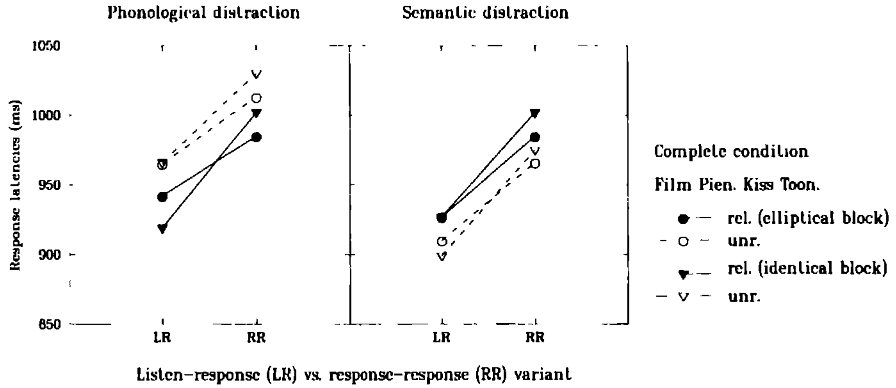


Figure 2.12 Mean response latencies (ms) for the complete utterance conditions for responding to the second picture of the listen-respond and the respond-respond variant.

The factor 'block' involved different utterance instruction for the identical or elliptical cases, but did not differ for the complete condition. Therefore no difference for the two complete conditions should be obtained between blocks. If differences were observed, this would be evidence for non-linguistic specific strategic effects.

The *semantic distractor condition for complete responses* revealed a significant main effect 'experiment'. $F_1(1,54) = 6.3$, $Ms_e = 39035$, $p = 0.01$; $F_2(1,44) = 10.9$, $Ms_e = 17988$, $p < 0.01$. Responding in the listen-respond variant was 67 ms faster than in the respond-respond variant. The main effect 'relatedness' was significant. $F_1(1,54) = 20.5$, $Ms_e = 1404$, $p < 0.01$; $F_2(1,44) = 12.7$, $Ms_e = 1871$, $p < 0.01$. This showed significant semantic inhibition. Neither the main effect 'block' nor interactions were significant.

The statistical analysis for the *phonological condition during complete production* revealed a significant main effect 'experiment'. $F_1(1,54) = 3.97$, $Ms_e = 50317$, $p = 0.05$; $F_2(1,44) = 9.0$, $Ms_e = 17077$, $p = 0.04$. Latencies in the listen-respond variant were 65 ms faster than in the respond-respond variant. The main effect 'relatedness' was also significant. $F_1(1,54) = 26.7$, $Ms_e = 2106$, $p < 0.01$; $F_2(1,44) = 25.4$, $Ms_e = 1801$, $p < 0.01$. This showed significant phonological facilitation. The interaction 'block x experiment' was significant across items, but not across subjects. $F_1(1,54) = 1.2$, $Ms_e = 8705$, $p = 0.26$; $F_2(1,44) = 15.3$, $Ms_e = 581$, $p < 0.01$. The significant interaction across items indicated that the main effect 'experiment' was greater in the identical block (74 ms) than in the elliptical block (45 ms). None of the other interactions were significant.

As can be seen from the figure and the statistical comparison, the data give a homogeneous pattern of phonological facilitation and semantic inhibition, without any unexpected side effects. The finding of having shorter latencies in the listen-respond variant than in the respond-respond variant has been discussed in the comparison section for identical production.

2.9 Conclusion

The comparison of the listen-respond and the respond-respond variant for elliptical utterance production clearly supports Klein's assumption of how speakers might produce ellipsis: Semantic and syntactic information should be taken over from the complete form. Only the phonological information should be reduced. Empirically, this assumption was tested by means of picture-word interference experiments. The observed finding of semantic interference is evidence for the speaker having the concept and the lemma active during the generation of ellipsis. The observed phonological zero effect showed that the form of the elided verb did not become activated.

In the second step, Klein's assumption of speaker-independency for the creation of the phonological reduction was investigated. Klein posits that if the context is the same, this should always lead to the same reduction rule. The reduction rule should be independent of who created the context (the speaker her- or himself, or the partner in the communication). To investigate this assumption, a second experiment was carried out that involved a change of speaker. This experiment was then compared with the first one, where the picture description was produced by the participants themselves. In the two experiments, the context remained constant (the content of the first picture description) which should lead to the same reduction rule: Evidence for semantic activation, and phonological reduction across the two experiments. This is what was observed. There was no difference between the experiments for elliptical utterance generation. This means that a change of speaker neither changed the semantic effect nor the phonological zero effect.

The phonological null effect supports the two-stage theory of lexical access (Levelt et al., 1991a, b; Roelofs, 1992). The referent noun is not selected for articulation. It should, therefore, not be phonologically active. In contrast, the phonological null effect is problematic for theories of cascaded processing (Dell & O'Seaghdha, 1991, 1992). The cascaded processing view predicts that an activated lemma automatically spreads activation towards its form, even if it is not selected for articulation. The observed semantic effect indicated that the lemma is activated. The lemma, therefore, should spread activation towards the phonological form, leading to an interaction with overlapping prime phonemes. However, no phonological effect, either inhibitory or facilitatory, was observed.

The elliptical condition was accompanied by other utterance conditions. One of them was the complete condition. This complete description served as a methodological control for the other conditions. It successfully indicated that the power of the experiment was sufficient to interpret the ellipitical description results. The second one was the identical response condition. This condition should give some insights into the complex activational processes during repeated production of the same verb. In contrast to the elliptical case, the identical condition is sensitive to modality. But the exact nature of the processes during this utterance condition are far from known. Some ideas were proposed in order to look somewhat deeper into the picture-prime-picture combination with regard to phonological effects.

LEXICAL ACCESS IN THE PRODUCTION OF PRONOUNS

CHAPTER 3

In this chapter, lexical access during the generation of pronouns will be addressed. Similar to ellipsis, pronouns are a form of speech reduction. But, whereas the use of ellipsis involves a complete articulatory deletion, the use of pronouns still needs overt articulation of the reduced form. This difference might involve different internal planning processes for the two forms of reductions. Therefore, a comparison of the assumed lexical access processes for ellipsis and pronouns might be interesting (as will be discussed in Chapter 5).

As in the preceding chapter, the focus of the present one is on the question of whether a referent (in this case a noun) becomes phonologically activated during the generation of its reduction (its pronoun). The phonological activation of the noun may give important information about lexical access during pronoun generation. First, a phonological activation of the noun during pronoun generation would indicate that the lemma spreads activation towards its phonological form, even if it is not selected for overt naming. This spreading of activation maybe predicted by theories that assume a cascading-like spreading of activation (Dell & O'Seaghdha, 1991, 1992). In contrast, strict two-stage theories would not predict phonological activation of words that are not selected for overt naming (Levelt et al., 1991a; Roelofs, 1992a, b). Second, a phonological effect would show indirectly that the lemma of a noun becomes accessed during pronoun generation. This assumption can be made because current theories of language production posit that phonological activation can only take place after lemma activation. As will be discussed in section 3.2, the lemma access is assumed to be necessary for pronoun generation, because the lemma carries the syntactic information of gender needed for the selection of the appropriate pronoun. However, this assumption has not yet been tested empirically. A series of experiments will be described that investigate this issue. But before I come to the experiments, I briefly introduce aspects of the speaker's discourse processing, because discourse is relevant to the generation of pronouns.

3.1 Discourse processing influences message encoding

The generation of an utterance, regardless of whether it is complete or reduced, initially involves the transformation of communicative intentions into preverbal conceptual messages. Levelt (1989) assumes that this conceptualization involves two processes: macro- and microplanning (see also Chapter 1). During *macroplanning* the speaker's intention is encoded. Among other things, he or she has to select the information to be expressed. An example for selection is a situation where the speaker is addressed by another person with a sentence like *I think your boyfriend is not happy with you.* The speaker can decide to go on with the current topic 'the boyfriend' by answering, for example, *Yeah, he is complaining about my mental absenteeism.* The speaker can also decide to make a shift towards a new topic, signaling that at the moment he or she is not interested in talking about potential personal problems, as in the reply *Um, more and more people get hallucinations these days.* During *microplanning* the speaker fits the message into the current discourse. For instance, in the first answer the speaker decided which of the selected concepts should be expressed as 'given' or 'old' information by selecting the pronoun *he* instead of a repetition of the entry *boyfriend* (Chafe, 1976; and see below).

As can be seen from the examples, the two phases of message encoding are context dependent: First, for a selection of the appropriate concepts the speaker must take into account his or her own interests and those of the addressee. Second, for an effective assignment to the discourse situation, the speaker has to keep track of what has been said already and what should be addressed as new information. Context dependent selection and keeping track consist of several aspects which are the subject of extensive research in discourse comprehension and production (for reviews see Kintsch, 1994; Clark, 1994). Here I will address only those aspects of the discourse processing that concern *how* and *why* a speaker generates reduced linguistic forms, such as pronouns. Following Levelt (1989) one relevant basic mechanism of discourse processing is the speaker's perspective taking in the ongoing discourse. In addition, the speaker's perception and 'book-keeping' of the ongoing continuous change in the discourse seem to be important for communication. Furthermore, the speaker's ability to create a common ground with the partner of the conversation is important. These aspects are addressed next.

3.1.1 Perspective taking

The speaker takes *perspective* in the discourse by choosing an anchoring point. Usually, the anchoring point is speaker-centric (Levelt, 1989). That means that personal, spatial, and temporal relations in the discourse context are seen from the speaker's point of view. He or she usually sees the world as *Me here and now.* Following Bühler (1934) this anchoring is standardly called *deixis.* Examples are the use of *you* and *me* in person deixis, *here* and *there* in place deixis, and *yesterday* or *tomorrow* in time deixis. An additional form of deictic expressions is the so-called discourse deixis. In the example *You have asked me* *that 42 times already!* the entry *that* refers to a question the partner in the communication asked before. By using *that* in this case, the speaker points to an earlier part of the discourse.

Deictic perspective taking can be seen as a device for the speaker to chose concepts during his or her message encoding that helps the listener follow the discourse. By using deictic terms, the speaker creates a coordinate system with his or her anchoring point as the starting position. During the utterance, the speaker then leads the listener through this coordinate system of time and space dimensions[1].

3.1.2 The discourse record

Although the speaker-centric perspective may be relatively constant, the discourse situation is a continuously changing situation. The speaker must keep track of this change by remembering what has been said already. Levelt (1989) called this 'book-keeping' of a *discourse record.* In the discourse record, the currently available discourse information is stored. This record consists of short-lived information, comparable to short-term memorization (following Baddeley, 1986). But the record's content can also be stored more deeply in long-term memory. The speaker continuously refers to this record during the encoding of his or her message (for a review of the interplay between short-term memory and long-term memory during text processing I refer to Ericsson and Kintsch, 1995).

One major reason to refer to the discourse record is to create a *coherent discourse* structure. Coherence means, that cooperating partners in a conversation choose a particular discourse *topic*, such as 'the weather' or 'the financial situation of Ph.D.-students', and talk about it by building a hierarchy of goals and subgoals (Grosz & Sidner, 1985). With regard to 'financial problems' a goal may be to inform about the possibility of getting tax breaks. A subgoal of this might be to talk about a specific person in the department who successfully managed to receive tax breaks in a more or less legal way. In a coherent discourse one part of the hierarchy is addressed by one speaker. Then it is taken over by the partner of the communication, who can either address the same issue again or can go up or down one level in the hierarchy.

---

[1] For details on the linguistic analysis of deictic systems, I refer to Jarvella & Klein, 1982; Levinson, 1983 For details on the acquisition of the deictic system, see Deutsch and Pechmann, 1978. For a review of psycholinguistic approaches to how a speaker produces spatio-temporal deictic utterances see Levelt, 1989, 1994. For a discussion on alternative perspective taking, such as intrinsic or absolute perspective see Brown & Levinson, 1993; Levelt, 1996, Levinson, 1996.

3.1.3 Accessing fragments of a common discourse model

The previously mentioned hierarchies may be seen as 'mental models' (Johnson-Laird, 1983) of persons, facts, their relations, and their properties. The *discourse model* can be based on what the speaker believes to be shared knowledge (shared between listener and speaker). A speaker's utterance that addresses a particular discourse model normally motivates the addressee to access his or her corresponding discourse model. The shared representations can then be elaborated more deeply during the ongoing discourse.

Depending on the complexity of such a discourse model, the speaker cannot attend to it completely during talking. According to Levelt (1989), the discourse model is stored as world knowledge in long-term memory, from which details can be accessed one at a time during conversation. Evidence for phases of information retrieval and macroplanning on the one hand, and phases of fluent speech production on the other hand, came from pause analyses during monologues (Henderson, Goldman-Eisler, & Skarbek, 1966; Butterworth, 1980; Beattie, 1983). The authors observed rhythmic alternations between speech phases that included frequent and long pauses and phases where nearly no pauses were present. The authors assumed that the observed hesitation phases were due to information retrieval processes (but see Power, 1983, for the view that the observed phases were not rhythmic but random).

The selection of a particular element from a complex discourse model may be necessary because the speaker and the listener do not have the capacity to make the whole complex discourse model available. The cognitive skill of selectively attending in order to circumvent capacity limitations has been addressed extensively in classical psychology (James, 1890; Miller, 1956; Broadbent 1958). A different reason for selection, the selection-for-action hypothesis, has been proposed more recently (Allport, 1987, 1993; Neumann, 1987, 1992). According to this view, selection is not due to capacity limitation but it is necessary because action can be carried out only sequentially. Allport (1987) gave the example of *picking apples*. Many fruit are within reach, and clearly visible, yet for each individual reach of the hand, for each act of plucking, information about just one of them must govern the particular pattern and direction of movements. The availability of other apples, already encoded in the brain, must be in some way temporarily decoupled from the direct control of reaching the target. The same may hold for speaking: Because the speaker can only talk about one thing at a time, he or she has to select a particular fragment from the discourse model.

3.1.4 The accessibility status of a referent

The selectively attended fragment of the discourse model is called the *focus* of the current discourse. The definition of what the focus of the discourse is has to do with the *accessibility status* of the particular referent (Levelt, 1989, p.144ff). According to Levelt, the speaker has to determine the accessibility of a referent. The assignment of the accessibility status takes place during microplanning. The speaker can derive the accessibility status of a referent from his or her discourse record by estimating, a) whether the referent is accessible to the listener or not; b) whether the referent is in the discourse model of the listener, and c) whether the referent is in the listener's focus. These three aspects can be depicted in the form of embedded sets (see Figure 3.1).

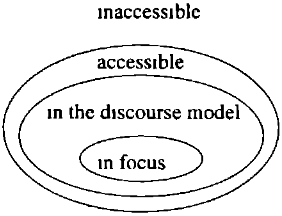


Figure 3.1
Accessibility status of referents in discourse (after Levelt, 1989, p.145)

The assignment of accessibility is relevant for further grammatical encoding of the speakers message. The following examples illustrate the syntactic consequences of the four types of accessibility depicted in the figure. First, if the speaker assumes a referent to be inaccessible in the listener's discourse model he or she would encode this referent as 'indefinite', as in *Marion is having some trouble with a dog.* Here the referent *dog* is introduced as 'new' (Prince, 1981; Chafe, 1976). Second, if a referent is not in the discourse model, but the speaker assumes that the addressee can infer the referent and make it accessible, the speaker can use a 'definite' expression, for instance, *Marion is having some trouble with the dog.* Here, the speaker might assume that the listener already focused on *Marion* in a particular visual scene. Therefore, the listener can probably also access the referent *dog*, because it is also visible. Third, if the speaker assumes that the referent is in the current discourse model of the listener he or she can mark it 'definite', as in, *Gosh, the dog is really big.* But, because the referent is in the discourse model it has no news value. This marking of 'old' information may then receive prosodic deaccentuation. Empirical evidence for this comes, for example, from a study by Fowler and Housum (1987). The authors analyzed monologues from radio programs with regard to the first and second mentioning of a word. They found that in a second naming of the referent, the duration of the word was shorter, and less loud (see also

MacWhinney and Bates, 1978; Marslen-Wilson, Levy, and Tyler, 1982; Terken, 1984). Finally, if the speaker assumes that the listener has the target referent in the current discourse focus, the speaker will deaccent this referent and make it definite, because it is also in the discourse model and accessible, as in the third example. But in addition, the 'in focus' feature leads to the reduction of the referring expression. An example of a reduction is the use of a pronoun, as in *The dog is really big. It even frightens our cat.*

These four examples show a referent's status concerning the accessibility features 'accessible', 'in the discourse model', and 'in focus'. The speaker might assign the accessibility status by marking the referent with the particular features. The assignment might take place in terms of a procedural IF/THEN rule, such as, *IF the referent is in focus, THEN assign (+)'in focus', ELSE assign (-)'in focus'* (Appelt, 1985; Levelt, 1989, for a review). Relevant for the present interest in pronoun generation is the idea that the accessibility status 'in focus' leads to the generation of pronouns, at least in Germanic languages, such as English, Dutch, and German.

3.1.5 Centering Theory

The previously mentioned accessibility features might provide each referent in the message with an index, for example, *(+)'in focus'* (Levelt, 1989). This index will be taken into account by the speaker during grammatical encoding. The preverbal features of the message encoding turn into linguistic devices, such as reductions. By doing so, the speaker provides the listener with cues about where to attend and where to locate the target referent in his or her own discourse record. Accessibility features, therefore, improve the discourse coherence between listener and speaker.

A formal way to describe how these linguistic devices affect discourse coherence has been developed within the framework of *centering theory* in computational linguistics (Joshi & Weinstein, 1981; Grosz, Joshi, & Weinstein, 1983; Gordon, Grosz, & Gilliom, 1993). According to Centering Theory, the conceptual referents of an utterance serve as *discourse centers* which are linked across utterances to create a coherent discourse. In Centering Theory, an utterance in a discourse can contain two kinds of centers, a backward-looking center (C_b) and a set of forward-looking centers (C_f).

The *backward-looking* center determines how the current utterance is to be incorporated into the preceding discourse. It is intended to capture the role of 'given' information (Prince, 1981; Chafe, 1976) and corresponds roughly to the linguistic notion of the 'topic' of a sentence (Joshi & Weinstein, 1981). For example, in the utterance *Markus kissed Annette* the backward-looking center is

Markus (if Markus was mentioned before), which is also the topic of the sentence. According to Centering Theory, each utterance has only one backward-looking center. This center must be realized linguistically. Joshi and Weinstein (1981) define appropriateness-rules that state what the linguistic form of an utterance must be to fit into the on-going discourse context. For example, an utterance is appropriate if the backward-looking center is identical to one of the forward looking centers of preceding sentences, and if the C_b is linguistically realized as a pronoun rather than a name or a definite description (Grosz et al., 1986). Gordon et al. (1993) found empirical evidence for this assumption in reading experiments. Reading time is elongated when the backward-looking center is linguistically realized as a name rather than a pronoun, as in *Markus was happy. Markus/he(*C_b*) kissed Annette.*

The *forward-looking centers* provide potential links to the subsequent utterance. In the example *Markus kissed Annette*, the forward-looking centers are *Annette* and *Markus*, which correspond roughly to the linguistic notion of 'focus' (Joshi and Weinstein, 1981). The members of a set of forward-looking centers can be ranked according to their prominence. Psychologically, prominence might reflect differences in accessibility from short-term memory (Gordon et al., 1993). Linguistically, prominence is thought to be expressed by factors such as surface position in the utterance, grammatical role, and pitch accent (see Levelt, 1989, p. 149ff.). This ranking is thought to provide default values for the interpretation of pronouns by the listener: The first pronoun in a sentence is usually interpreted as referring to the highest ranking member of the forward-looking center of the previous utterance (Gordon & Scearce, 1995). The prominence ranking might also give a default rule for the speaker's message planning, such as 'Take the most prominent forward-looking center of the previous discourse, put it in the first place of the next utterance, and reduce it to a pronoun'. Empirical evidence for this assumption during language production came from Marslen-Wilson, Levy, and Tyler (1982). The authors analyzed a speaker's telling of a story. They found that a speaker first introduces and establishes a highly focused entity (one forward-looking center), such as the actor of a particular scene. Once the topic is defined, the speaker tends to realize this topic with less marked forms, such as pronouns and ellipses. In contrast, the speaker tends to realize non-focused entities with more marked forms, such as definite descriptions.

During on-going discourse, a speaker, therefore, marks entities of his or her message as being backward- or/and forward-looking centers. These preverbal markers lead to a specific linguistic realization of the message that has the goal of making the discourse "well-formed" (Joshi & Weinstein, 1981).

This short introduction to the speaker's ability to establish and maintain reference in ongoing discourse should give an idea of why a speaker uses reduced forms, such as pronouns. Reductions seem to serve as cues for the listener to get optimal access to the referring entry in his or her own discourse model. By improving access to the listener's discourse model, linguistic devices, such as reductions, improve the coherence of the discourse structure. This coherence is, according to Centering Theory, the goal of every speaker. The theory describes how a speaker can create this coherence by looking forward and backward in time. The Centering Theory posits that speakers apply procedural rules during message encoding in order to fit a current utterance to the ongoing discourse. The speaker does so by transforming preverbal conceptual devices into linguistic devices. The question now is how the transformation of discourse dependent messages into their linguistic forms might look.

3.2 Lexical access of pronouns

The encoded message, as discussed above, specifies which concepts should be expressed. In addition, each concept is marked with its accessibility format. It, therefore, carries discourse dependent information about how it can be expressed linguistically, for example, as full noun or as a reduction. This discourse-marked message serves as input for the grammatical encoding stage (Levelt, 1989, and see Chapter 1). It activates corresponding lemmas which in turn deliver the syntactic information needed to generate an utterance that matches the required discourse constraints. The steps that are involved, from lemma access to pronoun generation, are outlined next.

3.2.1 The syntactic structure of lemmas

Following Levelt (1989), each lemma gives access to syntactic information that is relevant to form the surface structure of an utterance. For example, the lemma of an item carries the information about the syntactic category of this item, such as whether it is a noun, a verb, or an adjective. The lemma also carries diacritic parameters (Levelt, 1989, p. 191), such as tense, person, and number information. In addition, a lemma carries information about the lexical item's required grammatical functions. For instance, the verb *hand* requires a subject, a direct object, and an oblique object. In the sentence *Barbara proudly hands her pharmacology diploma to Dieter* these grammatical functions are fulfilled by the phrases *Barbara*, *her pharmacology diploma*, and *Dieter*. The lemma *hand*, therefore specifies that the conceptual agent (X) should be assigned the grammatical role of subject, the so-called theme (Y) should be realized as a direct object, and the recipient (Z) as an oblique object in the (as yet to be created) surface structure of the utterance. The lemma *hand* also allows for another argument (X, Y, Z)-to-syntactic function mapping, which figures in a sentence like *Barbara proudly hands Dieter her diploma*, where *Dieter* becomes the syntactic function of an indirect object.

This shows that the temporal ordering of the phrasal constituents of a surface structure during grammatical encoding can vary. This variation is determined by different factors, such as the saliency of the concepts in the message (cf. Bock & Warren, 1985), which in turn schedule the lemma retrieval, and through that the word order and/or grammatical functions in a sentence. In addition, the ordering is restricted by the grammar. For example, the positions of subject (S), verb (V) and object (O) phrases differ in SVO- or SOV- languages. The ordering of entities into specific positions in the sentence can be seen as filling empty S-, V-, and O-labeled slots with appropriate items. According to the so-called frame-and-slot models of language production, the retrieved lexical information, that is, the lemma and its syntactic information, is assigned to the slots of corresponding syntactic frames (see for example, Dell, 1986; but see also Fromkin, 1971, 1973; MacKay, 1972; Fay & Cutler, 1977; Shattuck-Hufnagel, 1979, 1987; Bock, 1982). Dell assumes that these frames are created by linguistic rules that only allow for acceptable combinations of items at the syntactic level (Dell, 1986, p. 286).

Following this view, a preverbal message that marks a particular concept as *(+)'in focus'* in the current discourse might lead to a specification of a 'pronoun-slot' instead of a 'noun-slot' at the syntactic level which has to be filled with the appropriate pronoun. How an appropriate pronoun may become selected is addressed in the next section.

3.2.2 From lemma to gender access

In languages with grammatical gender, the appropriate pronoun that must be selected for the syntactic frame has to be of the same syntactic gender as the noun to which it refers. How does the pronoun receive the right gender? Following current theories of language production, this gender information is directly accessed by the lemma[2] (Roelofs, 1992a, b, see also Chapter 2;

---

[2] This gender assignment differs between languages In English *chair* will activate neuter gender, in German it will activate masculine gender (der Stuhl), in Dutch it will assign the common gender (de stoel) A native speaker will simply acquire this assignment For an extensive discussion on whether gender assignment is random (without rules, for example, from the non-native speaker point of view) or systematic (involving phonological, morphological, or semantic constraints), and whether it is stored in the lexicon or has to be computed each time it has to be used, I refer to Van Berkum, 1996, in press (see also Zubin & Kopke, 1981, Corbett, 1991) In this thesis gender is assumed to be stored as grammatical information in the lexicon, following Roelofs, 1992a, b, and Schriefers, 1993

Schriefers, 1993; Jescheniak & Levelt, 1994). According to Schriefers (1993) the gender is used to determine the correct article in noun phrases, such as <u>*de*</u> *stoel* (the$_{\text{common}}$ chair), or <u>*het*</u> *bed* (the$_{\text{neuter}}$ bed). Schriefers generalized Roelofs' (1992a, b) model of lemma access to gender access. Following Dell (1986), Schriefers assumes that after the selection of the appropriate gender, the corresponding syntactic frames can be filled, which in turn leads to the phonological encoding of the utterance. He tested this assumption by means of picture-word interference experiments. Participants had to name pictures, and were asked to ignore distractor words. In the study, distractor words were presented that had either the same grammatical gender as the target picture word or had different grammatical gender. The author argues that when the distractor noun activates gender information different from the target noun's gender (at an critical SOA), the selection threshold for the correct gender will be reached later. This would lead to a delayed selection of the correct gender information. This, in turn, would delay the filling of the corresponding syntactic slots. The delay should result in longer naming latencies for noun phrases in conditions with gender-incongruent distractors than in conditions with gender-congruent distractors. The experimental data show exactly this result: a gender-incongruency effect. Schriefers attributes this effect to competition between gender information carried by the target lemma and gender information carried by the distractor noun lemma on the level of syntactic processing.

3.2.3 From lemma to gender to pronoun access

By extending Roelofs' theory of lemma access and Schriefers' idea of gender access to the pronoun case, the assumption can be made that pronouns should be accessed via the activated referent noun lemma and the noun's gender. In this view, as with gender information, pronoun information is assumed to be stored in the lexicon, possibly in terms of pronoun lemmas. Figure 3.2 shows the assumed nature of pronoun access outlined so far. Depicted is the generation of the German pronoun *sie* in the example *Die Blume ist rot. Sie wird blau. (The flower is red. It turns blue).* The example is chosen because it resembles the utterance format that was used in my own experiments described below. German was chosen because it has a clear grammatical gender assignment for things (as do, for example, Spanish and French; see Garnham, Oakhill, Ehrlich, & Carreiras, 1995). An object with feminine grammatical gender, such as *die Rose* (the rose), has to be referred to by the pronoun *sie* (she). A masculine noun, such as *der Klee* (the clover), has to be referred to by *er* (he), and a neuter noun, such as *das Veilchen* (the violet) by *es* (it). The gender of object names in German is grammatical because it does not refer to any biological gender. By using depicted objects that have no biological gender a potential confound of conceptual gender access and syntactic gender access can be excluded.

Figure 3.2 depicts the following process: The concept of BLUME (flower) becomes activated for the second time. This might happen in situations where, for instance, a previously presented picture had been described, which now re-appears. In the discourse record, this re-appearance is registered in terms of an accessibility assignment of (+)'*in focus*'. The (+)'*in focus*' feature should activate a procedure to produce a linguistically reduced form for the current concept. How can this be realized? According to Roelofs (1992a, b), the activated concept automatically activates its corresponding lemma *blume*. According to Schriefers (1993), the lemma, in turn, leads to the activation of its gender. So far, there seems to be no difference from the overt generation of the noun 'Blume'. However, because the discourse record signals the feature (+)'*in focus*', the selection of the noun for overt generation should be prevented and the pronoun should be activated instead. This discourse dependent switch in the processing mode is depicted in the figure as a gate between the connections of lexical gender and pronoun information. If the accessibility status in the discourse is (+)'*in focus*', the gate is open and allows the access from the gender information to the pronoun information. This leads to the selection of the appropriate pronoun *sie* that might fill a syntactic frame, according to Dell (1986).

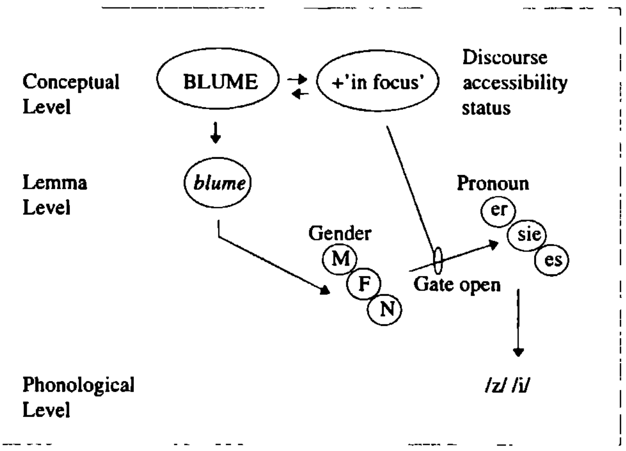


Figure 3.2 A lexical access view of the generation of the pronoun *sie* in *Die Blume ist rot. Sie wird blau.* (The flower is red. It turns blue.) M = masculine, F = feminine, N = neuter.

The selected pronoun activates its corresponding phonemes at the phonological level. The phonemes /z//i/ will be pronounced. Alternatively (not depicted in the figure), if the discourse record signals that the current concept is not accessible by previous discourse [with *(-)'in focus'*] the gate is closed, and access to pronoun information is not possible.

It has to be mentioned here that the proposed gating function is speculative in nature. But it can explain how the speech system generates different kinds of speech output given nearly identical visual input ('nearly' here relates to an experimental situation in which an identical object re-appears, but is depicted in a different color). Roelofs (1992a, b, see also Chapter 2 for details, and Chapter 4 for alternative models) hypothesizes how our speech system becomes active during picture generation: Confronted with a picture, for instance of a flower, the system activates its concept, lemma, and phonemes so that the noun *flower* will be named overtly. Roelofs simulated these activation processes in a computational network model that consists of nodes and connections between nodes, representing stored conceptual and syntactic information. The assumption of having such representations of stored knowledge in long-term memory at hand is generally accepted in current pycholinguistic research. But if a naming process would use stored knowledge only, the network would always deliver the same output given the same input (with regard to object naming). In our particular case, confronted with a sequence of two pictures that depict the same object, the network would always come up with the overt noun generation of this object, such as *The flower is red. The flower is blue.* As discussed above, this is not what speakers do. The speaker's skill in continuously implementing variable discourse information into the planning of his or her utterance should interact with the stored knowledge. The proposed gating mechanism enables such an interplay between procedural rules and stored conceptual and linguistic information within the same network architecture. It should be noted here that the gating mechanism is not the focus of the experimental section of this chapter. The present experiments focus on the first proposed step of pronoun generation, the lexical access of the referent. However, the gating mechanism will be addressed in Chapter 4 again, where it is implemented in a computational network.

3.3 The present research question

Is the phonological form of the referent noun active during pronoun generation? The current experiments focus on the nature of the lexical access of the referent noun. As in the preceding chapter on the generation of ellipsis, the question addressed is whether the lemma becomes activated at all during the generation of the reduction. As discussed above (see Figure 3.2), lemma access is needed in order to make gender information available, which in turn is crucial for selecting the appropriate pronoun. The proposed access mechanism for pronouns is an assumption. To the best of my knowledge, no alternative hypothesis exists. But this does not mean that no alternative mechanism can be conceived that does not involve lemma access during pronoun generation. The investigation of lemma access, therefore, can be seen as an existence proof of the proposed lexical access assumption.

One way to investigate lemma access would be to look at semantic activation of the referent noun at the moment the corresponding pronoun is uttered. As discussed in Chapter 2, a semantic effect could be interpreted as having the lemma active. However, as is discussed in the section about the first main experiment (section 3.6), there were problems finding appropriate materials for testing this issue.

A second way to investigate whether the lemma is active during pronoun generation is to look at the activation of its phonemes. Here, the underlying assumption is that phonological activation of a word is only possible via the lemma (Dell, 1986, 1988; Levelt, 1989; Levelt et al., 1991a; Dell & O'Seaghdha, 1991, 1992; Roelofs, 1992a, b). A phonological effect would indicate that the lemma has been accessed. Of course, this argument does not hold the other way around: A phonological null effect is neutral about the activation state of the lemma. This is the case because according to a strict two-stage theory (Levelt et al., 1991a; Roelofs, 1992a, b), a lemma could be active without becoming selected for naming, preventing further phonological activation (see below).

The aim of the present pronoun experiments is to look at what happens to phonological information of the referent noun during the generation of the corresponding pronoun. The phonological activation of the noun was investigated by comparing activation processes of two different utterance formats: Overt noun generation vs. pronoun generation.

As discussed in section 3.2, I assume that the conceptual activation of FLOWER in an utterance such as *The flower is red* is the same as for *it* in *The flower is red. It turns blue.* During the generation of the adjacent sentence pair, the concept FLOWER is therefore assumed to be accessed twice. The next processing step concerns lexical access of the noun. As has been addressed in the introduction (Chapter 1), according to theories of language production (Garrett, 1975, 1988; Stemberger, 1985; Dell, 1986, 1988; Levelt, 1989; Roelofs, 1992a, b), lexical access involves two separate stages: Lemma access and phonological access.

With regard to ***overt noun generation***, the theories agree on the assumption that during the naming of a picture, the lemma of the picture's noun becomes active (see Figure 3.3, left). This activated lemma, in turn, activates its corresponding phonemes, because they are needed for articulation.

With regard to ***pronoun generation***, the theories differ in their prediction of whether an activated lemma would lead to phonological activation of its phonemes or not (see Figure 3.3, right). On the one hand, theories that assume cascading spreading of activation would predict phonological activation of an activated lemma, even if it is not selected for naming. On the other hand, strict two-stage theories of lexical access would predict no phonological activation in the pronoun case.

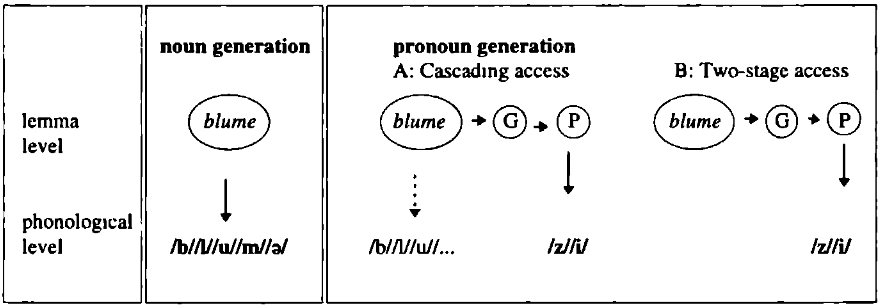


Figure 3.3 Phonological activation of the noun lemma *blume* (flower) during overt noun generation (left) and during the generation of its corresponding pronoun (right, A and B) A According to theories that assume a cascading spreading of activation, the noun's phonemes should become co-activated with the pronoun phonemes B· According to a two stage theory, phonemes of the noun lemma should not become active G = Gender, P = Pronoun.

According to cascading activation-spreading models of lexical access (Dell, 1986, 1988; see also Stemberger, 1985; Dell & O'Seaghdha, 1991, 1992; Harley, 1993; and see also Chapter 4 for more details), an activated lemma automatically spreads activation towards its phonological form. Because of the cascading spreading of activation, the phonological activation starts to increase immediately after the lemma gets accessed. Because a cascading model does not have a distinction between an activated and an selected lemma, the phonological form of the noun becomes available even if the lemma has not to be named overtly. As depicted in Figure 3.3 (right, A), the activation process of pronoun generation might involve the following steps: First, the lemma gets accessed by conceptual information. It, in turn, accesses its gender, and because of the discourse information the gender information is passed to the pronoun node. But in parallel to the accessing of the pronoun, the noun lemma also directly spreads activation to the phonological level, where its phonemes are co-activated with the pronoun phonemes. The co-activation of the noun's phonemes might become inhibited over time if the pronoun phonemes become more highly activated (see Chapter 4, Simulation 3, for a more detailed proposal on solving co-activation).

According to a two-stage theory of lexical access (Levelt et al., 1991a, 1991b; Roelofs, 1992a, b; see also Chapter 4 for details), a distinction is made between the mere activation of a lemma and its selection for naming. According to Roelofs' model of lemma retrieval in speaking, the activation level of the target lemma node must exceed that of other activated nodes in the lexicon by some critical amount. Once this level of activation has been reached, the actual selection of this target lemma is a random event with a probability that is given by the 'Luce-ratio' (Roelofs, 1992b, p. 47). Only a selected lemma spreads activation towards the phonological form, and, in turn, becomes articulated. Because in the pronoun case the noun is not selected for naming, the noun's phonemes should not become activated. This process is depicted in Figure 3.3 (right, B). A phonological null effect, according to a two-stage model, cannot be interpreted as having no lemma activation at hand. It can only be interpreted as having no lemma selection available. If a null effect were observed, further research would be necessary to look more directly into lemma access in the pronoun case.

3.4 The experimental paradigm

The phonological activation of the referent noun during pronoun generation was investigated empirically by means of dual task experiments, using the so-called 'lexical decision during picture naming' paradigm. The paradigm was chosen for three reasons.

First, its cross-modal nature has been proven to be sensitive to investigate on-line activation during sentence comprehension in the domain of anaphora resolution (for example: Cloitre & Bever, 1988; Nicol & Swinney, 1989; Fodor, 1989; MacDonald & MacWhinney, 1990; Osterhout & Swinney, 1993; McDonald & MacWhinney, 1995; Love & Swinney, 1996).[3]

Second, the dual task paradigm has been successfully applied in speech production (Levelt et al., 1991a). In the Levelt et al. study, it proved sensitive to measuring activation processes of lexical access during picture naming. Most importantly, it indicated clear phonological effects during picture noun naming. The paradigm served as a tool to distinguish between the two theoretical approaches, the cascading view and the strict two-stage view. It therefore should

---

[3] For a critical discussion about the advantages and disadvantages of the paradigm see McKoon, Ratcliff, & Ward, 1994, Balota & Abrams, 1995, Nicol, Fodor, & Swinney, 1994

be a useful paradigm to compare lexical access during noun and pronoun generation with regard to these two theories.

The third reason for using the dual task paradigm was a methodological one. In the ellipsis experiments I used the picture-word interference paradigm because it proved sensitive during single noun naming (for example, Schriefers et al., 1990, observed an approximately 60 ms phonological facilitation at a short SOA). However, during the more complex ellipsis experiments, it turned out that the size of the effects for the overt complete control condition was smaller. Therefore, I chose the dual task paradigm for the pronoun study, because the Levelt et al. (1991a) study showed a clear phonological effect with an effect size that was slightly bigger than that of the picture-word interference paradigm (at a short SOA Levelt et al. observed an approximately 80 ms difference between the unrelated and the phonologically related condition).

The 'lexical decision during picture naming' paradigm has the following characteristics. Its major purpose is to investigate on-line processes during speech production. The elicitation of normal speech is achieved by having the participants to describe pictures. It should be stressed that the undisturbed process of picture description is the default task of the participant. Only during a proportion of the experimental trials does this process become disturbed by the presentation of an acoustic probe stimulus. In these cases, the participant is asked to postpone describing the picture and to react to the presented probe first. The probe stimulus could either be a word or a pseudoword. A push button response is carried out to determine whether or not the stimulus is known as a word to the participant. This results in a 'yes' response if the presented probe is a real word, such as 'dog'. It results in a 'no' response if the probe is a pseudoword, such as 'dolk'.

The lexical decision latencies are supposed to reflect the current state of the naming process due to interactions of the encoding processes of the two tasks. Following Levelt et al. (1991a) the naming process consists roughly of four encoding stages: The visual encoding of the picture, the semantic and syntactic encoding, the phonological encoding, and the articulation. The lexical decision regarding the acoustic probe consists of two stages: Phonological and semantic/syntactic encoding. By the proposed serial encoding, Levelt et al. assume that a lexical decision always takes place after the semantic/syntactic encoding of the acoustic stimulus, regardless of the nature of the probe-picture relation.[4]

---

[4] The assumption of Levelt et al. that lexical decision involves semantic encoding (that is, at least lemma access) in all probe conditions is stronger than that made by theorists from the word recognition domain. Here, no explicit assumption has been made with regard to when the lexical decision can be executed. As, for example, Marslen-Wilson and Zwitserlood (1989) point out, the

The critical case during a dual task performance is one in which the corresponding encoding stages of the two tasks overlap in time, and probe and picture are related to each other. A semantic relation between probe and picture is defined as being members of the same semantic category (Miller & Fellbaum, 1991). For example, a probe word *dog* is presented together with a picture of a *cat*. A phonological relation is defined as sharing the same word onset (in the present case onset and nucleus of the first syllable). Based on the findings of Levelt et al., specific assumptions can be made about the to be expected effects:

Overlap of activation between the semantic stages of picture and probe would delay the lexical decision on the semantically related probe in comparison to an unrelated probe. Levelt et al. (1991a, p. 137) assume a Stroop-like character of inhibition: A semantically active item in the naming channel will interfere with the lexical decision for the meaning-related probe word. The tendency of the participant to react to the naming target instead of to the lexical decision probe has to be inhibited. More recently, Roelofs (1992a, b) postulated that the semantic inhibition may be due to competition between the two semantically related items at the syntactic encoding stage (see Chapter 4 for details).

Overlap of activation between the phonological stages for both picture-word and probe-word phonemes would delay the lexical decision on the phonologically related probe in comparison to an unrelated probe. The account of Levelt et al. (1991a, p. 136) for the phonological inhibition was that the partial phonological representation in the naming channel boosts phonological competitors to the lexical decision probe if it is compatible with their phonological representations. This is assumed to be the case when the lexical decision probe is phonologically similar to the picture name. The phonological competition has to be resolved, which takes time. The competition, therefore, delays the phonological encoding of the lexical decision probe in the phonological condition.

For the present experiments, it is important to notice that a phonological effect, according to the interpretation of Levelt et al. (1991a), can only be present, if the phonological stages of the probe noun and the picture noun overlap in time. A phonological effect during lexical decision, therefore, would indicate that the

---

localization of the lexical decision effect is dependent on the prime-target relation. A semantic relation of prime and target could clearly locate an observed priming effect at the lexical level (lemma/word level), which means that a lexical decision would be executed after semantic/ syntactic encoding. In contrast, however, a phonological relation of prime and target could either be located at the lexical level (lemma/word level) or at a sublexical level (phoneme or feature level). A lexical decision, therefore, could also be executed immediately after sublexical encoding, and does not necessarily involve lexical (lemma) access at all (see also Slowiaczek & Hamburger, 1992, Lively, Pisoni, & Goldinger, 1994).

phonemes of a picture's noun lemma are activated. Exactly this information is needed to test whether the phonemes of a noun lemma are activated if it has to be realized as a pronoun. An observed phonological activation would invite two interesting interpretations. First, it would show that the noun lemma is indeed active during pronoun selection. Second, an observed phonological effect would favor the cascading view of spreading of activation as opposed to the two-stage theory.

The following experimental section consists of four parts. First, pretests are described that were run to prepare the materials and determine the timing of stimuli in the main experiments. Second, the Levelt et al. (1991a) study which was carried out in Dutch, is replicated in German for the short SOA condition (SOA = 100 ms). Third, the pronoun generation experiment is described. And finally, a control experiment is reported that replicated the pronoun findings and checked for a potential residual activation that might also explain the results obtained during the first pronoun experiment.

3.5 Preparation of materials

The purpose of the pretests was to construct a homogenous set of materials for the main experiment with respect to mean picture naming latencies and mean lexical decision latencies. In the pretests, 87 line drawings of objects were combined with acoustic lexical decision probe words and pseudowords. The replication of the Levelt et al. (1991a) study involved the same picture-probe relations as the original study. The authors investigated four probe conditions: identical probes, semantically related, phonologically related, and unrelated probes. Therefore, each of the present 87 pictures was paired with four probe words. In addition, it was paired with four pseudowords. The pretests for picture naming and lexical decision were carried out separately.

3.5.1 The picture naming pretest

Method

Participants: The participants were 15 citizens of Münster, Germany, recruited by advertising in a newspaper. Their native language was German. They were paid for participating in the experiment. Their ages ranged from 18 to 35 years.
Materials: The visual stimuli were 87 white-on-black line drawings of target objects. Some drawings were selected from the picture pool available at the Max-Planck Institute, some created by the author. The drawings depicted objects with an equal number of German neuter, masculine, and feminine gender names.

From a norming study with 10 German speaking participants it was known with what noun the depicted objects were named spontaneously. The names included between one and three syllables. The frequencies of the target words' lemmas were determined using the MANNHEIM corpus of the CELEX database from the University of Nijmegen (Burnage, 1990). The MANNHEIM corpus includes 6 million German word tokens (see result section for frequency information).

Design: The mean response latencies and naming errors for each target picture were recorded. Each participant saw each picture once. The order of the items was random and different for each participant.

Apparatus: The experiment was run on a Hermac 386 SX computer. The pictures were presented on a Nec MultiSync 4FG screen. The participants' speech was recorded using a Sennheiser HMD 224 Headphone-Microphone combination and a SONY TCD D3 DAT recorder. Reaction times were measured by a voice key.

Procedure: Participants were tested individually. They were seated in a dimly lit laboratory at comfortable viewing distance in front of a monitor. They received a booklet with the instructions and the pictures. Below each object was the word printed that the participants in the norming study had spontaneously used most frequently as the object name. The participants were asked to use that name for the object. They were asked to name the picture as fast and as accurately as possible. As soon as they indicated that they had read the instructions and had studied the picture names, the experiment started with 20 practice trials and went on to the target pictures.

Trial structure: First, a fixation cross appeared for 500 ms. After an inter-stimulus-interval of 300 ms the picture was presented for 800 ms. The participants named the picture, and the response latencies were measured. The inter-trial interval was 2000 ms. Each experimental session took about 20 minutes.

Results

Three types of responses were categorized as errors: Incorrect naming of the object, a time-out (i.e., no response within 1500 ms after picture onset), and mouth clicks that triggered the voice key without involving speech onset for picture naming. Mean response latencies were analyzed for each item.

As selection criteria for the target pictures to be used in the main experiment, I chose a mean response latency lower than 1000 ms, a standard deviation smaller than 200 ms, and an error rate of no more than 3 out of 15 responses. An additional constraint was that the four acoustic probe words - which were combined with the picture later on in the main experiment - fulfilled their selection criteria (to be mentioned below) in the lexical decision experiment. This procedure resulted in 48 suitable target pictures with a mean lemma frequency of 205 out of 6 million (SD 388) and mean response latencies ranging from 643 to 985 ms.

3.5.2 The lexical decision pretest

Method

Participants: The participants were the same as in the response experiment.
Materials: The acoustic stimuli were 4x87 (= 348) nouns and 348 pseudowords. They consisted of one to three syllables. The word probes had the same grammatical gender as the depicted objects in order to avoid possible gender-incongruency effects (Schriefers, 1993). The semantic relation was defined as being members of the same semantic category. For instance, the picture name *Hund* (dog) was combined with the probe *Fisch* (fish). The phonological relation was defined as phoneme-overlap of the onset and nucleus of the first syllable of the probe and the picture name. For example, the probe *Husten* (cough) was phonologically related to the picture name *Hund* (dog). In addition, there was an identical acoustic probe *Hund*, and there was an unrelated probe (for example, *Tisch* (table) for the picture *Hund*). Semantic probes had no phonological relation to the noun of the semantically related picture. Phonological probes had no semantic relation to the picture names. The picture-probe relation was established by the unanimous judgment of five raters.
Design: The experiment consisted of 50% words and 50% pseudowords randomized per participant. The mean latencies for lexical decision were measured for four probe conditions.
Apparatus: The same machines were used as in the picture naming pretest. The probe words and pseudowords were spoken by a male speaker and recorded using a Sony 59ES DAT recorder. They were digitized with a sampling frequency of 16kHz and stored on the hard disk of the computer. The participants heard the stimuli on a Sennheiser HMD 224 headphone-microphone combination. Reaction times were measured by a yes-no-response on a push-button box.
Procedure: Participants were tested individually in a dimly lit laboratory. In the instructions, the participants were asked to decide as fast and as accurately as possible whether the acoustic probe was a word or a pseudoword by pushing the word or pseudoword button. As soon as they indicated that they had read the instruction, the experiment started with 20 practice trials. The experiment consisted of 6 blocks. Each block started with 4 warm-up trials, followed by 116 target trials.
Trial structure: First, a warning tone of 50 ms was presented over the headphones. Then, after 300 ms the acoustic probe was presented. The inter-trial interval was 2000 ms. The participants made the lexical decision and the

reaction time was measured. Including short pauses between the blocks, an experimental session lasted about 45 minutes.

Results

Errors were incorrect responses and time-outs (i.e., no response within 1500 ms after stimulus onset). Mean reaction times were analyzed for each item and probe condition. Criteria for selecting the acoustic probes as appropriate for the main experiment was a reaction time of faster than 1100 ms, a standard deviation smaller than 200 ms, and an error rate of no more than 3 out of the obtained 15 decisions per item. An item (i.e., picture plus four probes) could only be included in the main experiment if all five stimuli met the pre-specified criteria. This procedure resulted in 48 suitable probe words per condition. The mean lemma frequencies, mean word length, and mean lexical decision latencies of the acoustic probes per condition are depicted in Table 3.1.

Table 3.1 Mean lemma frequency (out of 6 million), mean word length (in ms), and mean lexical decision latencies (in ms) of the 48 acoustic probes per condition. The standard deviation is depicted in brackets.

	Probes identical	phonological	semantic	unrelated
Lemma frequency	205 (388)	192 (337)	169 (450)	178 (305)
Word length	720 (111)	697 (106)	723 (138)	748 (108)
Lexical Decision	848 (85)	899 (84)	899 (95)	897 (104)

The selected items matched on mean *lemma frequency* and mean *word length*. A one-way repeated measures ANOVA on 'word length' (4 probe conditions for every picture) was not significant, $F(3, 141) = 1.8$, $MS_e = 11674$. The mean duration for the 4x48 pseudowords was 743 ms (SD 134).

The mean *lexical decision latencies* for the selected 4x48 word probes were analyzed by a one-way repeated measures ANOVA on 'reaction times' (4 probe conditions for every picture). It revealed significant differences between probes, $F_2(3, 141]) = 3.7$, $MS_e = 8124$, $p = 0.01$. This result shows that lexical decision latencies on identical probe words were significantly faster than those for the other conditions. This might be due to the fact that words which can easily be represented in pictures are more concrete and, therefore, faster to access (Paivio, 1966; Paivio, Yuille, & Madigan, 1968; Bock & Warren, 1985). However, because in the main experiments each item has an individual control baseline (see section 3.6.1), this difference should not matter.

3.6 First main experiment: Lexical decision during single noun naming

The first main experiment was carried out to find the appropriate SOA timing and materials for the pronoun study. In addition, it served as a German replication of the study by Levelt et al. (1991a). As in the original study, participants were asked to name the pictures by using single nouns. On some trials, an acoustic stimulus was presented to which the participants had to make a lexical decision. For each target picture there were four acoustic word probes: an identical probe, a phonologically related one, a semantically related one, and an unrelated one. Each target picture was also paired with four pseudoword probes. This resulted in 50% potential word, and 50% pseudoword decisions in the lexical decision task. Of interest were only the word-response latencies to the four kinds of word probes.

3.6.1 The baseline experiment

Of interest in the main experiment was the amount of interference of related probes in comparison to unrelated probes. For this comparision the critical test probes were matched on frequency of usage, word length, and mean response latencies in the pretest. Although the probes were matched, they could still vary greatly between individuals. To control for this inter-individual variation, each test probe was made its own control by means of a baseline-experiment. Lexical decision latencies on the same acoustic probes as in the main experiment were collected again for each participant, without involving picture naming. The baseline-experiment was carried out about one week after the main experiment, in order to keep learning effects low. Apparatus, procedure, and trial structure were the same as in the lexical decision pretest.

3.6.2 Method

There are some methodological differences from the original study. One major difference is that the present experiment investigated a different language (German instead of Dutch). It therefore used different items. A second difference is that it tested only one SOA. This SOA (100 ms, presenting the picture first, and the probe second) corresponds roughly to the short SOA (mean = 73 ms) of the Levelt et al. study. But in the present study, the SOA was constant for all participants, whereas Levelt et al. matched the SOA to individual picture recognition latencies. Third, in the present experiment, a lexical decision had to be made in 50% of the trials, in contrast to a 33% proportion in the Levelt et al. study.

Participants: The participants were citizens of Münster, Germany, recruited by advertising in a newspaper. Their native language was German. They were paid for participating in the experiments. Their ages ranged from 18 to 37 years. The experiments were carried out with 52 participants. 6 of them had to be excluded from further analysis because they made more than 20% naming or lexical decision errors in the main experiment.

Materials: The visual stimuli consisted of 144 white-on-black line drawings. 96 pictures were filler and practice pictures. 48 were the critical test pictures selected as a result of the pretest; they are given in Appendix C. The acoustic stimuli were the 4x48=192 pre-selected critical word probes - including an identical, phonologically related, semantically related and unrelated probe for each of the critical pictures. Furthermore, there were 4x48 pseudowords for the main session along with 35 word and 35 pseudoword probes for practice, all selected from the first pretest set.

Design: In this experiment, participants named a series of pictures. In order to keep participants in a response mode, 50% of the trials were single noun naming without an acoustic distractor. This was realized by presenting 96 filler pictures. During the other 50% (96) of the trials, an acoustic probe was presented. The proportion of word and pseudoword presentation was also 50% (48). The 48 pseudowords were paired both with filler pictures (24) and with critical pictures (24). The pairing of pseudowords and critical pictures occurred only during a second presentation of the pictures. The repetition of pictures was done in order to keep the number of pictures low. But each critical picture-word pair was always presented first, so that repetition effects of picture presentation do not play a role in the data analysis. The order of trials was randomized, with the constraint that no more than two word probe presentations followed each other, and no more than five acoustic probe presentations. Each participant was presented with a different order.

The picture-probe pairings were counterblanced as follows: For 12 of the 48 critical pictures, a participant would receive the identical test word, for another 12 the phonological probe, and so on for the semantic and unrelated probe words. The pseudowords were presented with a randomly selected 24 of the 48 critical pictures, and with 24 filler pictures. The 52 participants were randomly divided into four groups of 13. All 13 participants in each group received the same picture-probe pairs. But the pairings were rotated among the four groups: The first 12 critical pictures were combined with the identical probes for the first group of 13 subjects, with semantic probes for the second group, with phonological probes for the third group, and with unrelated probes for the fourth group. The second 12 critical pictures were paired with different probes across groups. The same holds for the remaining pictures. This balancing procedure guaranteed that every critical picture was paired with every critical probe condition and every participant was tested in every probe condition. The procedure resulted in a one-factor repeated measures design with the four probe conditions as within factor for the subject and item analysis. The dependent variable is the differential score of the reaction times to the baseline minus the main experiment reaction times for each probe word.

Apparatus: The same equipment was used as in the pretest and in the baseline experiment.

Procedure: During the practice session, each participant saw a sequence of 144 pictures on the monitor. The sequence was randomized for each participant. The picture's name that was most frequently used in the pretest was printed below the object. The participant was asked to study the pictures and to use the printed name in the experiment. After familiarization with the picture and its name, the participant could press a button and the next picture appeared. In the second practice session, the pictures were presented on the monitor without the printed name, again randomized for each participant. The participant had to name the pictures. In a third practice session, the lexical decision task was introduced. The procedure was the same as in the lexical decision pretest. 15 acoustic words and 15 pseudowords were presented and the participant had to make a lexical decision. In a fourth practice block, the dual task situation was trained. The block consisted of 80 trials. 40 filler pictures were presented without acoustic probes, 20 pictures were paired with word probes, and 20 pictures with pseudowords. The trial timing was the same as in the main experiment and is described below. The participant was asked to name the picture as quickly and as accurately as possible. But the participant was told to delay the response when hearing an acoustic probe and to make the lexical decision first. The complete practice session lasted about 45 minutes and was followed by the main experiment. The main experiment consisted of 2 blocks of 96 trials each. Each block was preceded by 3 warm-up trials.

Trial structure: A fixation cross was presented for 500 ms, followed by an inter-stimulus interval of 300 ms. Then the picture was shown for 1500 ms. In the critical lexical decision trial, the probe was presented 100 ms after picture onset. The voice key was activated from picture onset for 1800 ms. The push button device was activated from the onset of the acoustic signal for 1500 ms. After the participant made a response, the next trial began. The inter-trial interval was 2.5 seconds. Each experimental block lasted about 8 minutes.

3.6.3 Hypothesis

By comparing the baseline lexical decision latencies with those during the dual-task situation, a general increase in reaction times was expected during the dual task session. This increase is due to the higher complexity of the dual task, which leads to an increase of cognitive load that slows down the lexical decision during a picture naming task. The size of the cognitive load effect, in addition to possible learning effects, can be seen in the reaction time difference in the baseline and main experiment for the unrelated probe condition, because no other linguistic effects were expected here.

For related probes, however, in addition to the general cognitive load effect, specific linguistic effects were expected. If there is specific interference for the semantically related probes during picture naming, this interference should add to the general inhibition. This means that the difference between baseline and dual-task performance for semantically related probes should be bigger than the difference for the unrelated probes. The same logic holds for phonologically related and identical probes. The hypotheses about specific linguistic effects were as follows (see the introduction of the experimental paradigm for details):

An overlap of activation between the semantic stages of picture naming and lexical decision should delay the lexical decision on the *semantically related probe* in comparison to an unrelated probe due to semantic competition.

An overlap of activation between the phonological stages for picture description and lexical decision should delay the lexical decision on the *phonologically related probe* due to phonological competition.

With regard to lexical decision on *identical probes*, interference was expected for SOA = 100 ms relative to the unrelated condition. According to Levelt et al. (1991a, p. 138), this interference is explained by both the Stroop-like semantic interference effect and the phonological competition due to the picture's boost of phonological alternatives to the probe.

3.6.4 Results

For a comparison, Figure 3.4 (top) depicts the results of Levelt et al. (1991a) for the short SOA condition. Figure 3.4 (bottom) shows the results of the present replication, which are also presented in Table 3.2. The figure displays mean lexical decision latencies for the participants on the critical test probes (I, P, S, U) in the baseline session and main session. The statistical analyses to be reported are based on differential scores. They were obtained by pairwise subtracting from each participant's reaction times in the baseline session the reaction times for the main experiment. Because of the longer lexical decision latencies in the dual task main experiment, this subtraction resulted in negative values. The mean differential scores are also presented in Table 3.2. Missing values were defined pairwise. That is, whenever a participant's reaction time score in the baseline or the main session was missing, the differential scores was also treated as a missing value. The analyses are based on the remaining data.

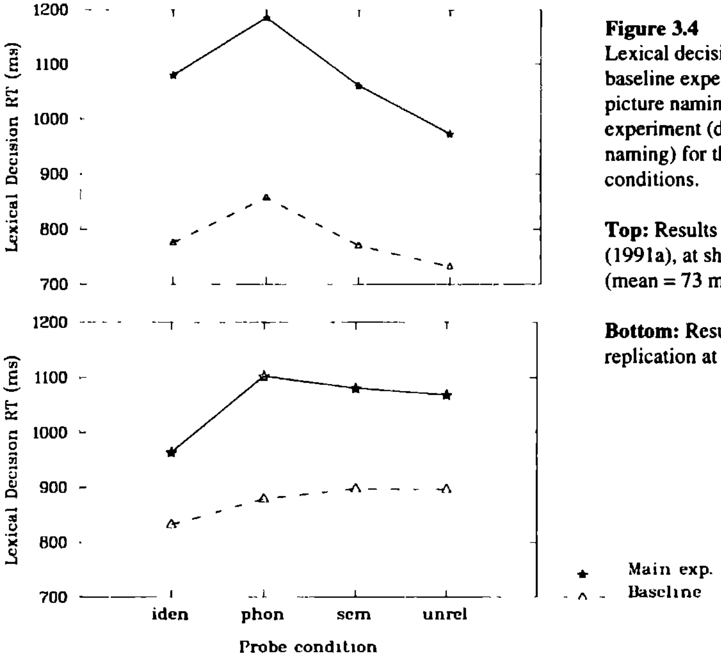


Figure 3.4
Lexical decision latencies of the baseline experiment (without picture naming) and the main experiment (during picture naming) for the four word probe conditions.

Top: Results of Levelt et. al (1991a), at short SOA (mean = 73 ms).

Bottom: Results of the present replication at SOA = 100 ms.

Table 3.2 Lexical decision latencies (in ms), differential scores (in ms), standard deviation (SD) of the differential scores, and percentage of missing values for the baseline and the main experiment for the four word probe conditions.

	Probes identical	phonological	semantic	unrelated
Baseline	832	875	896	889
Main session	968	1092	1086	1068
diff (baseline - main)*	-135	-217	-190	-178
SD	132	105	110	96
% Missing	8	20	16	11

* Any apparent inaccuracies in means are the result of rounding.

A missing value in the baseline session occurred when no lexical decision response or an incorrect one was given for an item, or when the reaction time was longer than 1500 ms. A missing value in the main experiment arose if no lexical decision was made, if an incorrect decision was made, or if the lexical decision was longer than 1500 ms. In addition, incorrect naming or speech onset before the onset of the acoustic probe were defined as missing. The percentage of missing values is also presented in Table 3.2.

The assumption behind using differential scores is that the baseline and the main session experimental effects are additive. That means that fast baseline test probes should not behave differently from slow baseline items. They both should be equally sensitive to the experimental manipulation in the main experiment. However, there are reasons to believe that the additivity assumption might not hold. For example, a constant SOA of 100 ms between picture and probe presentation might have a different effect for slow and fast probes. Generally, fast baseline items might be encoded faster than slow probes in the main experiment. As a consequence, the time course of picture naming might affect the encoding of fast probes differently in contrast to slow items, regardless of specific linguistic relations between pictures and probes. This difference is no problem for the interpetation of the differential scores if fast and slow items are equally distributed across conditions. However, there might be problems in interpreting the differential scores when the baseline for the four probe conditions differ.

As can be seen in the table, the baseline for the identical probes was the fastest. This was already the case in the pretest data. In the present experiment, the identical condition revealed the smallest effect, indicating that fast items might be less sensitive to experimental manipulations than slow items. This result, therefore, might cast doubt on the additivity assumption. However, the additivity assumption was supported by the results of the phonological condition. This condition revealed the greatest differential score, although its baseline is slightly faster than those of the semantic or unrelated condition. Therefore, I assume that no differential scores are artifacts of differences in the baseline. Instead the differential scores were supposed to mirror specific linguistic effects for each probe-picture relation.

Errors: The error analysis was carried out on the arcsin transformed error proportions (see Winer, 1971). The error proportions per probe condition represented the sum of errors of main experiment and baseline. An ANOVA for repeated measurements was performed with the factor 'probe condition' as within factor for the subject and item analysis. The main effect of probe condition was significant, $F_1(3,135) = 9.9$, $Ms_e = 0.074$, $p < 0.01$; $F_2(3,141) = 5.8$, $Ms_e = 7459$, $p < 0.01$[5]. Planned pairwise comparison between the unrelated (U) probe condition as the baseline and the semantic (S), phonological (P), identical (I) probes indicated a significant difference between the U and the P probes only, $F_1(1, 45) = 13.4$, $Ms_e = 0.13$, $p < 0.01$; $F_2(1, 47) = 5.06$, $Ms_e = 0.33$, $p = 0.03$. Participants made more errors in the phonological condition than in the unrelated condition. Because this condition also revealed the longest decision latencies in the main experiment, a speed accuracy trade off can be rejected. In

---

[5] In repeated measurement analysis with $df > 1$, an assumption for ANOVA is that the variances of treatment-differences are homogenous Instead of testing whether this assumption is violated, it was assumed that it is violated and the p-values were adjusted following the Greenhouse-Geisser lower bound correction Only these p-values are reported here (Kirby, 1993).

order to exclude that the higher proportion of errors was due to specific items and not due to a specific linguistic phenomenon in the dual task situation, a separate error analysis for the baseline experiment was carried out. The error proportions in the baseline experiment were 1% (I), 3.6% (P), 4.2% (S) and 2.9% (U) across subjects. An ANOVA of repeated measures on the arcsin transformed error proportions showed no differences between the probe conditions in the baseline experiment, $F_1(3,135) = 1.3$, $Ms_e = 0.003$; $F_2(3,141) = 0.73$, $Ms_e = 0.075$. This result indicated that the observed higher proportion of errors in the phonological condition of the dual task experiment was not due to item-specific artifacts, but could be interpreted as having a systematic linguistic interference at hand in the dual task situation

Differential scores: Repeated measurement ANOVAs were performed on the differential scores (with the four probe conditions as within-factor) for subjects and items. The main effect of 'probe condition' was significant, $F_1(3,135) = 11.3$, $Ms_e = 4659$, $p < 0.01$; $F_2(3,141) = 9.0$, $Ms_e = 7459$, $p < 0.01$. Planned pairwise comparisons were carried out in order to locate specific differences (Kirby, 1993). The unrelated condition can be considered as a baseline for the evaluation of the S, P and I probe conditions. The planned comparison of unrelated vs. semantically related probes revealed no significant difference, $F_1(1,45) = 1.0$, $Ms_e = 6770$; $F_2(1, 45) = 0.3$, $Ms_e = 10023$. The planned comparison of unrelated vs. phonologically related probes was significant, $F_1(1,45) = 14.02$, $Ms_e = 4872$, $p = 0.01$; $F_2(1, 45) = 0.39.9$, $Ms_e = 12611$, $p < 0.01$. The phonologically related probes showed a differential score of -217 ms between baseline and main experiment (see Table 3.2), whereas the unrelated probes revealed a differential score of - 178 ms, resulting in a significant 39 ms difference between phonologically related and unrelated probes. The planned comparison of unrelated vs. identical probes was significant, too, $F_1(1,45) = 6.6$, $Ms_e = 12473$, $p = 0.01$; $F_2(1, 45) = 4.1$, $Ms_e = 19293$, $p = 0.05$. The identical probes (differential score = -135 ms) became on average 43 ms less inhibited than the unrelated probes (differential score = -178 ms).

In summary, phonological inhibition and identical facilitation but no semantic effect were found.

3.6.5 Discussion

The aim of the first main experiment was to find appropriate materials for the pronoun-study. In addition, the experiment served as a German replication of the Levelt et al. (1991a) study. According to Levelt et al. (1991a), inhibition was expected for the S, P and I probes in comparison with the U probes at short SOA.

The *semantic inhibition* observed by Levelt et al. (1991a) could not be replicated for German speakers. One reason for this might have been the variability in semantic relation between pictures and probes within the 48 item pairs. As mentioned in the pretest section, the semantic relation was selected by only using same category members, and excluding associative relations. That was done because other studies showed that same category distractors inhibited, but associated distractors facilitated, the responses (see LaHeij, Dirkx, & Kramer, 1990). We wanted to avoid this interaction of the two effects, and expected inhibition only. The prediction of inhibition was supported by an observed semantic inhibition using same category members in the material set (Meyer, 1996). However, by looking into the present means of individual picture-probe pairs, 24 revealed inhibitory effects in comparison to the unrelated probes, 24 showed facilitation. One account for this difference may be that facilitatory association might be possible for semantic relations. However, this issue has not yet been investigated for German. It has to been shown in future research whether the observed differences in the direction of the semantic effect are related to association in general or whether it is specific for individuals. This question seems to be a research topic of its own, and during the ongoing experimental work I made the decision not to go into the details of semantic relations any further. The present pronoun experiments will, therefore, focus on phonological activation processes.

The results for the *identical condition* showed significant facilitation. In contrast, Levelt et al. (1991a) found identical inhibition at a short SOA. The interpretation of Levelt et al. (p.138) was that identical inhibition at an early SOA might be due to both a Stroop-like interference related to semantic overlap of picture and probe, and phonological interference because of phonological overlap. The Stroop-like interference was explained as the participant's tendency to suppress the naming of the picture, which causes inhibition of the lexical decision of identical probes. The phonological inhibition of identical probes was explained as a boosting of phonological alternatives by the picture's partial phonological representation. Partially activated phonemes in picture naming were supposed to activate potential phonological competitors to the lexical decision probe.

In contrast, an explanation for pure identical facilitation found in the present study might be that both the semantic encoding and the phonological encoding of the probe become speeded up. With regard to semantic encoding, the identical picture pre-activates the meaning and the lemma of the probe which then can be recognized more quickly. With regard to phonological encoding, at the short SOA the picture has activated parts of its own form. This helps the form selection of the identical probe. It becomes facilitated. In contrast to the Levelt

et al. interpretation, no boost of phonological alternatives takes place - because there are no alternatives in the identical condition.
However, the different empirical results of the Levelt et al. study and the present experiment could also be due to methodological differences that were mentioned in the method section. One crucial difference might be the use of different proportions of identical test probes in both studies. Whereas in the original study 1/8 of 33% (4.1%) of the lexical decision trials included identical probes, in the present study it was 1/8 of 50% (6.3%). Although this difference is small, there is evidence that the proportion of related stimuli matters for the direction and the size of interference effects. This comes from empirical studies of picture naming (McEvoy, 1988) and spoken word recognition (Goldinger, Luce, Pisoni, & Marcario, 1992). The different proportion of trials per condition, therefore, might have affected the present results. Further empirical evidence is necessary in order to resolve the issue. This will not be addressed here, but see Chapter 4 for simulation results that support the facilitatory effect.

Most importantly, the *phonological condition* showed inhibition in both the present study and in the Levelt et al. (1991a) study. Levelt et al. explained the effect as follows: "A partial phonological representation is present during the phase of phonological encoding in the naming channel, and that representation will boost phonological competitors if it is compatible with their phonological representations. This will be the case when the lexical decision probe is phonologically similar to the picture name." (p. 136).
The results of the present study show that the phonological inhibition effect of the noun naming study of Levelt et al. can be replicated in German. The phonological inhibition effect can be explained by the fact that the picture's lemma and its form are both active and interfere with the form of the probe. The phonological inhibition effect may, therefore, be a good indicator of activation processes of a noun lemma during pronoun generation.

3.7 Second main experiment: Lexical decision during pronoun generation

This experiment was carried out in order to investigate whether the phonological form of the referent noun is active during the generation of its corresponding pronoun. As discussed in section 3.3 (see also Figure 3.3), a theory that assumes a cascading spread of activation predicts phonological activation, whereas a two-stage theory would not. These assumptions were tested using the dual task paradigm.

3.7.1 The extension of the dual task paradigm

The present experiment tests whether in a pronoun construction, such as *The flower is red. It turns blue,* the phonological form of the noun *flower* is still active when the pronoun *it* is generated by the speaker. But how to elicit a pronoun utterance in a picture naming experiment? The idea was to present a sequence of two colored pictures that depict, for example, a red flower first followed by the same flower in a different color (see Appendix E for an illustration). The participant's task was to describe such a sequence as *The flower is red. It turns blue.* Time-locked to the onset of the second picture, the acoustic probe was presented. The probe was either phonologically related to the object's name or unrelated.

In order to avoid a default pronoun generation, a noun control naming was added to the pronoun condition. Here again, a sequence of two pictures was presented. But now two different objects followed each other (see Appendix E). For instance, a red sun was followed by a blue flower. The desired description format was *The sun is red. The flower is blue.* Again, time-locked to the onset of the second picture, the acoustic probe was presented. The probe was either phonologically related to the second picture's noun or it was unrelated. Note that in a critical trial, the same object with the same color was presented as in the pronoun condition. It was paired with the same probes as in the pronoun condition. The only difference between the two utterance conditions was that in the noun condition the noun had to be produced overtly. Hence, in addition to being an anti-strategic control, the noun generation revealed important information: A comparison of overt (noun) and latent (pronoun) generation was possible with regard to a phonological inhibition effect.

3.7.2 Method

Participants: The participants were native speakers of German, which were recruited by advertising in a newspaper. They were paid for participating in the experiments. Their ages ranged from 17 to 38 years. The experiment was carried out with 32 participants. 3 of them had to be excluded from further analysis because they made more than 20% naming and lexical decision errors in the main experiment and in the baseline.

Materials: From the 48 pictures of the preceding study, 16 pictures were selected as critical pictures. They were paired with the same phonological and unrelated acoustic probes as in the preceding study. The selection was based on the desire to find an optimal baseline. Across items and subjects of the preceding experiment the selected unrelated and related probes did not differ in mean lexical decision latencies [for items: related probes = 893 (SD 73) ms, unrelated probes = 887 (SD 71)]. The material is listed in Appendix D. In addition, the pictured objects were counterbalanced with regard to grammatical gender (5 masculine, 5 feminine, 6 neuter). Instead of the white-on-black line drawing of the preceding study, the pictures were colored. Four colors were chosen so that they were easily recognizable: red, yellow, green, and blue.

Design: In this experiment, two utterance conditions of the second picture description (pronoun and noun generation) were compared with regard to two probe conditions (phonologically related and unrelated probes). To elicit a noun or pronoun, a sequence of two pictures was presented. In the noun condition two different pictures followed each other, in the pronoun condition the same object re-appeared again in a different color. The same critical target pictures (picture 2), as well as the same critical target probes were used in both utterance formats. For each of the four probe presentations per target picture, a color change from picture 1 to picture 2 was the same. Hence, within trials that use one particular target picture, differences of visual perception due to color change was excluded. The color changes were counterbalanced across target pictures, so that anticipation of color changes was not possible.

To keep participants in a naming mode, 50% of the trials were simple picture descriptions without an acoustic distractor. This was realized by presenting 64 filler pictures. Of these, 32 trials were noun conditions, 32 trials were pronoun conditions. During the other 50% of the trials (64), an acoustic probe was presented. The proportion of word and pseudoword presentation was again 50% each. This leads to 16 noun and 16 pronoun generation trials each for the word and the pseudoword conditions. The 32 pseudowords were paired with filler pictures (16) and with critical pictures (16). The order of trials was randomized with the constraint that not more than 2 word probe presentations followed one another. In addition, no more than 5 acoustic probe presentations and no more then 5 successive trials of the same utterance format were presented. Each participant obtained a different randomized order of trials.

Each critical "second" picture was presented twice to each participant. For a randomly selected set of 8 of the 16 critical pictures, a participant received the phonological and unrelated probe words in the noun condition. For the other 8 critical pictures the same participant received the related and unrelated probe in the pronoun condition. The 32 participants were randomly divided into two groups of 16. All 16 participants in each group received the same picture-probe pairs. But the items were rotated across the 2 groups: 8 critical pictures were combined with the related and unrelated word probes in the noun condition for one subject group, and in the pronoun condition for the second subject group, and vice versa for the other 8 critical items. This balancing procedure guaranteed that every critical picture was tested under every critical probe condition and every participant was tested in every probe condition. The design resulted in a two-factor repeated measures design with the utterance condition and the probe condition as within factors for the subject and item analyses, respectively. Items were defined as the depicted object of the critical second picture.

As in the preceding experiment, in addition to the lexical decision latencies during picture naming, baseline lexical decision latencies without picture naming were collected. The baseline experiment was run one week after the main experiment. The procedure for the baseline was the same as in the preceding experiment. The dependent variable was the differential score 'baseline minus main experiment' for lexical decision on each probe word.

Apparatus: The same equipment was used as in the preceding experiments (see pretest section).

Procedure: The practice procedure was basically the same as in the first main experiment. First, participants were trained to name the pictures by using single nouns. But in addition, the utterance conditions were introduced and trained on in 32 practice trials (16 noun generations, 16 pronoun generations). In the next practice session, the lexical decision task was introduced. The participants were asked to name the sequence of two pictures as practiced beforehand. But if a probe was presented in a trial, they were asked to postpone the naming process and make the lexical decision first. This practice block consisted of 80 trials: Embedded in 40 filler trials (picture description only), 40 acoustic probes (20 words and 20 pseudowords) were presented that were not used in the main experiment. The practice probes were counterbalanced across utterance conditions. The trial timing was the same as in the main experiment and is described below. The practice session lasted about 40 minutes and was followed by the main experiment. The main experiment consisted of 2 blocks of 64 trials each, preceded by 3 warming-up trials within each block.

Trial structure: A fixation cross was presented for 500 ms followed by an inter-stimulus interval of 300 ms. Then, the first picture was presented for 1500 ms. When a participant started to describe the picture the voice key was triggered, and 1500 ms after the voice key was triggered, the second picture was presented. This was done in order to keep the duration between onset of the first picture description and onset of the second picture constant. In the lexical decision trial, the probe was presented 100 ms after the onset of the second picture. The push button device was active for 1500 ms from the onset of the acoustic signal. After the participant made a response the next trial began. The inter-trial interval was 2.5 seconds. Each trial lasted about 7 seconds. Each experimental block lasted about 7 minutes.

3.7.3 Hypothesis

As in the preceding experiment, the basic assumption of this pronoun study was that the probe's activation interferes with that of the naming process. The lexical decision latencies, therefore, should probe the current state of the naming process. More specifically, the lexical decision times should indicate whether or not phonological information of a referent noun is available when a participant starts to generate the corresponding pronoun. A phonological inhibition effect would be evidence for having form information of the noun available during the generation of a pronoun for that noun. In addition, the effect would indicate that the lemma is accessed during pronoun generation.

It was also assumed that *overt noun generation* in connected speech behaves in the same way as overt single noun naming in the Levelt et al. (1991a) study and in the replication, described above. The phonological information of the noun must be available in overt generation, and will interfere with the phonological activation of the acoustic probe. Therefore, phonological inhibition is expected. If this effect was not observed in this control condition, the results of the pronoun condition could not be interpreted.

The outcome of the experiment for *pronoun generation* was open. According to theories that assume cascading-like spreading of activation, the phonological form of the referent should become active if the lemma is activated at all. According to two-stage models of language production, no phonological activation is expected (see Figure 3.3).

3.7.4 Results

The main results are shown in Table 3.3 and Figure 3.5. The table displays mean lexical decision latencies for the participants for the critical test probes (P, U) in the baseline session and main session for the two utterance conditions. The statistical analyses are based on differential scores. As in the preceding experiment, they were obtained by pairwise subtracting each participant's main reaction times from that of the same item in the baseline session for each utterance condition. The mean differential scores are presented in Table 3.3. Missing values were defined as in the preceding experiment and are presented in the table. The analyses are based on the remaining data, that is, when both baseline and main experiment decisions on each word probe were correct.

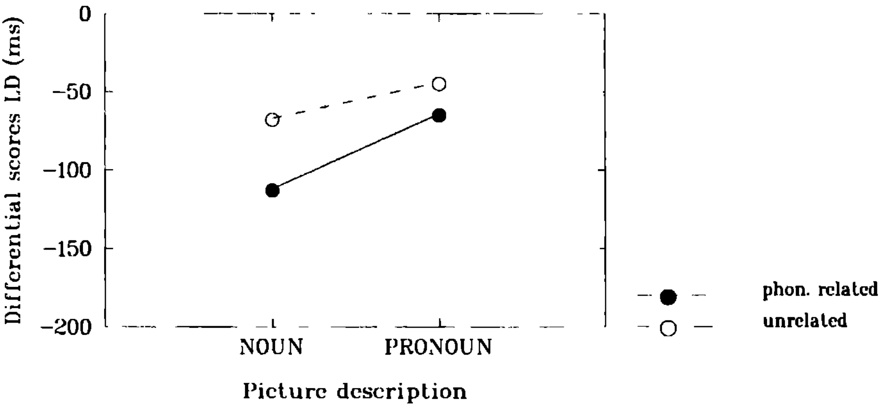


Figure 3.5 Mean differential scores of lexical decision latencies: Baseline experiment (without picture description) minus main experiment (during picture description) for the two word probe conditions for each utterance condition. First pronoun experiment. Negative values indicate inhibition.

Table 3.3 Lexical decision latencies and differential scores (in ms), standard deviation (SD) of the differential scores, and percentage of missing values for the baseline and the main experiment for the two word probe conditions for each utterance format (for subjects).

	Pronoun generation		**Noun generation**	
	phon.	unrel.	phon.	unrel.
Baseline	972	952	972	943
Main session	1036	995	1086	1007
diff (baseline - main)*	-64	-44	-112	-67
SD	107	150	108	149
% Missing	13	7	18	12

* Any apparent inaccuracies in means are the result of rounding.

As in the noun naming study described above, the assumption for using differential scores was that the baseline and the main session experimental effects are additive. As discussed before, different baselines might complicate the interpretation of differential scores. The mean reaction times for the baseline (see Table 3.3) look different. However, they are statistically equal. This was shown by an ANOVA for repeated measurements on the factor condition (4 measures per subject or item). $F_1(3, 84) = 2.2$, $Ms_e = 2657$, $p = 0.10$; $F_2(3, 45) = 0.72$, $Ms_e = 3985$.

Errors: The error analysis was carried out on the arcsin transformed error proportions in a 2 (utterance condition) x 2 (probe condition) repeated measurements ANOVA. Only the factor 'probe condition' was significant: $F_1(1,28) = 5.5$, $Ms_e = 0.048$, $p = 0.03$; $F_2(1, 15) = 8.6$, $Ms_e = 0.032$, $p = 0.01$. This result indicated a higher percentage of errors in the phonological condition. Because this condition also had showed the longest decision latencies in the main experiment a speed-accuracy trade off could be excluded. In order to exclude the possibility that the higher percentage of errors in the phonological conditions was due to item specific effects and not due to their specific behavior in the dual task situation, a separate analysis was carried out for the baseline lexical decision only. The error percentages across subjects in the baseline experiment were as follows: Pronoun generation 3.5% (related), 1% (unrelated); noun generation 3.8% (related), 4.3% (unrelated). The arcsin transformed error proportions of the baseline experiment were tested in a repeated measures ANOVA with the four conditions as within factors. Statistically, the probe conditions in the baseline did not differ with regard to error performance, $F_1(3, 84) = 2.2$, $Ms_e = 0.011$, $p = 0.10$; $F_2(3, 45) = 1.3$, $Ms_e = 0.02$, $p = 0.28$. These results support the assumption that the higher percentage of errors in the main analysis is due to specific phonological effects in the dual task situation.

Differential scores: Repeated measurements ANOVAs were performed on the differential scores with 'utterance condition' (noun and pronoun generation) and 'probe condition' (related and unrelated) as within factors, separately for subjects and items. Five extreme outliers were excluded from the analysis, following Tukey (1977; see also Kirby, 1993). The main effect 'utterance condition' was significant, $F_1(1,28) = 8.9$, $Ms_e = 4142$, $p < 0.01$; $F_2(1, 15) = 4.96$, $Ms_e = 2860$, $p = 0.04$. Lexical decision become more inhibited in the noun condition than in the pronoun condition. The main effect 'probe condition' was significant, too; $F_1(1,28) = 4.2$, $Ms_e = 7137$, $p = 0.05$; $F_2(1, 15) = 7.3$, $Ms_e = 2851$, $p = 0.01$. Lexical decisions were slower to related probes than to unrelated probes. The interaction was not significant, $F_1(1,28) = 1.2$, $Ms_e = 3652$, $p = 0.27$; $F_2(1, 15) = 1.6$, $Ms_e = 820$, $p = 0.22$. The lack of an interaction showed that phonological inhibition was found in both the noun and the pronoun condition. It also showed that the inhibition was statistically of equal size for both utterance formats. In summary, phonological inhibition was found in noun generation as well as in pronoun generation.

3.7.5 Discussion

The *noun condition* revealed phonological inhibition. This result is in line with the results of Levelt et al.'s study (1991a) and the data of the above mentioned semi-replication. It showed that the method of lexical decision during picture naming is sufficiently sensitive for investigating phonological activation during more complex utterances. As discussed above in the single noun naming study, the phonological inhibition can be accounted for by phonological competition between the picture's form and the probe's form (Levelt et al., 1991a). The active phonological form of the picture name boosts alternatives to the phonological form information of the probe word. These alternatives have to be suppressed in order to select the appropriate probe phonemes. This suppression takes time and, in turn, delays the lexical decision on phonologically related probes.

During *pronoun generation* phonological inhibition was also observed. This finding supports a cascading-spread-of-activation view of lexical access (Dell, 1986, 1988; Dell & O'Seaghdha, 1991, 1992). According to this view an activated lemma spreads activation towards its form, even if it need not be uttered overtly (see Figure 3.3). In contrast, the observed phonological inhibition contradicts the assumption made by the strict two-stage model. According to this model (Levelt et al., 1991a, Roelofs, 1992a, b) no phonological effects should be observed, because the pronoun condition should not lead to competition at all. This prediction was made on the assumption that the picture's noun should not become phonologically encoded at all during pronoun generation, because its lemma is not selected for overt naming.

In addition to providing insights into phonological processing of the referent noun during pronoun generation, the data indirectly showed that the lemma has become activated. This conclusion can be drawn because current theories of language production agree on the idea that phonological encoding can only take place if the corresponding lemma has become activated (Dell, 1986; Levelt, 1989; Levelt et al, 1991a; Dell & O'Seaghdha, 1991, 1992; Roelofs, 1992a, b).

An alternative explanation for the observed phonological inhibition is that it might be due to the participants' strategies to use latent (or internal) speech. The participant might plan the overt noun generation as a default, and switch to the pronoun generation if necessary. This switch could happen independently and externally of the speech planning process (lexical access). Thus, lexical access in noun and pronoun generation might not differ at all, leading to the same phonological inhibition effect. In highly standardized experimental sessions such response strategies might indeed play an important role.

However, the strategy of default internal noun planning should lead to a high proportion of noun description if pronoun description were required. This was not the case. Incorrect noun naming instead of pronoun generations occured in less than 1% of the critical trials. In addition, one (indirect) argument against the existence of the latent-speech-strategy is given by the results of the ellipsis experiments (see Chapter 2). In these experiments, a phonological effect was not observed for the ellipsis condition. Given the same amount of reduced utterance formats in both experimental designs (25% of all picture descriptions), it is not clear why participants would use the latent-naming-strategy in one experiment, but not in the other.

3.8 Third main experiment: Residual vs. re-activation during pronoun generation

The observed phonological inhibition during pronoun generation can be explained by assuming that the lemma and the phonological form of the noun are re-accessed. However, there might be an alternative explanation of the phonological inhibition. It involves potential residual activation of the form of the referent's noun during pronoun generation. In an utterance, such as *The flower is red. It turns blue*, the referent noun is overtly articulated in the description of the first picture. Because of this overt articulation, the phonological form of the noun must be available in the speech system. This phonological information might still be active at the moment the acoustic probe comes into play, which is shortly after the onset of the second picture. The observed phonological effect might therefore be due to interaction of the residual form activation of the first picture description with the activation of the phonologically related probe.

The assumption of such an activation trace in the speech system comes from word-word priming studies (O'Seaghdha, Dell, Peterson, & Juliano, 1992). In the experiment of O'Seaghdha et al., a word had to be named that was preceded by a prime word. The prime could either be phonologically related to the target or unrelated. One finding was that target response latencies were increased if the prime was phonologically related to the target word and the target was high-frequency; similar phonological inhibition is found in word-picture naming (see Eberhard, Bock, & Griffin, 1994). According to O'Seaghdha et al., this interference might result from the activity of a residual trace of the prime in a proposed word frame (following Dell, 1986). The residual activation of the phonological segments of the prime (for example, the 'g' in *pig*) can create competition with the target's segments (like the 'n' in *pin*), because both want to fill the same slot in the same word frame. According to O'Seaghdha et al., competition can only arise for fast (high-frequentcy) targets, because for slow targets the prime's residual trace would have already decayed.

In the first pronoun experiment, the previously articulated noun might have played the same role as the prime in the word-word priming experiments. It could have left a trace at the phonological level that led to inhibition of the encoding of the related lexical decision probe later on during pronoun generation. According to this view, the previously observed phonological inhibition during pronoun generation might not indicate a re-activation of the noun's phonemes during pronoun generation. Instead, the observed effect might indicate residual activation of the noun's phonemes that has nothing to do with pronoun generation at all.

To test the residual activation hypothesis, a second pronoun experiment was carried out. Basically, it had the same paradigm and design as the preceding experiment. It involved two utterance formats, the noun and the pronoun condition, and involved the same material and picture-probe timing as the preceding experiment. But in this experiment, the noun generation was the critical condition for determining whether residual activation of the first picture description influenced the lexical decision for the probe during the second picture presentation. In order to investigate residual activation, the noun condition was modified by reversing the presentation of the two different pictures (see Appendix F). The sequence *sun - flower*, now became *flower - sun*. This led to an utterance such as *The flower is red. The sun is blue*, instead of the previous *The sun is red. The flower is blue*. As before, the probe was presented 100 ms after the onset of the second picture. The probe could either be phonologically related or unrelated to the first picture noun. Because the onset of the second picture was time locked to the response onset of the first picture (as in the preceding study with a delay of 1500 ms) the probe was presented exactly 1600 ms after the response onset of the first picture description. It was the same probe as in the first pronoun experiment. However, by reversing the presentation of the pictures the probe was now phonologically related or unrelated to the first picture noun. It was always unrelated to the second picture noun in the noun condition.

3.8.1 Method

Participants: The participants were German speakers, which were recruited by newspaper and paid for participation. The age ranged from 18 to 36 years. The experiment was carried out with 35 participants. Five of them had to be excluded from further analysis because they made more than 20% naming and lexical decision errors in the main experiment and the baseline experiment.

Materials, apparatus, and procedure: The same materials were used as in the preceding experiment. The materials are listed in Appendix D. Apparatus and procedure were the same as in the first pronoun experiment.

Design: The same design was used as in the first pronoun study, except that the sequence of the two pictures were reversed in the noun condition (see description above for details).

3.8.2 Hypothesis

The assumption behind this manipulation of the noun condition was that if the form of the noun *flower* is still active during the presentation of the probe, phonological interference should be found. If there is no residual activation left at the form level, the latencies for lexical decisions on related and unrelated probes should be identical.

The pronoun condition was included in addition for three reasons: First, to keep the experiment as similar as possible to the first pronoun experiment. Second, the pronoun generation should serve as a replication of the pronoun results of the first study. Third, the pronoun condition served as a statistical control condition for the noun condition. If no effect were observed in the noun condition, this could be due to two things. The first possibility is that there is no residual activation left in the system. The second possibility is that a null effect could be due to methodological problems. The interpretation of a null-effect in the noun condition as resultung from there being no residual activation at hand is only valid, from my point of view, if the same participants and items showed phonological inhibition in the pronoun condition within the same experiment.

3.8.3 Results

The main results are shown in Figure 3.6 and Table 3.4. The table displays mean lexical decision latencies for the participants for the critical test probes (P, U) in the baseline session and main session for the two utterance conditions. The statistical analyses based on differential scores are shown in Figure 3.6. Differential scores and missing values were defined as in the first study, and are also reported in the table. Missing values were excluded from further analyses.

Table 3.4 Lexical decision latencies and differential scores (in ms), standard deviation (SD) of the differential scores, and percentage of missing values for the baseline and the main experiment for the two word probe conditions per utterance format (for subjects) Experiment Residual vs re-activation

	Pronoun generation		**Noun generation***	
	phon	unrel	phon	unrel
Baseline	910	894	921	905
Main session	1060	988	1003	991
diff (baseline - main)**	-148	-95	-82	-87
SD	95	102	90	90
% Missing	13	10	14	7

* Probes were related to the noun of the first picture

** Any apparent inaccuracies in means are the result of rounding

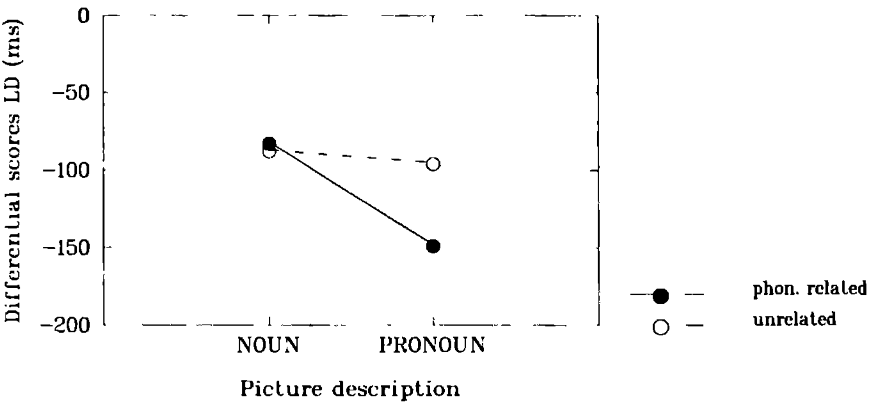


Figure 3.6 Mean differential scores of lexical decision latencies: Baseline (without picture description) minus main experiment (during picture description) for related and unrelated probes for each utterance condition. NOUN depicts the performance of probes that are related or unrelated to the first picture noun generation. It should indicate the amount of phonological residual activation of the overtly spoken first picture noun during the presentation of the probe after the onset of the second picture. Negative values indicate inhibition.

The mean reaction times for the baseline (see table) were statistically equal. This was shown by an ANOVA for repeated measurements on the factor 'condition' (four probe conditions per subject or item). $F_1(3, 87) = 1.1$, $Ms_e = 3188$; $F_2(3, 45) = 0.48$, $Ms_e = 4798$.

Errors: The error analysis was carried out on the arcsin transformed error proportions in a 2 (utterance condition) x 2 (probe condition) repeated measurements ANOVA. The factor 'probe condition' became significant: $F_1(1,29) = 7.7$, $Ms_e = 0.056$, $p = 0.01$; $F_2(1, 15) = 9.5$, $Ms_e = 0.037$, $p < 0.01$. More errors were produced in the phonological condition. This condition also showed the longest decision latencies in the main experiment. Therefore, a speed-accuracy trade off could be rejected. In order to exclude the possibility that the higher percentage of errors in the phonological conditions was due to item-specific effects, a separate analysis was carried out for the baseline lexical decision only. The error percentages across subjects in the baseline experiment were 2.5% (pronoun related), 2.1% (pronoun unrelated), 3.7% (noun related), 1% (noun unrelated). The arcsin transformed error proportions of the baseline experiment were tested in a repeated measures ANOVA with the four conditions as the within subject or item factor. The four conditions in the baseline did not differ with regard to error performance, $F_1(3, 87) = 1.7$, $Ms_e = 0.009$, $p = 0.17$; $F_2(3, 45) = 2.09$, $Ms_e = 0.01$, $p = 0.21$. As in the preceding experiment this result supports the assumption that the higher percentage of errors found in the main analysis is due to specific phonological effects in the dual task situation.

Differential scores: Repeated measurement ANOVAs were carried out on the differential scores with 'utterance format' (noun and pronoun generation) and 'probe condition' (related and unrelated) as repeated measurement factors. 6 outliers were excluded from the analysis, following Tukey (1977). The main effect 'utterance format' was significant, $F_1(1,29) = 5.2$, $Ms_e = 7572$, $p = 0.03$; $F_2(1, 15) = 6.03$, $Ms_e = 3845$, $p = 0.03$. This result showed that lexical decision became more inhibited in the pronoun condition than in the noun condition. The main effect 'probe condition' did not reach significance, $F_1(1,29) = 2.9$, $Ms_e = 5772$, $p = 0.10$; $F_2(1, 15) = 1.5$, $Ms_e = 5903$, $p = 0.23$. However, the interaction was significant, $F_1(1,29) = 5.2$, $Ms_e = 4405$, $p = 0.03$; $F_2(1, 15) = 9.74$, $Ms_e = 1554$, $p < 0.01$. This interaction indicated that phonological inhibition was obtained in the pronoun condition but not in the noun condition. In summary, phonological inhibition on lexical decision was found in the pronoun condition. No phonological effect was found for noun generation.

3.8.4 Discussion

One purpose of the present experiment was to replicate the phonological inhibition found in the first pronoun experiment during the generation of pronouns. The result of the pronoun condition shows that the replication was successful. In addition, the pronoun condition in the present experiment served as a control for the noun condition. A potential null effect during pronoun generation would make the interpretation of a null effect of residual activation of the first picture description impossible. Fortunately, this problem did not occur.

The second aim of the present experiment was to investigate a potential residual phonological activation of the overt noun generation of the first picture. As previously discussed, this potential residual activation might have caused the phonological inhibition effect during pronoun generation: The overt generation of the first picture noun might have left a phonological trace that could lead to competition with the phonological encoding of the probe. If the first picture description influences the lexical decision latencies of the probe later on, the observed phonological inhibition during pronoun generation could not be interpreted in terms of re-accessing the phonomes of the referent noun. The present result showed a null effect in the noun condition: The lexical decision latencies of probes that were presented during the second picture presentation were not systematically effected by the description of the first picture.

The present finding suggests that during the presentation of the probe no residual phonological activation of the first picture noun was present. The results showed that activated phonological information decayed over time and was not available anymore 1600 ms after the onset of overt articulation. The observed phonological null effect in the noun condition, therefore, supports the idea that the observed phonological inhibition during pronoun generation taps into the nature of lexical access during pronoun generation.

3.9 Conclusion

In the introduction of this chapter, a theoretical view about accessing pronouns was proposed. This view assumed conceptually driven access of the referent's noun lemma, the noun's gender access, and a discourse-driven access of pronominalization (see Figure 3.2). According to this view, the activation of the referent noun lemma is the initial step during the syntactic encoding of pronouns. Lemma access is also the first step in the syntactic encoding of overt noun naming (Levelt, 1989). The specific question addressed in the present experiments was whether lexical access of the referent noun differs between the two utterance formats. The two formats were compared by focussing on the phonological activation of the referent noun. A theory that favors a cascading spread of activation would predict phonological activation of the referent noun during pronoun generation - if the lemma has been activated before. In contrast, a strict two-stage model of language production in the sense of Levelt et al. (1991a) would predict no such phonological co-activation.

The two pronoun experiments revealed phonological inhibition during pronoun generation. The second pronoun experiment investigated the interpretation of the effect in terms of residual activation from the first picture noun generation. The results revealed no phonological effect in the noun condition. The observed null effect showed that no phonological activation from the noun generation of the first picture was available during the presentation of the probe. This interpretation of the null effect led to the conclusion that the observed phonological inhibition during pronoun generation must be due to re-accessing the phonemes of the referent noun.

This leaves us with two different accounts for the observed phonological inhibition. Both locate the effect during lexical access. However, the first one locates it during phonological encoding, the second one during lemma access.

Interference at the phonological level due to competition

The observed phonological inhibition can be explained in terms of *interference at the phonological level due to competition*: According to theories of language production that assume a cascading spread of activation, the activation of the picture's noun lemma automatically leads to the activation of the noun's phonological form. The phonological form of the picture's name becomes available even if the noun is not to be produced overtly. Form-overlap of the picture's name and the probe may lead to interference at the form level during pronoun generation, as proposed for single noun generation by O'Seaghdha, Dell, Peterson, and Juliano (1992, see also Peterson, Dell, & O'Seaghdha, 1989; O'Seaghdha & Marin, submitted). O'Seaghdha et al. (1992) developed an interactive activation model of form-related priming. This model showed that phonological interference may result from the activity of a residual trace of the prime. It successfully simulated empirical results of Colombo (1986, see below). The idea of a phonological trace was already introduced in the discussion of the second pronoun study. We considered the possibility that the *overt generation of the first picture* served as a prime for later probe processing. The observed null effect of residual phonological activation excluded this possibility. However, the idea of a trace could still be adapted to the dual task situation. The *generation of the second picture* might create a trace (not in terms of past activation but in terms of present co-activation). This trace influences the probe performance: The second picture's concept activates its lemma. The lemma sends activation towards its phonological form, regardless of the utterance format (noun or pronoun). The partially available form information leads to competition with the encoding of the probe if the picture name and the probe are phonologically related. Following O'Seaghdha et al., interference arises when two segments compete to fill the same slot in a word frame. For example, the phonemes /ɑ//ʊ//ə/ in the picture noun *flower* can create competition with the acoustic probe's phonemes /ɪ//n//t/ in *flint* because these segments want to fill the last three slots within the same word frame starting with /f//l/.

As discussed in section 3.3, the *location of the execution of the lexical decision* task is not clear if it involves phonologically related probes. In order to see the just-outlined phonological competition in the behavioral data of a lexical decision task, two different possible assumptions about the nature of the task should be considered. One possibility is that the lexical decision could immediately be carried out *after phonological encoding* of the probe. According to this view, delayed phonological encoding of the probe directly increases lexical decision latencies. Alternatively, the lexcial decision may be carried out after the semantic/syntactic (lemma) encoding of the probe (as proposed by De Groot, 1985; Levelt et al., 1991a; O'Seaghdha et al., 1992, model 2). If lexical decision takes place after lemma access of the probe (and picture words and probe words share the same lemma level),[6] a feedback flow of phonological information towards the lemma level has to be assumed. Only if the system has the possibility of feeding back the 'delayed' phonological information to the lemma level can lexical decisions can be effected by competition at the phonological level.

---

[6] A shared lemma representation for pictures and probes was not explicitly assumed by Levelt et al., but it was assumed, for instance, by Roelofs (1992a, b).

Competition of phonologically related words at the lemma level

According to a strict two-stage theory of lexical access, the picture's noun lemma is not selected for overt naming, because it is expressed as a pronoun. Therefore, the noun's lemma should not activate its phonological form (Levelt et al., 1991a, Roelofs, 1992a, b, for the selection criteria, and see Chapter 4 for further details with regard to mediated semantic priming). From this strict two-stage point of view, it follows that the explanation of Levelt et al. for the phonological inhibition during overt *noun generation* does not hold for *pronoun generation*. In 1991, the authors interpreted the phonological effect as being due to interference at overlapping phonological stages.[7] But, a two-stage model would predict that such a phonological competition should not arise at all during pronoun generation.

However, more recently an alternative account to an inhibitory process at the form level has been proposed. It can handle the observed phonological inhibition even in a strict stage model. This account comes from the domain of word recognition research. It assumes a *competition process of phonologically related words at the lemma level*. The phonological inhibition during pronoun generation may be explained in terms of cohort processes (Marslen-Wilson & Welsh, 1978; Colombo, 1986). In her word recognition study, Colombo presented prime-target word pairs. The prime preceded the target. The prime was either orthographically related or unrelated to the target word. Participants had to carry out a lexical decision on the target. Colombo found that lexical decision latencies on orthographically related targets were inhibited if the target was high frequency. Facilitation was observed if the target was low frequency. Colombo proposed a connectionist model to explain the inhibition of high-frequency targets. In this account, prime identification is assumed to initially involve the activation of the lexical units of orthographic neighbors (for example, the prime *flower* also activates the target *flint*). The model also assumes that successful identification of the prime also requires the inhibition of strong (and in particular, high-frequency) competitors. The result is that if a high-frequency neighbor is then presented as a target, its processing will be slowed. Low-frequency neighbors do not reach a level of activation that allows for inhibition.

---

[7] The explanation of Levelt et al (1991b) of the phonological inhibition effect was. "If (the picture of a) *goat* is phonologically active and *goal* is presented as lexical decision item, the activation of *goat* delays reduction of the cohort to the single element *goal* in comparison with the control condition " (1991b, p. 616) According to the authors' stochastic model of the strict two-stage theory of lexicial access, this delay is located at the form level. "If the target word in naming is phonological similar ... to the lexical decision item and the phonological stages overlap, the rate of the phonological stage in lexical decision is reduced for as long as the overlap of stages lasts " (1991a, p 137 The inverse rate, 1/r, is supposed to be the duration of the stage).

This line of reasoning for phonological inhibition can now be generalized and adapted to the dual task paradigm. The incoming acoustic probe activates a cohort of word candidates at the lemma level (e.g., the acoustic input /fl/ of the probe word *flint* activates the lemmas of *flower, flesh, flint...*). When the complete word has been heard, this cohort is normally reduced to one target element, and a lexical decision can be executed. The lemma of the phonologically related picture name *flower* that needs to be expressed as a pronoun later on is also initially a member of the probe cohort. In addition, it also becomes activated by its concept due to the initiation of the picture description. Because the picture's noun lemma is a member of the cohort and gets additional activation from the naming process, it is a strong competitor to the acoustic probe lemma. The picture lemma might delay the cohort reduction process towards one single element, the target probe. As a result the lexical decision on the acoustic probe is delayed.
To my knowledge, there is only one study (Roelofs, Meyer, & Levelt, 1996) that has recently adapted a phonological cohort process to language production. Roelofs et al. investigated a potential interaction of semantic inhibition and phonological facilitation in a so-called mediated priming condition (see also Chapter 4). For example, the description of a picture of a *cat* could be semantically inhibited by a distractor word *calf*, because both words represent animals. In addition, however, *calf* is also phonologically related to *cat*, which normally (without semantic relation) would lead to a decrease of reaction times. This interaction of effects in the mediated priming condition might lead to a null effect in the behavioral data, as observed by Starreveld and La Heij (1995, 1996). Roelofs et al. simulated this interaction by means of a computational model (Roelofs, 1992a, b). The simulation results showed that the bigger the phonological overlap, the smaller the observed semantic inhibition effect becomes. This decrease of inhibition was explained in part by a facilitatory cohort effect that Roelofs et al. located at the lemma level, and in part by facilitatory effects at the phonological level. However, because the authors proposed an explanation for phonological facilitation (not inhibition), it remains to be shown whether a phonological cohort effect might lead to inhibition as well, as observed in the present experiments by using the dual task paradigm.

The question about the location of the phonological inhibition will be addressed in more detail by means of a connectionist model. It is described in Chapter 4. At this moment, the simplest explanation for the observed phonological inhibition seems to be in terms of a model that assumes cascading spreading of activation. A preliminary comparison of the results obtained for the two different forms of reductions, ellipsis and pronouns, will be addressed in Chapter 5.

As an attempt to interpret the phonological interference effect observed in the pronoun experiments, a modular connectionist model was used to simulate activation processes and reaction times during a dual task situation. Especially the model should give some insights into the nature of the phonological inhibition effect observed by Levelt et al. (1991a) and during the first main experiment of Chapter 3. The present simulation compared two versions of the computational model: One version involved feedback from the phonological level to the lemma level (Simulation 1), whereas the second one has no such feedback (Simulation 2). The feedback model tested the assumption that the phonological inhibition is due to segmental mismatch effects, as postulated by O'Seaghdha et al. (1992). The second version investigated whether cohort-effects at the lemma level can explain the empirical data (as assumed by Colombo, 1986). The feedback version then was trained to generate pronouns in order to simulate the phonological effects found during the pronoun study (Simulation 3).

The present model is a parallel distributed processing (PDP) network (Rumelhart, Hinton, & Williams, 1986). PDP models have been applied in a wide range of language processing: In speech perception (for example, McClelland & Elman, 1986; Norris, 1992, 1994), speech production (for example, Seidenberg & McClelland, 1989; O'Seaghdha et al., Model 1, 1992), and in language development (for example, Rumelhart & McClelland, 1986; Seidenberg & McClelland, 1989; Indefrey & Goebel, 1993; Plunkett & Marchmann, 1993; Westermann & Goebel, 1995).

The model consists of modules. These modules reflect the different stages of speech processing that are supposed to be involved in the dual task of lexical decision and picture naming. The model's architecture follows basic assumptions of theories of language production (Garrett, 1975, 1988; Stemberger, 1985; Dell, 1986, 1988; Levelt, 1989; Roelofs, 1992a, b; O'Seaghdha et al., 1992). It is comparable to the architecture of Roelofs' (1992a, b) spreading-activation model.

However, the present PDP model differs from spreading-activation approaches with respect to how it acquires the performance of a task, such as picture naming: A spreading-activation model is completely built up by the researcher with regard to architecture and pattern of connections within and between speech levels. In contrast, the present model learns to map a specific input onto a desired output pattern by self-organized learning of the weights of connections between its modules. This self-organized learning mechanism has the advantage that a model's performance can be easily adapted to new tasks, in terms of developmental acquisition of more complex tasks. One example will be addressed in Simulation 3 with regard to the acquisition of pronoun generation in a model that was trained on noun naming first. In addition to the developemental aspect, the self-organized learning mechanism often produces surprisingly simple and, for the researcher post hoc, plausible solutions for a specific task performance. Models that use self-organized learning, therefore, serve as fruitful tool to create ideas and to improve theories.

The current model simulates reaction time latencies for correct lexical decisions. It, therefore, differs from models that also adress the issue of phonological competition during lexical decision, but use the probability of misselecting a critical phonological segment as a dependent variable (O'Seaghdha et al., 1992; Harley, 1993). A high probability of misselection of phonological segments (activated by a distractor word) was indirectly interpreted as inhibition of a lexical decision response to the target word. The current model chooses a direct approach for reaction time simulation by assuming a threshold of activation of lemma units. When a unit reaches the threshold, a lexical decision is carried out. Latencies are measured in terms of time steps (cycles) it takes to reach the threshold.

Before I come to the description of the model's architecture and the simulations, I will briefly summarize the two-stage view and the cascaded processing view with regard to feedback, mediated priming, and their assumptions concerning phonological inhibition. This theoretical introduction should make the rationale of the simulations more transparent.

4.1 Strict two-stage vs. cascaded processing view

The two-stage and the cascaded view differ with regard to assumptions about feedback and the nature of activation spreading (in serial stages vs. cascading), shown in the literature by means of so-called mediated priming experiments. These different theoretical assumptions are briefly outlined next, because they are implemented in the two computational model versions discussed later on.

4.1.1 Feedback assumptions

The strict two-stage view and the cascaded processing view of language production agree on the *separation of the two stages* of lexical access, lemma access and phonological access. Evidence for a separation of the semantic and phonological stages came from speech error data. Garrett (1975, 1976) distinguished between two classes of errors: Word and sound exchanges. Word exchanges were assumed to take place during lemma selection. An example may be 'I would like to have some *coffee* with my *sugar'*. In contrast, sound exchanges have to do with the retrieval of phonemes at the phonological level, for example saying *lemporal tobe*.

The two different theoretical approaches also agree on the *general time course* of the access of the two stages: Lemma access precedes phonological access.

However, the two theoretical approaches differs in some respects. For example, the existence of *feedback* connections is under debate. According to Garrett (1988) so-called mixed errors, which show both form and meaning relations between target and error (such as saying *rat* instead of *cat)* are rare in contrast to the two above mentioned error categories of word and sound exchanges. The rarity of mixed errors was seen as evidence that the two types of errors might evolve in relative independence (see also Martin, Weisberg, & Saffran, 1989). The independence was interpreted as absence of feedback from the phonological level to the lemma level: Feedback would lead to phonological influence on word selection. The error data, together with empirical evidence from reaction time data (Schriefers et al., 1990; Levelt et al., 1991a) and tip-of-the-tongue studies (Brown & McNeill, 1966; Brown, 1991, for a review) favor a two-stage view without feedback between form and lemma level.

Activation-spreading models of lexical access differ with regard to a feedback assumption. The cascade model, for example, of Humphreys, Riddock, and Quinlan's (1988) only assumes forward spreading of activation. However, some authors from the domain of cascading-activation models assume additional backward spreading of activation between form and lemma level (Dell, 1986, 1988; Dell & O'Seaghdha, 1991, 1992). Because the feedback assumption is important for an explanation of the phonological inhibition effect observed in Chapter 3 with regard to competition at the phonological level, I will focus only on the latter class of activation-spreading models in this chapter.

Dell and colleagues argued that the interactive nature of lexical access (i.e., feedback) in the cascaded processing view explains a variety of speech-error phenomena, such as the mixed errors. In contrast to Garrett, Dell argued that quantitative analyses of error collections showed a higher proportion of mixed errors than would be predicted from the independent contributions of phonological and semantic similarity (Dell & Reich, 1981; Martin et al., 1989; Dell & O'Seaghdha, 1991; Hinton & Shallice, 1991; Martin, Gagnon, Schwartz, Dell, & Saffran, 1996). Dell and O'Seaghdha explained mixed errors as products of activation spreading between lemmas and phonological units before the selection of a lemma. For example, if the semantic units for *cat* are active, the word (lemma) *cat* will become active, which, in turn, activates its corresponding phonemes. These active phonological units then send activation back to all words connected to them, such as *cap*, *cash*, *rap*, *rat*. Thus, phonologically related words acquire some activation and may be erroneously selected, creating a mixed error, such as saying *rat* instead of the intended *cat*.

An alternative account for mixed errors according to a strict two-stage model is that mixed errors could result from a mechanism of self-monitoring. This mechanism can be seen as a feedback spreading of information via the comprehension system (see figure 1.1). According to Levelt (1989), self-monitoring can begin as soon as there is a phonetic plan for the word, and so before articulation is initiated. If *rat* is internally but erroneously planned, sound-related words will be activated in the speaker's comprehension system, among them *cat*. Its meaning can be activated via this phonological route, and in addition via a semantic route, allowing the acivated word *rat* to prime the related word *cat*. Since that is the intended meaning, the monitor may not notice the error and pass the (erroneous) item. Levelt et al. (accepted) argue that the monitoring mechanism explains why it is more likely that a mixed error, such as *rat*, will occur in contrast to a semantic substitution error, such as saying *dog* instead of the intended word *cat* (as observed by Dell and Reich, 1981). Both errors are lemma selection errors because of the semantic relation to the target word. The monitor would equally likely detect this error. However, because *rat* is also phonologically related to *cat*, and *cat* is in the cohort of *rat*, it is less likely that the monitor notices the error *rat* in comparision to *dog*, which has no phonological relation to the intended word.

The discussion of feedback in a lexical decision during naming paradigm has been addressed in more detail by Levelt et al. (1991a, 1991b), Dell & O'Seaghdha (1991, 1992), and Harley (1993). For a discussion of this issue in a picture-word interference paradigm I refer to Roelofs et al. (1996), and Starreveld and La Heij (1995, 1996).

4.1.2 Mediated priming

Mediated priming has been addressed in the literature as a way to distinguish between a serial lexical access and a cascading access. According to the strict two-stage view, during picture naming the semantic information becomes available first. Only if the lemma access is completed does the form information become activated. Direct empirical support for a strict stage theory came from experiments 5 and 6 of the Levelt et al. study (1991a, see below). In these experiments, the question addressed was whether phonological encoding is restricted to selected items only, or whether any semantically activated item will, to some extent, become phonologically active. The authors argued that a strict

two-stage assumption predicts no phonological activation of non-selected items, because only selected lemmas are supposed to spread activation to their form.
In contrast, a cascaded processing view predicts an activation of a non-selected element: During picture naming the concept activates the lemma. Partly activated lemmas can already spread activation through the system towards the phonological level. Furthermore, in some cascading models via the backward spreading of activation, the phonological information can influence the ongoing lemma processing. This feedback has consequences for a lexical decision if a lexical decision task involves lemma access (which is assumed by Levelt et al. and in the present simulations, see also the general conclusions of Chapter 3).

Levelt et al. (1991a) tested the strict-stage hypothesis by means of *mediated priming*. They investigated whether during the preparation of a naming response, not only the target but also close semantic associates or same-category members are phonologically activated. For example, they assumed that during the naming of a picture of a *sheep* the semantically related concept of a *goat* becomes co-activated, but not the phonological form of *goat*. This was tested by presenting lexical decision probes that were phonologically related to the semantic associate, for example *goal* for *goat*, given a picture of a *sheep*. The results showed no phonological effects on *goal*, indicating that *goat* does not become phonologically active. This result was interpreted as support for the strict stage theory.
In their reply to Levelt et al., Dell and O'Seaghdha (1991) argued that the amount of semantic activation of the semantically related alternative is only a fraction of that of the target item. The amount of phonological activation of the alternative is then in turn only a proportion of this fraction, which leads to 'multiplicative diminution' (p. 607). This small amount of phonological activation has to be sent back to the lemma level, in order to influence a lexical decision on *goal*. Dell and O'Seaghdha posit that the amount of mediated priming should be very small and should not be discovered empirically, as shown in the Levelt et al. data (but see also Levelt et al., 1991b, for a reply).
However, direct empirical support of the cascaded processing view came from a picture-naming study of Peterson, Shim and Savoy (1993; see also Peterson & Savoy, in press; O'Seaghdha & Marin, 1997). In the experiment of Peterson et al. (1993), the participants were presented with a picture. Shortly following the onset of the picture (at different SOAs), one of two types of stimuli was displayed in a small frame. On half the trials, the stimulus was a question mark. This question mark served as the signal for the participant to name the picture. On the other half of the trials, a target word was presented. In this case, participants had to read the word instead of naming the picture. The pictures consisted of objects with two nearly synonymous names (for example a

couch/sofa) where *couch* was the dominant name, and *sofa* the subordinate name (that in a prestudy was only used in 15% of the responses). The target word could be either phonologically related to the dominant name, such as *count* to *couch*, or to the subordinate name, such as *soda* to *sofa*. The target word could also be unrelated to the picture name. The authors assumed that the two synonyms should become semantically co-activated. According to a cascaded processing view, both dominant and subordinate names should receive early phonological activation. According to the two-stage view, only a single name for the picture should become phonologically active. The authors found evidence for early phonological activation of both *count* and *soda*, after presenting the picture of a couch. The phonological activation was observed in terms of faster reading times of phonologically related items in comparison to unrelated items. This result was interpreted in such a way that the two semantically activated items (*couch* and *sofa*) already activate their phonological segments, even if only one of the items become selected for naming, supporting the cascaded processing view.

In a second experiment, Peterson et al. tested the phonological activation of non-synonymous semantic alternatives (i.e. a same category member, such as *bed* for a picture of a couch). Here the phonological probe would be *bet*. With regard to the used picture-word relations this experiment was a replication of the Levelt et al. (1991a) experiment of mediated priming. As in the Levelt et al. study, Peterson et al. observed no phonological activation of same category semantic alternatives. Peterson et al. explained the difference of the synonymous experiment and the same category experiment in terms of different amounts of co-activation. They assumed that a synonym will receive much more activation than will a category associate that is not a possible name of the picture. The synonym name will be able to activate its phonological form to a much greater degree than will the category member, and thus phonological priming might more easily be obtained for these items. Phonological priming might be too small to be reliably observed in a priming experiment.

Levelt et al. (accepted) proposed an alternative explanation for the observed difference of using synonymous items or same category members. This explanation is based on Roelofs' (1992a) account for word blends (errors, such as saying *my* when *man* or *guy* was intended). He assumed that word blends might occur when two lemma nodes are activated to an equal level, and both get selected, because the selection criterion in spontaneous speech is satisfied simultaneously by both nodes. The parallel selection of both lemmas leads to a co-activation of their phonological form. Roelofs argue that this account would explain why these blends mostly involve near synonyms.

Thus, with regard to the described mediated priming results, the question of the nature of information flow is still open.

An alternative method to distinguish between the two theories is to look at *pronoun activation*. As for mediated priming, according to a strict two-stage theory no phonological activation of the non-selected item was expected, in this case the not overly spoken noun referent. In contrast, a cascaded processing view would predict such activation.

As discussed in the general conclusions of Chapter 3, the observed phonological inhibition during pronoun generation in my own experiments would favor the cascading view, if we still lived in 1991. This view locates the phonological effect at the phonological level due to competition of mismatching segments (O'Seaghdha et al., 1992). Via backward spreading of activation to the lemma level the lexical decision latencies are delayed, as observed in the experiments in Chapter 3.

However, recently, theorists who favor the two-stage model introduced a phonological cohort mechanism at the lemma level which can explain a phonological effect (Roelofs et al., 1996). According to this view, a phonological effect does not violate the two-stage lexical access assumption, because it does not reflect mere phoneme activation but also phonologically driven lemma activation. According to a phonological cohort account that locates the phonological effect - at least in parts - at the lemma level, the phonological effect does not necessarily show that the non-selected noun lemma spreads activation to its form during pronoun generation.

The presence of the alternative account means, in turn, that the observed empirical effect for pronoun naming does not clearly distinguish between the two theories anymore. However, because the proposed phonological cohort effect at the lemma level explains facilitation (in picture-word interference tasks), and not inhibition, it still has to be shown how the inhibition observed in the pronoun experiments should be explained. This issue will now be addressed by means of a connectionist model that contrasts the cohort processing and the phonological mismatch assumption by means of two different model versions: The first simulation tests the phonological mismatch assumption of the cascaded processing view by involving backward spreading of information from phonemes to lemmas. The second simulation tests the cohort view of the two-stage theory.

4.2 The empirical data to be simulated

In order to simulate lexical decision during pronoun naming a model should be capable of simulating single noun naming first. For the simulation of lexical decision during noun naming the empirical data of the Levelt et al. (1991a) study and of the replication of the Levelt study (first main experiment in Chapter 3) were considered. Both studies measured lexical decision latencies on phonologically related, semantically related, and identical acoustic word probes, and compared them to an unrelated probe condition. Although the phonological effect is of major interest here, the model should - for the sake of completeness - capture the empirical semantic and identical effects, as well as the previously reported results of mediated priming. The following results were obtained in the experiments (see also Figure 4.4, top panel, p. 150):

- In the phonological condition, the behavioral data of Levelt et al. showed phonological interference across a broad range of SOAs (73 ms - 673 ms, presenting the picture first, the acoustic probe second) during single noun naming. The results are in line with my own findings (phonological inhibition at SOA = 100 ms) during noun and pronoun generation.
- In the semantic condition Levelt et al. found inhibition at short SOA but not at long SOA. In my own single noun naming experiment, I could not replicate this early semantic inhibition, but this might be due to material problems (as discussed in Chapter 3). Therefore, for the simulation, the results of the Levelt study were the relevant ones.
- In the identical condition Levelt et al. found interference at a short SOA and facilitation at a long SOA. In contrast, the replication study showed significant facilitation at a short SOA. Because of these contradicting findings the outcome of the simulations with regard to the identical condition was kept open.
- In the mediated priming condition no phonological effect should be observed, according to the findings of Levelt et al. (1991a, experiment 5 and 6).

4.3 The general architecture of the lexical decision model

In spite of the fact that the strict two-stage model and the cascading model disagree on the way the activation spreads from one level to the next, the models agree with respect to the levels that are involved: Semantic, syntactic, and phonological representation. These speech levels are implemented in the model in terms of modules. The two theoretical approaches also agree on the general time course: During picture naming the meaning becomes activated first, and later on the phonology. This time course is implemented in the current model by means of temporal dynamic spreading of activation from one level to the next.

4.3.1 Shared representations

The computational model should provide an answer to the question why inhibition occurred at all in the lexical decision task during picture naming.

During the task, the participant had to name a picture. In addition, he or she had to pay attention to an acoustically presented probe word. So, two different input modalities were involved. The basic assumption for using the dual-task paradigm in language studies was that the lexical decision is systematically influenced by the picture naming process. According to the mathematical model of Levelt et al. (1991a) semantic activation of the picture slows down the semantic activation of the acoustic word if picture and probe are semantically active at the same time. Partial phonological activation of the name of the picture delays the phonological process of the acoustic probe due to competition at overlapping phonological stages. The slowing-down of a lexical decision means that the two input modalities seem to share - at least in part - internal representations during the speech process. The slowing-down of lexical decision could also mean that two corresponding representations are closely connected. One possible way of overlap and interaction of speech processing stages was implemented in the present architecture of the computational model, which is described next.

4.3.2 The stages of speech processing in the model

Figure 4.1 shows the stages that I assume to be engaged in a dual-task experiment of lexical decision during picture naming (roughly following Roelofs, 1992a, b).

In the *naming of a picture* at least four processing stages are involved. First, based on visual input, the object has to be identified. This identification takes place at the concept level by activating the object's concept. At the concept level the meaning of the word is stored. Second, the concept activates its lemma at the lemma level. Here, syntactic information is linked to the target lemma, for example, the lemma has a connection to its pronoun. Third, the lemma activates the object's phonological form at the phonological level. This level consists of phonemes. The fourth stage involves the articulation of the name of the object.

The *recognition of the acoustic probe word* involves at least three stages. First, the acoustic signal activates the phonological form. Second, the phonological information activates the lemma. Third, at the concept level the meaning of the acoustic probe gets identified.

The phonological stage of the recognition process is assumed to be separated from the phonological stage of the naming process because of empirical evidence reported by Shallice, McLeod and Lewis (1985). The authors found that a dual task of detecting a name in an auditory input stream, while reading aloud visually presented words showed only little single- to dual task performance decrement. Shallice et al. interpreted the relative ease of the dual task by assuming two different phonological representations.

In addition, if the lemma activation of the acoustic word reaches a specific threshold, a positive *lexical decision* can be carried out. The lexical decision is assumed to be directly linked to the activation of the lemma because of empirical findings of Levelt et al. (1991a, experiments 3 and 4). Experiment 3 was the dual task paradigm, where participants had to carry out a lexical decision during picture naming. The results of this experiment showed a clear semantic interference effect at a short SOA. In experiment 4, all else being equal, participants had to carry out *a lexical decision task during picture identification* - instead of picture naming. They had to decide whether or not a presented picture was known from a pre-session. No naming response was required. In this dual task situation the lexical decision latencies on semantically (to the picture) related probes were not delayed. This finding is in contrast with the outcome of experiment 3. The lack of semantic interference in the picture identification task was seen as evidence for ruling out recognition processes as explanation for the semantic inhibition. The semantic interference in experiment 3 was assumed to be purely an effect of lexical access (lemma selection), because it only occurred if naming was involved.

As has been discussed in Chapter 3, the *location of the lexical decision execution* is not that obvious for phonologically related probes. It could for example take place immediately after phonological encoding. However, by stipulation, it was assumed that no two different execution mechanisms for the same task were at hand, but only the lexical decision after lemma selection.

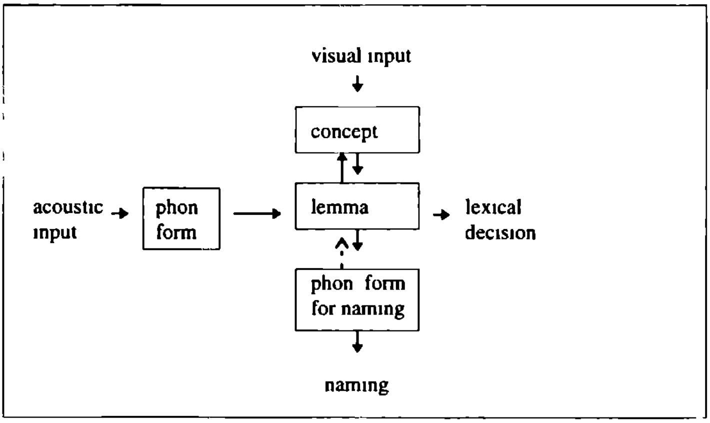


Figure 4.1 The basic architecture of the dual task computational model Depicted are the speech levels assumed to be involved in the dual task. The arrows indicate directions of activation flow. Two versions of the model differ with respect to whether or not there is phonological feedback from the phonological form to the lemma level (dashed line).

4.3.3 Two different model versions

The two different accounts for phonological inhibition (segmental mismatch or cohort effects at the lemma level) will be modeled by means of two different computational network architectures (see Figure 4.1). The first architecture includes phonological feedback flow of activation from the phonological form to the lemma level (dashed arrow). It therefore resembles the assumption made by some cascaded processing theories, and should test the phonological mismatch account. The second architecture has no feedback connections from the phonological form level to the lemma level. It tests the cohort assumption of inhibition during lemma activation, made by a two stage-model.

4.4 Details of the connectionist model

The above mentioned assumptions about the dual task paradigm were implemented in a modular computational network model. The model was built using the software package 'Neurolator', developed by Goebel (1995). The type of model is a recurrent one. Recurrence is needed in order to obtain temporal dynamics of activation spreading between levels (Pineda, 1987).

The backpropagation model contained no pre-wired inhibitory or exitatory connections between speech levels. Rather, it learned to map between input patterns of stimuli (word and pictures), their internal representations at each speech level, and their output patterns on its own. The model learned this mapping by using a learning algorithm called the delta rule (Rumelhart, Hinton, & Williams, 1986). The goal of learning was to find an optimal pattern of connection weights between the levels of processing. An optimal pattern of connection weights was given if every input led to its desired output. The obtained self-organized connection weights could either be of exitatory or inhibitory nature for the spread of activation through the system.

When the model had learned successfully to satisfy the representation of the acoustic and visual pattern, it had to carry out a dual task: Lexical decision during picture naming. As in the real experiments, picture-probe presentations in different SOA conditions were simulated during testing. In order to achieve a gradual spread of activation from level to level over time in the testing mode, the Pineda mode was implemented (Pineda, 1987, see testing section below). This leads to a cascading-like activation, which means that a slightly activated unit already spreads activation towards the next level of processing.

Whereas the cascaded processing is assumed and desired in the feedback version of the model (Simulation 1), it does not play an important rule in the cohort version (Simulation 2) in terms of contradicting the two-stage theory. The crucial difference between two-stage and cascaded processing theories is that in a two-stage theory a specific selection criteria for a target lemma is assumed. Only if a lemma reaches a certain activation threshold does it become selected and spreads activation to its phonemes (see Roelofs, 1992a, b). The nature of selection criteria is relevant for simulating naming responses, because depending on the criteria the phonological form for naming becomes activated in different ways. However, in a model without phonological feedback (as in Simulation 2), the selection criterion does not play a role for lexical decision. This conclusion can be drawn because the lexical decision effect is supposed to be located at the lemma level. In a two stage-model it could only be influenced by an interaction between conceptual and cohort like-activation at the lemma level. This interaction of conceptual and lemma activation, however, is continuous according to the two-stage model (Roelofs, 1992b, p. 43), as it is in the present model (Simulation 2).

4.4.1 Description of the network levels

Design of the *input levels:* As depicted in Figure 4.2, left panel, the acoustic input is realized as mock speech, allowing for an input of mono- or disyllabic words. Each syllable consists of onset (first consonant), nucleus (vowel), and coda units (final consonant), following current theories on phonology (for example, Selkirk, 1984; Booij, 1995). Onset and coda units represent the 20 phonemic consonants of Dutch. The nucleus units represent the 16 phonemic vowels of Dutch. For example, an acoustic input of the word /bureau/ (desk), the phonemes /b/ (onset 1st syllable), /y/ (nucleus 1st syllable), /r/ and /o/ (onset and nucleus 2nd syllable) become active. The visual input consists of 16 units. Each unit represents one target picture.

The design of the *phonological form level* for naming is identical to that of the phonological form input level.

Design of the *lemma level:* The lemma level includes 16 units (i.e., there are 16 lemmas in the simulation). Each unit represents one lemma of an acoustic probe and its picture. Given the acoustic input of /byro/ or the visual input of the picture "bureau", the "bureau-unit" at the lemma level should become active. The lemma units represent items that are needed to simulate a dual task experiment. Therefore, the lemma level of the model contains units of pictures and their semantically related, phonologically related, and unrelated acoustic probes. In addition, some units represent items that are phonologically related to semantic alternatives.

Design of the *concept level:* Each visual input (or each activated lemma) leads to an activation of two units at the concept level (see Figure 4.2, right panel). One of these units could also be activated by semantically related concepts, leading to partial overlap of semantic information.

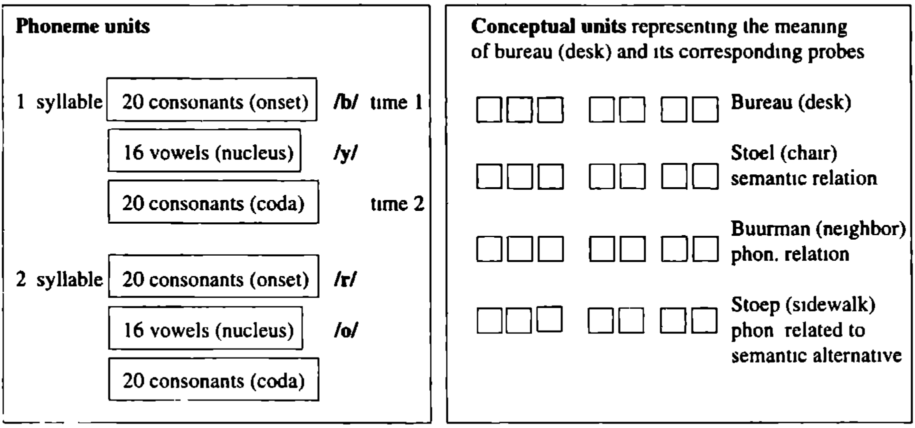


Figure 4.2 *Left panel* Phoneme units at the acoustic input level (an identical representation is implemented for the form level of naming) This input level allows for a one- to two-syllabic input. As an example, the input of the acoustic probe *bureau* (desk) is shown. The dashed line indicates the way in which the sequential input was realized At time step 1, the word initial phonemes were presented At time step 2 the remaining phonemes of a word were presented *Right panel.* Depicted are four times the same 7 conceptual units in order to illustrate the meaning representation of one item and its related probes. A grey unit indicates that this unit should be active during meaning encoding (see text for further explanation).

This form of coding is a very simplified version of distributed representation (Hinton, McClelland, & Rumelhart, 1986). Each unit can be regarded as a carrier of micro-feature-information of the target. This representation is supposed to represent the *core meaning* of an item (Miller, 1969; for a review see Levelt, 1989, pp. 212-214). A core meaning involves the idea that each meaning of an item is unique. Therefore, no second item exists that has the same meaning. The model is trained in such a way that a lemma only gets selected if its core meaning is satisfied by its specific conceptual activation. Because of unique representation no hyperonym problem arises.[1]

This form of representation differs from so-called decomposed representations where words can be retrieved on the basis of a combination of single primitive concepts (e.g., Bock, 1982; Stemberger, 1985). In that view, for example, the lemma *dog* is retrieved on the basis of conceptual representations like ANIMAL and BARK. But, as discussed in Levelt (1989), here the hyperonym-problem arises: If the conceptual primitives are active in order to select the lemma *dog*,

---

[1] This hyperonym problem (Levelt, 1989) addresses the issue that if the conceptual conditions of a hyponym (for example, dog) are met, then those of its hyperonym (for example, animal) are also satisfied The sharing representations might lead to activation and naming of the word "animal" instead of the intended word "dog" This problem does not hold for core representations. They distinguish between hyponyms and hyperonyms following the Gestalt-psychological idea that wholes are not simply the sums of their parts (for a review see Hinton, McClelland, & Rumelhart, 1986; Hinton & Shallice, 1991)

they would also activate their lemmas *animal* and *bark*, because their conceptual-syntactic mapping is not unique.

The present form of meaning representation also differs from so-called local representation, where a single concept node stands for a particular meaning (Collins & Loftus, 1975; see also Roelofs, 1992a, b, 1997, for reviews). According to this view, an abstract representation DOG is used to retrieve the lemma *dog*. Properties such as ANIMAL remain background information, which is linked to the concept DOG in terms of labeled connections, such as DOG 'is a' ANIMAL (where 'is a' indicates such a labeled link).

In the present model, the semantic relation is coded in a combination of three units for every target-probe pair. For example, the concept of BUREAU (desk) is coded as 110, which means that for these three units the first unit and the second unit has to be active, the third unit has to be inactive. In Figure 4.2, right panel, activated units are depicted in grey. In contrast, the concept of the semantically related probe word STOEL (chair) is coded within the same units as 101. So, the first unit is active for both, desk and chair, and can be interpreted as common information of "chair" and "desk". This unit might also be part of the concept of FURNITURE, but it is not the entire concept. The second unit might represent features related to "desk", whereas the third unit represents specific features for "chair". The semantic overlap, therefore, is given by the overlap of activation in the first unit. The semantic mismatch is given by the second and third unit, because it distinguishes between the two items.

As depicted in Figure 4.2, the phonological probes are coded in two separate units. They should be active when the phonological probe is presented. The phonologically related probes to the semantic alternative are represented in another two separate units. Not depicted in Figure 4.2 are unrelated probes. This is because an unrelated probe was randomly selected from the remaining items. It is coded as two activated concept units in a different group of 7 units.

The *output* consists of two domains: 16 units for the naming output for the pictures and two units for a lexical decision output. One of the lexical decision units is a Yes-response, the other one is a No-response (the latter will not be addressed any further).

4.4.2 The connections between speech levels

No connections within speech levels exist. The between-level connections are as follows: The *acoustic input* level is fully connected to the lemma level, which means that every phonological input unit may send activation to all lemma units. The acoustic input level is also fully connected to the form level of naming:

Every phonological input unit may send activation to all units of the form level of naming. No feedback connections exist towards the phonological input.
The assumption that a perceived word can directly activate its articulatory program is intuitively supported by a speaker's ability to repeat acoustically presented pseudowords that have no meaning representation. Empirically, the direct route was supported by Glaser and Glaser (1989) and La Heij et al. (1990). They observed an inhibitory effect of word distractors on reading a word aloud. However, there was no additional effect of semantic relatedness between distractor and target. This latter result indicates that a word can be - but does not have to be - read aloud without explicitly selecting the word's lemma. In addition, if instead of a distractor word, a distractor picture was given, almost no effect on reading was observed (Glaser & Düngelhoff, 1984). This finding leads to the interpretation that the observed inhibitory effect is due to direct connections of phonological form input and phonological form output.

The *form level of naming* is fully connected to the picture naming output units and with all lemma units. All connections from units of the form level to phonologically unrelated lemma units were set to zero in order to realize that there is no relation, as assumed in experiments using unrelated control items. The *lemma level* is fully connected to the phonological level, to the lexical decision output units, and to the conceptual level. The *conceptual level* is fully connected to the lemma level. The *visual input* is connected to the conceptual level in a 1:2 way, meaning that every target picture is only linked to its two conceptual units. The connection weights here were fixed to a value of +10 (w = +10) in order to prevent variation of conceptual activation due to picture input. All other connection weights were set to an initial value of w = 0.5 and were free to be changed during learning.

Table 4.1 Material used to train and test the computational model

	semantically related to first column	**phonologically related to first column**	**phonologically related to second column**
bureau (desk)	stoel (chair)	buurman (neighbor)	stoep (sidewalk)
cactus (cactus)	stekel (sting)	kakker(lak) (cockroach)	steen (stone)
fietspomp (inflator)	band (tire)	file (queue)	bank (bank)
geweer (rifle)	oorlog (war)	gewei (antlers)	oorbel (ear-ring)

4.5 The training phase

During the training phase, the modular network had to learn input-output mappings between adjacent speech levels for each item (see Table 4.1). For example, the phonological input of the word *desk* had to be mapped onto the 'internal target' output *desk* at the lemma level. In addition, the representation of the lemma *desk* was an input that had to be mapped onto the conceptual output, again given as internal target units. The network learned this modular input-output mapping by specifying weight-patterns of the between-level connections in such a way that a given input would lead to the desired internal output. This modular input-output mapping was trained in terms of sequences of input-output patterns for each task.

The model was trained first to acquire a lexical decision task on 16 acoustically presented words. After this training step, the obtained weights of the connections from form input to lemma level and to the form level for naming were frozen, as well as the connections from the form level for naming to the lemma level. In the next training step, the model learned to perform the picture naming task.

4.5.1 The internal representations in terms of internal target units

The lexical decision task on an acoustic word probe, such as "bureau" (desk), was initiated in the form input by the activation of the phonemes /b/ /y/ /r/ /o/. The model then had to build up the connection weights from the form input to the form level of naming in a way that exactly the phonemes /b/ /y/ /r/ /o/ at the form level of naming became activated. These units were specified as so-called internal target units that should become active. Furthermore, the model had to activate the internal target node *bureau* at the lemma level. It also had to activate the output target unit for a positive lexical decision response, because "bureau" is a word. During the picture naming task, the activated input node for the picture "bureau" should lead to the activation of the concept units for BUREAU, the lemma units *bureau*, the phonological nodes /b/ /y/ /r/ /o/, and the naming output node for "bureau".

4.5.2 Sequential acoustic input

For each acoustic word, for example "bureau", the model got a sequence of two phonological inputs (see also Figure 4.2, left panel). First, it got the word-initial two phonemes as input pattern, /b/ /y/, and in a second time step it got the remaining phonemes, /r/ /o/. This sequential input is a simplified version of the more natural phoneme-wise input. But it is sufficient to create the desired cohort representation at the lemma level. This is the case because the list of training patterns includes words that share the same phonological onset. For example, if the network had to train the internal representation of "bureau", it also had to train that of "buurman" (neighbor), which also begins with the phonemes /b/ /y/. The mere presentation of /b/ /y/, therefore, is related to two potential candidates at the lemma level. For "bureau", this input should satisfy the target representation of the lemma *bureau*, for example as 10 (1 represents the activated target lemma *bureau*, 0 represents the phonological related lemma *buurman;*, ignoring other lemma units for the moment which should be deactivated). For "buurman", this input should satisfy the target representation of the lemma *buurman*, for example as 01 (within the same two units as mentioned above). How does the model solve the ambiguous input problem? In order to minimize the error during learning, a network trained with the delta rule chooses the solution to activate both lemma candidates to an amount of 0.5 instead of 0 or 1[2]. This learning behavior creates a cohort-like activation pattern at the lemma level, given an initially ambiguous input. This cohort is reduced towards a single element, when at a next time step the remaining acoustic input phonemes are trained that favor one specific target lemma. In order to have every item as a potential member of a cohort, the material list was created in such a way that for each item a phonologically related item exists (see Table 4.1).

4.5.3 The structure of the training patterns

The internal target units of every speech level were specified in the training pattern, separately for each task. The training was carried out sequentially for individual adjacent submodules of the network. For example, in order to train a lexical decision task, a word pattern consisting of three parts was needed. The first part of an acoustic learning-pattern trained the model to map the word-initial phonemes to adjacent speech levels. The model trained the connection weights of the form input to the form level of naming, and from the form input to the lemma level. During this step, the units of the remaining speech levels were specified as "don't care values" (Jordan, 1986), meaning that they were allowed to produce an arbitrary value during training at this time. A second

---

[2] According to the learning procedure of the delta rule, the measure of error for each pattern is $E_p = 1/2 \, \Sigma \, (t_{pj} - o_{pj})^2$, where Σ is the sum across all j units of the pattern, t_{pj} is the desired target output, and o_{pj} is the actual output of unit j. To give an example, an actual 00 solution creates the difference between 00 and the desired 10 (bureau), which leads to $E_p = 1/2 \, [(1-0)^2 + (1-1)^2] = 0.5$ In addition, a 00 solution creates a difference between 00 and the desired 01 (buurman), again leading to an error $E_p = 0.5$. The overall measure of the error $E = \Sigma E_p$ is than $E = 0.5 + 0.5 = 1$ The same holds for an actual 11 solution Alternatively, an actual activation of 0.5 and 0.5 leads to a sum of overall errors $E = 0.5$ for the two alternative inputs, which is prefered by the network in order to minimize the overall error.

pattern part let the model train the connection weights for the word-final acoustic input in the same way. In addition, it trained the connections from the lemma level towards the lexical decision units. A third part of the training pattern let the model train the connection weights from the lemma level to the conceptual level. The same principle holds for the picture naming task: The first pattern part trained concept-to-lemma connection weights. A second part trained lemma-to-form level-to-naming output.

4.5.4 Task units

Acoustic and visual stimuli had exactly the same internal representation when they represented the same object (for example, the picture of a desk, and the acoustic probe 'desk'). The same internal representation usually resulted in the same output. So, normally the model would give both a naming and a lexical decision response regardless of the modality of input. To circumvent this problem, the experimental instruction was simulated by two so-called task units. They were fully connected with the input layer and the output units. Connections from the input layers towards the task units were learned by the network so that any visual stimulus activated task unit 1 and any auditory input activated task unit 2. The connections from the task units to the output units were pre-wired. The instruction of the experiment "Name the picture if no acoustic stimuli is given" was realized by task unit 1. Connections from that unit to the naming outputs were excitatory ($w = +5$), while its connections towards the lexical decision units were inhibitory ($w = -5$). The instruction "Don't name the picture when you hear a word, but do a lexical decision instead" was given by a task unit 2. Its connections towards the naming units were inhibitory ($w = -5$), while those towards the lexical decision unit were excitatory ($w = +5$). In general, this construction gave the desired outcome for single task performance, for example, a lexical decision whenever an auditory word came into the system.

4.5.5 The activation function

The activation of a unit served as the output value sent to other units. The unit itself gets activation from other units that are active and connected to the target unit. The amount of activation a target unit gets is called the net input of this unit. This input depends on two factors: First, the activation of the units that are connected to the target unit, and second, the weights (the strength) of connections between these units and the target unit. The net input can be defined as the sum of the product of these factors, as follows:

$$net_i = \sum_j w_{ij} a_j \quad , \qquad (1)$$

where w_{ij} refers to the weights on the connections from nodes j to node i, and a_j indexes the activation of nodes j which send to node i.

Given this net input, the activation of a unit can follow specific activation functions. The activation of the units of the model follows the logistic activation function:

$$a_i = \frac{1}{1 + \exp^{-(net_i)}} \quad , \qquad (2)$$

where a_i refers to the activation (output) of node i, exp is the exponential, and net_i is the net-activation flowing into node i.

4.5.6 The learning rule

As a learning rule the delta rule was used (see Rumelhart, Hinton, & Williams, 1986): The model started with a randomly selected weight configuration (in all simulation initial $w = 0.5$). After one episode (all pattern-sequences were activated once) the model calculated the difference between the desired representation and the obtained representation. The difference was interpreted as learning error (see note 3). This error was then propagated back through the levels and was decreased by changing the weights of the connections. The model did this according to the delta rule for recurrent networks, which implements a gradient decent principle in order to minimize the error. In this simulation an improved back-propagation algorithm was used, called the quick-prop-algorithm according to Fahlmann (1990). This algorithm chooses the best fitting learning parameters automatically, such as the learning rate. Training was finished when an error minimum had been reached.

4.6 The testing phase

After finishing the training, the testing phase began. This testing phase involved the simulation of the dual task experiment. A trained model can be seen as one pseudo-participant. The model was confronted with virtual acoustic and visual input patterns and had to carry out a lexical decision task at various SOA conditions. The dependent variable was the number of cycles (time steps) it took the model to make the lexical decision. A positive lexical decision was defined to be executed if one lemma reached an activation threshold. This threshold was set to 0.95 during all simulations. It indicated that a lemma was nearly fully activated - and therefore selected as a lexical decision target.

4.6.1 The testing algorithm

One cycle of testing means sending activation through the model for one step in time. A second cycle adds activation to the already existing activation in the system, and so forth. The activation of each speech level unit follows the logistic activation function, as described above. In addition a so-called Pineda mode (Pineda, 1987) was used in order to implement gradual temporal dynamics. The Pineda mode leads to summation of activation over time according to the following formula:

$$a(\mathrm{t}) = (1 - \mathrm{tau})\, a(\mathrm{t}-1) + \mathrm{tau} * \mathrm{sigmoid}\,(\mathrm{net} + \mathrm{bias}). \qquad (3)$$

The activation (a) of a unit at one moment in time (t) can be seen as a sum of two kinds of activation. The first term of this sum includes the activation of the same unit at one time step before t, at t-1, multiplied with a constant (1-tau). The second term of the sum is the current activation that the unit gets from other units multiplied by tau. Net is the net input. Bias means possible fixed activation values of the unit.

If the constant tau equals 1, the first part of the sum becomes zero and the unit's activation function is the usual sigmoid one. In this case the unit only has to deal with the net input it gets from other units at time (t). If the constant tau equals zero, the second part of the sum becomes zero. The means, the unit would not change its activation because $a(\mathrm{t}) = a(\mathrm{t}-1)$. If tau is $0 < \mathrm{tau} < 1$, then the activation of the preceding time step and the incoming new activation sum up. This leads to an activation function that implements temporal dynamics. In all current simulations tau was set to 0.1.

If the net input to a unit is zero, the unit's activation decays over time, following the Pineda-rule

$$a(\mathrm{t}+1) = (1-\mathrm{tau}) * a(\mathrm{t}), \text{ if } \mathrm{net}_{\mathrm{t}} = 0\ , \qquad (4)$$

4.6.2 Time delayed activation spreading

The standard Pineda mode leads to a spread of activation from one level to the next within one single time step, or one testing cycle. For example, if a unit becomes activated at the concept level it starts to activate its corresponding unit at the lemma level already during the next time step. At time step 3, the phonological units start to increase their activation. This leads to nearly identical activation of all corresponding units at all speech levels shortly after the presentation of the picture. The Pineda mode, therefore, reveals only a very limited time course of activation spreading between the speech levels.

With such a limited time course the aim of the dual task experiment could not be realized properly in the model. The main purpose of the dual task study of Levelt et al. (1991a) was to tap into the time course of the naming process by presenting acoustic probes at different SOAs. Depending on the SOA, the probe enters the speech system at different activation states of the naming process: At a short SOA, the lemma of the picture's name should be active, whereas its phonological form should not be active. The naming process, in this case, should influence the semantic/syntactic access of the acoustic probe. In contrast, at a long SOA the picture naming should have a clear form activation. In this case, the naming process should hamper the phonological encoding of the acoustic probe.

In order to obtain a more salient time course between the speech levels in the model, the activation change of a unit (following Pineda) was propagated to the next speech level in a delayed fashion. This delay was realized in terms of delayed connections. For example, if a connection from unit x at the lemma level to unit z at the phonological level obtained a time delay td, the activation a(t) sent from unit x will arrive at unit z at time t + td. This delay leads to a spread of activation from one unit to units at the next level that does not start immediately after one time step, but after several time steps. All present model versions have connections with a time delay td = 20 between all speech levels in all directions.

4.6.3 Measuring lexical decision latencies

As in the real experiment, two kinds of latencies should be obtained. First, the baseline lexical decision latencies during single task performance. Second, the lexical decision latencies during dual task performance. The most direct way to collect these data was to define that a lemma gets selected for lexical decision if it reaches a certain activation threshold. The number of cycles it takes to reach the threshold was used to model response latencies.

The collection of the data was carried out by two additional units that were added to the model. These units did not influence the spreading of activation within the internal layers of the model. Their task was simply to measure the activation of the most active lemma unit. Given that more than one lemma could be active at the same time, the selection of the most active lemma was relevant. This selection was realized in the following way: All lemma units were connected to a 'max-unit' ($w = +1$). The max- unit had a so-called maximum activation function. This function guarantees that the max-unit can only be activated by the most active unit j. Its activation value is given by

$$a_i = \max (w_{ij}a_j) \ , \qquad (5)$$

The activation of the max-unit grows with increasing activation of the most active lemma. The max-unit projects its activation state further to a 'decision-unit' ($w = +1$). This unit defines its activation by means of a threshold function. If the incoming stimulation is higher than the decision-unit's threshold the unit's activation value is 1.0, otherwise it is 0.0. The threshold was set high (to 0.95 for all simulations), because firing of the decision-unit was used to indicate that one lemma unit finally won the competition. The firing was interpreted as a lexical decision on this item. The counting of the cycles started at the onset of the first acoustic input and stopped with the lexical decision.

4.6.4 Task specific attention shifts

During the dual task situation a participant had to follow the instruction to delay naming if an acoustic stimuli was presented, and was asked to carry out a lexical decision instead. This involves a *selection* of a different action. According to theories of attention this selection-for-action (Allport, 1987; Neumann, 1987, 1992) might lead to the inhibition of alternative responses: The naming task inhibits the lexical decision task, and vice versa. This general inhibition was found in the real data in terms of the difference between single task and dual task decision latencies. The observed longer reaction times during dual task performance may be due to relatively late response execution processes (Allport, 1987; Neumann, 1987). In contrast, it is also possible that the inhibition happens relatively early, namely already when the acoustic stimuli is perceived, but not yet fully identified (Neisser, 1967). As Allport (1987) pointed out, the question of 'late vs. early' selection is not yet answered in the literature. He even argues that "the controversy regarding 'early' versus 'late' selection has systematically confused 'selection' as selective *cueing* and 'selection' as selective *processing*.[3] Once the distinction is made clear, there may even be no controversy." (p. 409).
A computational model that has to perform a dual task should be able to selectively attend to the incoming acoustic probe instead of the picture name. According to the just mentioned assumption of selective attention, this process could either be located outside of the speech system (that is, 'late' after speech

---

[3] Selective *cueing* operates predominantly in terms of 'early' physical, or precategorical sensory attributes, whereas *processing* of both cued and noncued information proceeds at least to 'late' categorical levels of analysis (Van der Heijden, 1978)

encoding) or 'early' inside the speech system, or both. By stipulation, the selection was implemented in the model in terms of an attention mechanisms *within* the speech system. This was done because the main focus of the current simulation study was to investigate the nature of speech internal processes, and not speech external ones. The shift of attention from naming towards lexical decision due to incoming acoustic information was implemented in the model in terms of a short-term inhibition at the lemma level. It has to be noticed that, although the mechanism is speculative in nature, it affects all lemmas in all SOA conditions in the same way. Its technical details are described next.

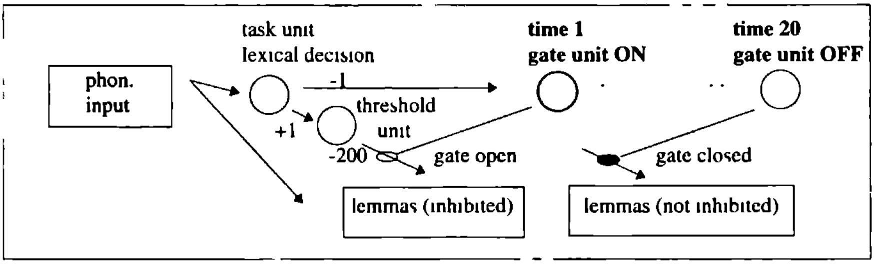


Figure 4.3 The attentional "lemma hemmer". An acoustic input signals not to name the picture but to carry out a lexical decision task instead. This leads to a short-term inhibition of all lemmas by an activated threshold unit Depicted are two moments in time At time 1 (immediately after acoustic input presentation), the gate unit is active and allows the threshold unit to inhibit all lemma units At time 20, the gate unit is de-activated and the gate for inhibition is closed Filled circles represent activated units Values denote connection weights.

The implementation of a "lemma hemmer": At the moment the acoustic stimuli enters the system the lemma level is generally inhibited for a short amount of time. This general attention-based inhibition mechanism is implemented by two units: A threshold unit and a so-called gate unit. The two units are connected to the task unit for lexical decision (see above in the learning section). The task unit is stimulated by the acoustic input. It sends activation to the threshold unit (w = +1.0). The threshold unit normally is inactive if no acoustic input is given. It has a bias of 0.1, which leads to its activation immediately after the acoustic input is presented. This unit has inhibitory connections towards all units of the lemma level (w = -200). The lemmas, therefore, would become inhibited as long as the threshold unit is active. However, a permanent inhibition of the lemmas would prevent the lemmas from becoming re-activated later on. Therefore, the inhibition has to be de-activated again. This is realized by the gate unit. A gate unit has the ability to open and close connections. This happens by so-called multiple weights from the gate unit to the inhibitory connections (for various applications of gate units see for example Goebel, 1996; Rumelhart, Hinton, & McClelland, 1986, p. 73ff; Van Kuijk, Wittenburg, & Dijkstra, 1996). The impact of a unit j (in this case the inhibitory effect of the threshold unit) on the unit i at the lemma level is dependent on the state of activation of the gate unit, following

$$net_i = w_{ij}^{a} a_j a_{s_{ij}} \quad , \qquad (6)$$

where $a_{s_{ij}}$ is the activation of gate unit s_{ij}. The activation of the gate unit s_{ij}, therefore, decides whether or not the activation of the threshold unit a_j can inhibit the unit i at the lemma level. If the gate unit is not active ($a_{s_{ij}} = 0$) the connection is closed. If the gate unit is active ($a_{s_{ij}} = 1$), the connection is open, and inhibition takes place. The gate unit follows a linear activation function involving activation values in the range of 0.0 and 1.0,

$$a_i = \min(\max(\sum w_{ij} a_{ij}, 0), 1) \quad , \qquad (7)$$

The gate unit has a bias of 1.0, which means that it is active before an acoustic stimulus presentation. However, it gets inhibited by the lexical decision task unit (weight from task unit to gate unit = -1.0) at the moment the acoustic input is presented. This leads to a continuous de-activation of the gate unit towards 0.0, which starts with the acoustic stimulus presentation. The gate unit is de-activated after 20 cycles. It then completely closes the inhibitory connections from the threshold unit towards the lemma level. In short, the lemma inhibition starts with the acoustic input and stops 20 cycles later.

Basically the same mechanism is used to decay the conceptual activation after a constant amount of time, assuming a constant decay of the concept that relates to the picture. The inhibition of the concept starts 60 cycles after the picture was presented. Here the inhibition is activated by stimulating a threshold unit that has inhibitory connections to all units at the concept level (w = -200). This threshold unit is stimulated by the visual task unit (w = 1.0, delay 60). The inhibition is canceled 20 cycles after the beginning of the acoustic presentation in order to allow the concepts to become active again due to the attention change toward the new task. The cancellation is carried out by a gate unit that is active first, and gets inhibited by the lexical decision task unit. The lexical decision unit gets active with the acoustic input, and inhibits the gate unit by a negative weight (w = -1.0) that is delayed by 20 cycles. This delay covers the longer pathway from acoustic input to the concept level in comparison to the lemma level.

Before we come to the simulation, a short remark has to be made concerning the pathway between acoustic input phonological level and the phonological level for naming. These connections were needed during the training phase in order to learn bottom-up connections from the form level of naming to the lemma level. As discussed above, there was also empirical support for the existence of these connections by picture-word-interference studies (for example, Roelofs, 1992a, b). However, the two paradigms, picture-word-interference and lexical decision, differ. The difference between the two tasks lies in the quality of treating the acoustic stimuli. During a picture-word interference task this stimulus has to be ignored, whereas during a lexical decision an active decision on the stimuli has to be executed. This active decision makes the direct access from the input towards the lemma level important. However, it was not clear beforehand whether the pathway of form input towards phonological level of naming was used at all during lexical decision. In pre-simulation it was found that these connections lead to a facilitation for phonologically related probes, as usually found in picture-word-interference experiments (for example, Schriefers et al., 1990). This preliminary finding was not in line with the observed inhibitory effect during lexical decision. This divergence led to the hypothesis that during a lexical decision task the pathway between phonological input and phonological level for naming might not be relevant - in contrast to a naming task that involves distractor words which have to be ignored (Roelofs, 1992a, b). Therefore, during all reported simulations the connections from the form input level towards the form level of naming were set to zero. By stipulation, the closing of the connections could be interpreted as being temporal and task specific.

4.7 Simulation 1: Noun naming including phonological feedback

The goal of the simulation was to investigate the impact of feedback connections from the phonological form of the naming process towards the lemma level. This feedback was assumed to be needed if the phonological inhibition effect found in the experiments was due to phonological mismatch effects, as assumed by O'Seaghdha et al. (1992).

4.7.1 The baseline simulation

As in the real experiment, first a baseline of lexical decision latencies was collected. Because in the baseline experiment no switch between tasks were required, no attention specific inhibition at the lemma level was assumed. Following this assumption, the lemma inhibition was not activated during single task performance.

The word initial phonemes were presented for 20 cycles. Then, the word final phonemes were presented for the remaining time. This led to an activation of all phonologically related lemma units up to 0.30 after 40 cycles. After this time, the activation of the non-target item decayed to zero. The cohort-like activation of the lemma units also activated the corresponding concept units (0.50 after 50 cycles). However, a potential top-down feedback from this conceptual activation towards the lemma units did not play a role during this single task performance, because the lexical decision was carried out before this feedback could reach the lemma level. The mean lexical decision latencies for all 16 items were 66 cycles.

4.7.2 The dual task simulation

During the dual task simulation the SOAs were chosen according to the internal activation state of the picture naming process. The SOA variation should test effects of the naming process onto the lexical decision latencies. Mainly, three different states were of interest: First, only the concept of the picture should be active. Second, the lemma should be active, but not the phonological form. Third, the phonological form of the picture's noun should be active. Therefore, three different SOAs were chosen, as in the Levelt et al. (1991a) study.

SOA = 10, because 10 cycles after presenting the visual input only the concept was active during picture naming (activation of the concept = 0.60, lemma activation 0.0)

SOA = 30, because 30 cycles after presenting the picture its lemma was activated (activation of the concept at this time = 0.94, of the lemma = 0.50, and of the phonological form = 0.0)

SOA = 50, because 50 cycles after presenting the picture the phonological activation was 0.30 (activation of the concept = 0.84, of the lemma = 0.90)

The visual input during the SOA simulation was presented for 40 cycles. The word initial input was presented for 20 cycles, the word final input for the remaining time, as in the baseline simulation. The picture-probe pairs used in the simulation are shown in Table 4.2.

Table 4.2 Pictures and probes used to test the computational model in the dual task situation. Each row shows the 5 probe words presented with the target picture.

identical to picture	semantically related	phonologically related	unrelated	phonologically related to sem.
1 bureau (desk)	stoel (chair)	buurman (neighbor)	cactus (cactus)	stoep (sidewalk)
2 cactus (cactus)	stekel (sting)	kakker(lak) (cockroach)	fietspomp (inflator)	steen (stone)
3 fietspomp (inflator)	band (tire)	file (queue)	geweer (rifle)	bank (bank)
4 geweer (rifle)	oorlog (war)	gewei (antlers)	bureau (desk)	oorbel (ear-ring)

The table depicts the experimental picture-probe-relations. Four different conditions of relatedness between picture and acoustic probe were investigated: Identical (bureau-bureau (desk)), phonologically related (bureau-buurman (neighbor)), semantically related (bureau-stoel (chair)) and unrelated picture-probe pairs (bureau-cactus (cactus)). It should be mentioned here, that additional picture-probe pairs were tested that satisfied the conditions (for example, presenting the item *buurman* as the picture, and *bureau* as the phonologically related probe). They basically behave in the same way as the four combination depicted in Table 4.2.

For each probe the number of cycles from probe onset until reaching the 0.95 threshold at the lemma level were measured. Every picture-probe pair was tested under each SOA condition. Errors were defined as selecting the wrong lemma, or as time-outs. A time-out was defined as lexical decision times that were longer than 150 cycles. One error per category occurred. These two errors were excluded from the data.

4.7.3 Results

The obtained lexical decision latencies during the dual task were subtracted from those of the baseline simulation. As in the real experiment the subtraction was carried out pair-wise, for example, the lexical decision for bureau (dual task) was subtracted from bureau (baseline). Figure 4.4 (bottom) and Table 4.3 show the mean lexical decision cycles for each probe condition for the three SOA conditions in terms of differential scores (baseline - main experiment).

4.7.4 Descriptive analysis of the reaction cycles

The present network model can be seen as one trained pseudo-participant. Its performance in the dual task situation does not involve variation: If the model is presented with the same input, it will produce the same output. As a consequence of this lack of variance, each difference in reaction cycles within the model is significant. This is the case because the cycle values are real values without noise.[4] The model's performance will now be descriptively compared to the real data. In addition to the descriptive analysis of the reaction cycle data, the model will also be analyzed in terms of weight patterns and time course of activation.

---

[4] As depicted in the result figures, some differences between the unrelated and the related conditions were more salient than others In order to have a qualitative decision criterion about what is a clear difference and what is not, in the discussion differences smaller than 15 cycles were seen as not salient.

Table 4.3 Mean differential scores (baseline - main experiment) of lexical decision cycles for each SOA during noun naming in a model that includes phonological feedback. Negative values show inhibition during the dual task performance in comparison to the single task performance.

	Probe conditions				
SOA (in cycles)	**identical to picture**	**semantically related**	**phon. related**	**phon. rel. to sem.**	**unrelated**
10	12.5	-20.3	-52.8	-10.0	-4.0
30	15.5	-21.0	-63.8	-12.0	-3.5
50	17.8	- 8.0	-67.5	-10.0	-3.3

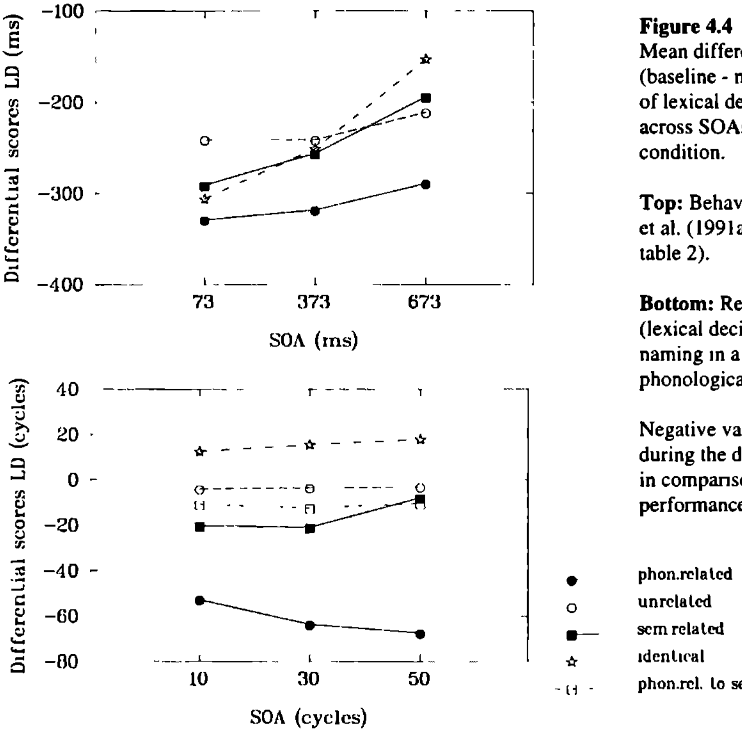


Figure 4.4
Mean differential scores (baseline - main experiment) of lexical decision latencies across SOAs for each probe condition.

Top: Behavioral data of Levelt et al. (1991a, Experiment 3, after table 2).

Bottom: Results of Simulation 1 (lexical decison during noun naming in a model that includes phonological feedback).

Negative values show inhibition during the dual task performance in comparison to the single task performance.

The *unrelated condition* [for example, the picture-probe pair *bureau* (desk) - *cactus* (cactus)] did not show an SOA effect. This result is in line with the empirical data of the Levelt et al. study (1991a). The constant number of cycles across all SOAs (mean: 3.5 cycles) can be interpreted as a baseline for lexical decision of acoustic probes in a dual task situation.

The *phonological condition* [for example, the picture-probe pair *bureau* (desk) - *buurman* (neighbor)] did not show an SOA effect either. However, as can be seen from the figure, the reaction times are inhibited in comparison to the unrelated condition. This result is also in line with the empirical data of both the Levelt et al. study and the German replication (first main experiment in Chapter 3).

The *semantic condition* [for example, the picture-probe pair *bureau* (desk) - *stoel* (chair)] revealed an SOA effect. As can be seen in the figure, the reaction times were inhibited at the short SOA in comparison to the unrelated condition. This inhibition nearly disappeared at the long SOA. The same SOA behavior was found in the data of Levelt et al.

The *identical condition* [for example, the picture-probe pair *bureau* (desk) - *bureau*] did not show an SOA effect. As can be seen in the figure, the reaction times were facilitated in comparison to the unrelated condition. At the short SOA the amount of facilitation was 16.5 cycles, at long SOA it was 21 cycles. This indicated a trend for an increase of facilitation with increasing SOA. This result differs from the Levelt et al. data. The authors found identical inhibition at the short SOA and facilitation at the long SOA. However, during the replication experiment (Chapter 3) a significant facilitation was observed at the short SOA.

The *mediated priming condition* [for example, the picture-probe pair *bureau* (desk) - *stoep* (sidewalk), which is phonologically related to *stoel* (chair)] did not show an SOA effect. As can be seen in Figure 4.4 (bottom), the reaction times were slightly inhibited in comparison to the unrelated condition. However, there is no clear difference between the two conditions. This result is in line with the Levelt et al. data (experiment 5 and 6) where phonological related probes to semantically related alternatives of the presented picture did not show an inhibitory effect at a short SOA.

In summary, the model covers the major empirical findings: No SOA differences for unrelated probes, clear phonological inhibition across all SOAs, and early but no late semantic inhibition. It produced early identical facilitation, which was found in the replication study (Chapter 3, first main experiment), but not in the Levelt et al. experiment.

4.7.5 The descriptive analysis of the weight patterns

The connection weights between concept and lemma level for one item are presented in Figure 4.5. It shows the concept of bureau (desk) and its connections to the relevant lemma units.

As can be seen from the Figure 4.5 the model set the weight values during training so that corresponding units stimulate each other. Unrelated conceptual and lemma units inhibit each other.

The pattern of connection weights - in part - explains the facilitatory effect found for the identical condition at short SOA. Because only exitatory connections are involved, the already developed activation due to the visual input cumulates at the lemma level together with the matching phonological input (see below). As a result, the identical lemma reaches the lexical decision threshold fast.

Also shown in the figure is the solution the model choose for the semantic representation. The conceptual unit that is shared by the two semantically related concepts has nearly no effect on the semantically related lemma. The connection weight is relatively small (*w* = -1). In addition, the semantic mismatch between the concept of "bureau" and the lemma "stoel" is also shown. The mismatch is given by the second conceptual unit (depicted in gray) that is not shared by the semantically related concept representation. The mismatch is inhibitory in nature, given a relatively large inhibitory connection weight between the conceptual unit and the lemma "stoel" (*w* = -28).

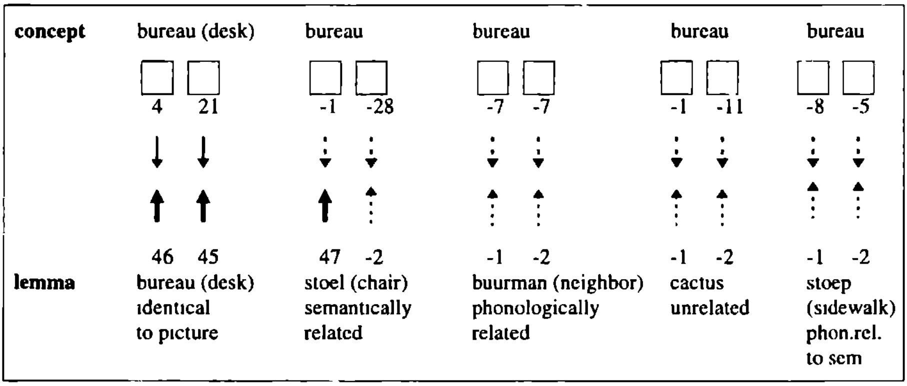


Figure 4.5 Connection weights between the two concept units of "bureau (desk)" and the relevant units at the lemma level. Arrows that point down represent connections from the concept to the lemma, and vice versa. Exitatory connections are shown by line arrows, inhibitory connections by dashed arrows. The weight value of each connection is written next to the corresponding arrow. The gray colored conceptual unit indicates the unit that is not shared by the semantically related probe "stoel (chair)".

This strong inhibitory connection leads to early semantic inhibition (at short SOA). This is the case because at short SOA the concept of the picture is fully active and continuously inhibits the semantically related lemma. This inhibition suppresses the acoustic-input-driven activation process of that target lemma. As a result, the lexical decision on semantically related probes is inhibited. At long SOAs, however, the impact of conceptual inhibition becomes smaller due to conceptual decay of the picture's concept. Feedback of activation of the picture's lemma towards the concept, and from there back to the lemma level again does not effect the lexical decision anymore. This is the case because early pictorial lemma activation was inhibited by the attention gating mechanism, leading to nearly no feedback later in the process. No obvious differences in the

weight pattern between lemma and concept level were observed for the unrelated, the phonological and the mediated priming condition.
The connection between lemma level and phonological level of naming are shown in Figure 4.6. The descriptive analysis of the weight pattern of connections from the lemma level to the form level of naming revealed a systematic pattern of the model's solution concerning phonologically related items. The connections from overlapping phonemes (e.g. /b/ and /y/ for the lemmas of "bureau" and "buurman") got only a small weight value (mean *w* = -0.2). Whereas the non-overlapping phonemes (/r//o/) inhibited the lemma "buurman" by choosing huge negative weights (mean *w* = -18).

This weight pattern explains the phonological inhibition at late SOA. The more the pictures' phonemes become activated (at long SOA), the more they inhibit, via feedback, the phonologically related probe lemma unit. This means that the non-overlapping phonemes were responsible for the inhibitory effect at late SOAs. This result is in line with findings in a lexical decision study of Marslen-Wilson (1990). In this study, a lexical decision had to be carried out for a target that was preceded by a prime. If prime and target had an initial overlap of several segments, such as the prime-target pair "streak-street" Marslen-Wilson observed inhibition of the lexical decision to "street". In contrast, if the prime was just a word fragment, such as "stree", facilitation was found. The author interpreted the inhibitory effect due to mismatch of the final segments.[5]

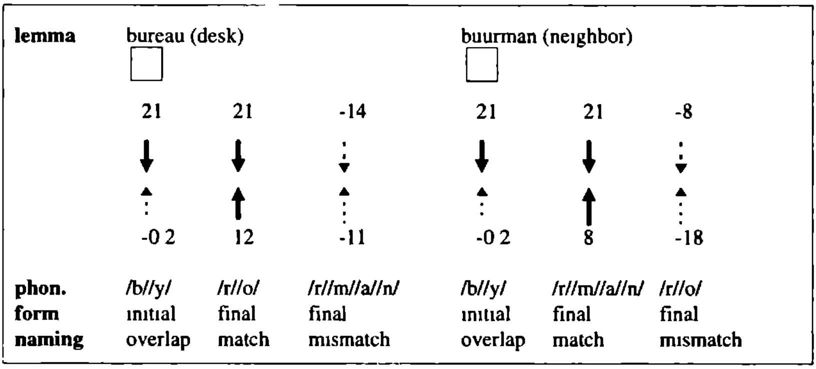


Figure 4.6 Mean connection weights between the two phonologically related lemmas "bureau (desk)" and "buurman (neighbor)" and the relevant phonemes at the phonological form level for naming Arrows that point down represent connections from the lemma to the phoneme units, and vice versa Exitatory connections are shown by line arrows, inhibitory connections by dashed arrows The mean weight value, averaged across connections between lemma and the depicted phonemes, is written next to the corresponding arrow.

---

[5] It should be mentioned here that the pattern of results for form priming and inhibition in the literature is quite divergent It depends on the task, on the amount and position of overlapping segments, on the modality of target and prime presentation, and on the SOA (for a review see Zwitserlood, 1996)

The early phonological inhibition (at SOA = 10) was initiated due to conceptual inhibition. As shown in Figure 4.5, the concept of the picture gives moderate inhibition to unrelated lemmas. This moderate inhibition leads to a suppression of the appropriate selection of the target lemma for lexical decision. This in turn gives the phonological mismatch units more time to become active, in the sense of having an under-cover long SOA at hand. The interaction of conceptual inhibition and phonological feedback inhibition led to the constant phonological effect.

The weight pattern between lemma and phoneme level of naming also explains the facilitatory effect found for the identical condition. As for the concept-to-lemma connections, all matching connections are positive, leading to a support of the selection of the identical lemma.

The connections between phonological input and lemma level are depicted in Figure 4.7. The model activates the cohort at the lemma level by using positive connection weights from the input of the word initial overlap towards the corresponding lemmas (mean w = +3). The word final input, then, disambiguates the activation at the lemma level by stimulating its own lemma unit, and by inhibiting the alternative unit. However, the cohort-like activation did not play a role in the current simulation. This was the case, because the activation of the picture's concept led to inhibitory effects at the lemma level that were more powerful than the positive activation via the acoustic input. As a result, no cohort activation occurred during SOA simulation in this model version. Therefore, effects of cohort mechanism can be excluded as interpretation of the observed phonological inhibition effect during single noun naming.

4.7.6 Discussion

The weight pattern analysis in combination with the observed activation processes revealed an explanation for early semantic inhibition, the lack of mediated priming, continuous identical facilitation, and the phonological effect observed in the model's performance.

The *semantic effect* is in line with empirical findings of semantic inhibition during picture naming (Schriefers et al., 1990; Roelofs, 1992a, b). It also covers the SOA function for semantic inhibition during lexical decision found by Levelt et al. (1991a). The analysis of the model's performance corresponds to Levelt et al.'s explanation for early semantic inhibition due to co-activation of semantically related words. In addition, because the conceptual activation decays over time at long SOAs the inhibition effect disappears. However, the exact nature of the inhibition effect slightly differs between assumptions of Roelofs (1992, p.51) and the present model. Roelofs assumes a local representation for each concept, and connections between semantically related concepts in a semantic field, for example, FISH and DOG. A presented picture of a dog, therefore, activates its concept DOG, and due to positive connections it also activates the related concept FISH. In turn, both concepts activate their lemmas, *fish* and *dog*. A presented acoustic word, for example "fish" activates its lemma, then its concept, then the semantically related concepts, which in turn activates its lemma.

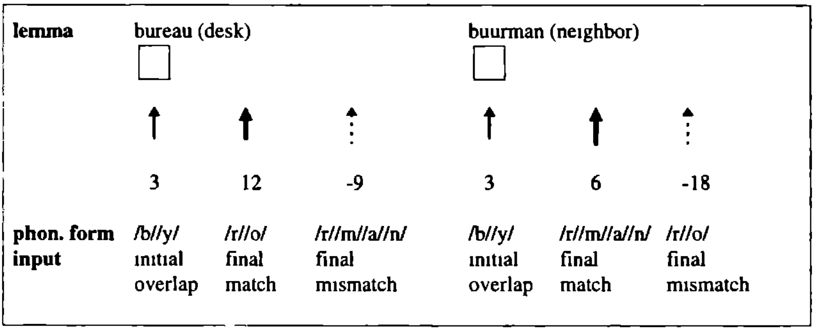


Figure 4.7 Mean connection weights from the phonological input level to the two phonologically related lemmas "bureau (desk)" and "buurman (neighbor)" Exitatory connections are shown by line arrows, inhibitory connections by dashed arrows The mean weight value, averaged across connections between the depicted phonemes and the corresponding lemma, is written next to the corresponding arrow

According to Roelofs, the semantic inhibition is due to a trade-off between the priming of a distractor lemma node by the picture, and the priming of a target lemma node by the distractor. Due to different path-length of the encoding of acoustic and visual stimuli, the picture will prime the distractor lemma node faster (DOG -> FISH -> *fish*) than the distractor word will prime the picture lemma (*fish* -> FISH -> DOG -> *dog*). In addition, the picture lemma is highly active because of the naming process. It now has to compete with the highly active distractor lemma. This competition takes time, which leads to semantic inhibition in comparison to an unrelated condition where this priming does not take place.

In contrast, the current model does not have between-concept connections. Here, inhibition is mainly a top-down process from the concept to the lemma level. This inhibition is stored in the connections between the concept and the lemma level. The model had to solve the task - during training - to disambiguate between a partially overlapping concept representation and its distinct representation at the lemma level. The model solved this task during training by setting negative connection weights between not shared conceptual units and the lemma of the semantically related concept (see Figure 4.5, the connections between the concept "bureau" and the lemma "stoel"). This negative connection weight leads to semantic inhibition if the picture concept is active and a semantically related acoustic word lemma has to be selected. The picture's concept, "bureau", inhibits the activation of the lemma of its semantically related concept "stoel".

The null effect for the *mediated priming condition* is in line with the Levelt et al. findings. Levelt et al. interpreted the lack of mediated priming as evidence for a strict two stage theory. They assumed that the semantically related concept is co-active with the picture, for example "goat" with the picture of a sheep. But, because it is not selected for naming, its phonological form does not become active - shown as a null effect in the mediated priming condition for "goal". The present model also produces a null effect. However, the model's null effect is due to the fact that the semantically related concept is only partially active, namely by the unit that is also shared by the active picture's concept. The partially activated concept, however, has no systematic inhibitory effect to phonologically related lemma. As can be seen in Figure 4.5, both conceptual units of "bureau" inhibit the lemma "stoep" to the same amount, comparable to other conceptually unrelated conditions.

The *identical facilitation effect* of the model and my own data are in contrast with the finding of Levelt et al. (1991) and needs further empirical investigation in order to determine its nature.

The *phonological inhibition effect* observed in the model's reactions times also matches the empirical findings. In the model, it is due to an interplay between early moderate conceptual inhibition and phonological feedback. The moderate conceptual inhibition prevents the lemma for a while becoming selected for lexical decision. This delay, in turn, gives the phonological feedback some time to develop - even at an short SOA. The phonological inhibition at long SOA is due to phonological feedback only.

The explanation of the phonological effect due to feedback of phonological information is basically in line with the explanation of O'Seaghdha et al. (1992). The authors also assume a competition of mismatching segments between prime noun (in our situation the picture's noun) and the lexical decision noun. However, the exact nature of the phonological effect differs between the O'Seaghdha et al. model and the present one.

O'Seaghdha and colleagues locate the segmental competition directly at the form level due to residual activation of the prime at that level. The authors' explanation will be outlined by describing their interactive activation model of form-related priming. The model has three distinct levels: letters (input), lemmas, and segments. It processes stimuli in a syllabic CVC (consonant, vowel, consonant) structure, such as *cat*, *cap*, *cad*, *peg*, *pen*, and *pez*. If a word is presented, activation spread throughout the model, from letters to lemmas, from lemmas to segments and letters, and from segments to lemmas. Nodes at one level activate nodes at adjacent levels via facilitatory connections if they

correspond. Inhibition is obtained via inhibitory connections between non-corresponding nodes. The input at the letter level for the word *cat* is coded as $C_1A_2T_3$, where the numbers indicate the position of each letter in the word. The same holds for the segment level, which serve as a naming output. The words are coded at the lemma level in terms of one unit per word.
The model consists of words that realised the conditions of a priming experiment. Thus, the word *cat* serves as a related prime, and *peg* serves as an unrelated prime to the target *cap*. A presented prime activates its target lemma, and orthographically similar lemmas to a lesser extent. In the case where *cat* is the prime, the lemma *cat* would be highly activated, and *cap* would be slightly activated. After some time steps, links are created from the prime's letter nodes to a so-called episodic node, and from the episodic node to the prime's segments. In this way, the episodic node is recruited as a temporary memory of the prime. A presented target, for example *cap*, spreads activation throughout the system in the same way as the preceding prime, except that now the episodic links are also present. If the target is related to the prime, i.e. it shares letters with the prime, the target activates the episodic node: The presentation of a related target re-activates the prime, and therefore, indirectly reinstates the prime's segments, which influences the activation of the lemma nodes in terms of a memory trace. Response latencies were simulated indirectly in terms of the probability of misselecting the critical third segment in an additional word frame that is linked to the lemma nodes (in the case the /t/ instead of the /p/) at a point when the target's lemma had reached full activation.
The present model explains the phonological inhibition by an interaction of activation of mismatching units and specific weight patterns of feedback connections. Mismatching units get activated by the picture naming process. This activation is then sent back to the lemma level. Because of the learned inhibitory weights of connections from mismatching units at the form level to the corresponding phonologically related target lemma, the selection of this target lemma is delayed. The model accounts for the empirical data without assuming an additional feature, such as an episodic node.

In summary, the present model, which includes feedback connections between form level of naming and lemma level, accounts for the empirical data. It presents an explanation for the effect of phonological inhibition during noun naming. The inhibition obtained in the model is due to phonological feedback of mismatch information. A more detailed discussion is postponed to the end of this chapter. In the next section, a second version of the model is described that addresses the alternative explanation for the inhibition effect. As discussed in the preceding chapter, a second account for the observed phonological inhibition effect is the cohort-account.

4.8 Simulation 2: Noun naming including cohorts at the lemma level

According to Colombo (1986) the incoming acoustic probe activates a cohort of word candidates at the lemma level. When the complete word enters the system, this cohort is normally reduced to one target element and a lexical decision can be executed. The phonologically related noun lemma of the picture's object is also initially a member of the probe cohort. In addition, it also becomes activated by its concept due to the initiation of the picture naming process. Because the picture's noun lemma is a member of the cohort and gets additional activation from the naming process, it is a strong competitor to the acoustic probe lemma. As a result, the cohort reduction process and the lexical decision on the acoustic probe is delayed.

The architecture and details of the model, as well as the material, the presentation times and the SOA conditions remained the same as in the preceding simulation. Only two modifications were carried out. First, the feedback connections from the form level for naming to the lemma level were cut out, because a strict stage theory does not assume such a feedback. Second, in order to obtain a strong cohort effect, all connections of initial phonemes at the form input level towards the lemma level were increased by a factor of 10. As discussed in the previous section, the trained connection weights from the form input to the lemma level were always smaller than the connection weights from the concept to the lemma level, leading to no cohort effect in the self-trained model. This is due to the fact that the model has not been trained on the dual task, but was trained on single task performance. It had, therefore, no information during the training phase about balancing parallelly presented inputs of conceptual and phonological information at the lemma level.

The above mentioned manual manipulation of the connection weights from phonological input units to the lemma units, however, led to a salient cohort effect during the dual task performance. After presenting the initial word input for 20 cycles, both members of the cohort became activated up to 0.90. By presenting the word final input for the remaining time, the cohort was reduced to one element.

4.8.1 The baseline simulation

As in the preceding simulation the baseline latencies for lexical decision were collected without involving picture naming. This resulted in a mean lexical decision baseline of 50 cycles.

4.8.2 Results

Again, differential scores were obtained by pair-wise subtracting each items latencies during the dual task situation from that of the baseline simulation. They are depicted in Table 4.4 and Figure 4.8. As can be seen in the figure, the performance of the model revealed comparable results to the preceding version with respect to the *unrelated condition* and the *mediated priming* condition. This means that cohort effects did not influence the latencies in these conditions. The model performed differently with regard to the identical, the semantic, and the phonological condition. The *identical condition* does not show facilitation anymore in comparison to the unrelated condition. The effect size for *semantic probes* at short and medium SOAs is larger than in the preceding model. It is also larger than that of phonological probes. The form of the SOA functions remained the same.

The *phonological condition* showed the typical phonological inhibition at early and medium SOA. It is decreased in comparison to that of the former model. Interestingly, a clear phonological inhibition was not obtained at long SOA, showing only an inhibition of -8.7 cycles in comparison the -2.3 cycles of the unrelated condition. In contrast, at this SOA the former model version showed a related-unrelated proportion of -67.5 : -3.3.

Table 4.4 Mean differential scores (baseline - main-experiment) of lexical decision cycles for each SOA during noun naming of the cohort model (Simulation 2). Negative values show inhibition during the dual task performance in comparison to the single task performance

	Probe conditions				
SOA (in cycles)	**identical to picture**	**semantically related**	**phon. related**	**phon. rel. to sem.**	**unrelated**
10	0.0	-48.3	-27.8	-10.0	-5.3
30	1.0	-43.8	-29.5	-12.0	-3.8
50	2.0	-9.8	-8.7	-10.0	-2.3

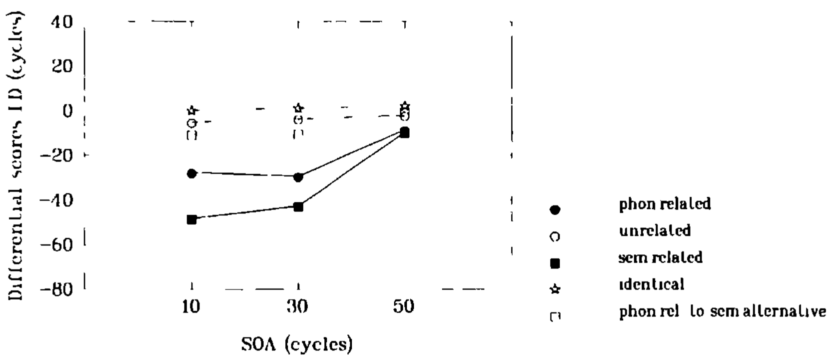


Figure 4.8 Mean differential scores (baseline - main experiment) of lexical decision cycles per probe condition at each SOA during noun naming in a cohort model. Negative values show inhibition during the dual task performance in comparison to the single task performance.

4.8.3 Discussion

Because the unrelated condition, as well as the mediated priming condition did not differ to the former version with regard to their SOA functions, they are not discussed here.

The increase of *semantic interference* at short and medium SOA (in comparison to the semantic effect size of Simulation 1) can be explained in terms of the increased cohort process. The target lemma became initially inhibited by the picture concept. This inhibition prevent the target lemma to become selected early (i.e., as early as in the baseline lexical decision). The inhibition also enables the cohort members to become fully active at the concept level. The concepts of cohort members - in addition to the picture concept - inhibit the selection for the target lemma at short and medium SOA. However, because of the decay of the picture's conceptual activation at long SOA, the initial conceptual inhibition of the target lemma (that has been observed at shorter SOAs) did not take place anymore. The conceptual activation (at 20 cycles after acoustic input presentation) is about 0.40 at long SOA, in comparison to 0.84 and 0.94 at medium and short SOA. The smaller amount of conceptual activation at long SOA (in contrast to shorter SOAs) had the consequence that the target lemma was not (or only slightly) inhibited. The target lemma, therefore, became selected before the cohort members could activate their concepts and feedback this information.

In the *identical condition* this model version behaved nearly as the unrelated condition. The identical probes became slightly speeded up at the lemma level in comparison to unrelated probes. This is due to support of the lemma selection process by conceptual activation of the identical picture naming process. The facilitatory effect is smaller than in the preceding simulation, because no additional stimulation due to phonological match feedback takes place.

In principle, the *phonological condition* worked in the same way as the semantic condition, except for the effect size. In comparison to the semantically related conceptual inhibition, the general conceptual effect is smaller. Concepts inhibit unrelated lemmas by inhibitory connections of the weight range of -12 to -14, whereas concepts inhibit semantically related lemmas by negative connections of the size -28 (see Figure 4.5).

To summarise, the model that includes a cohort mechanism simulates the relevant empirical data (Levelt et al., 1991a, see also the top of Figure 4.4) with respect to the semantic SOA function, the zero effect during mediated priming, and with regard to phonological inhibition at short and medium SOA. However, this model lacks late phonological inhibition (-8.7 at SOA = 30 cycles in contrast to -27.8 and -29.5 at shorter SOAs, see Table 4.4 and Figure 4.8). Such a late phonological effect was observed in the experimental data of Levelt et al. In fact, in the Levelt et al. study the phonological effect, defined as the difference between the unrelated and the phonologically related condition, is salient across all SOAs (see Figure 4.4, top panel).

In contrast to the cohort model, the feedback model (Simulation 1) showed a clear phonological effect across all SOAs (-52.8, -63.8, and -67.5 cycles at SOA = 10, 20, and 30, see Figure 4.4, bottom panel). A comparison of the two computational models described above, the one with feedback and the one including cohort processes, therefore, favors the feedback model.

4.9 Simulation 3: Pronoun naming including phonological feedback

The comparison of the preceding simulation results favors the model variant that includes phonological feedback from the form level to the lemma level. Therefore, this model version was trained to distinguish between noun and pronoun naming. In order to handle this, the architecture of the model was slightly changed.

4.9.1 Additional network features for pronoun generation

As can be seen in Figure 4.9, two processing stages were added to the already known model architecture: A discourse input stage and a pronoun unit. Furthermore, the naming output was extended by an additional unit that represents a pronoun naming output.

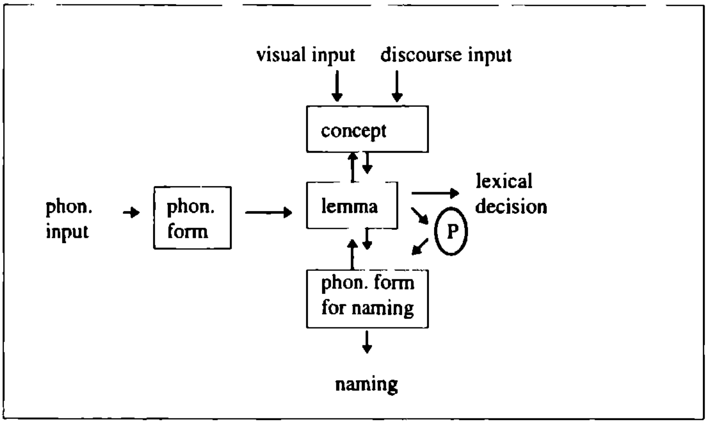


Figure 4.9 The architecture of the computational model for pronoun generation. P depicts the pronoun node.

The discourse module

The discourse module consists of 16 units - as the visual input. It serves as a memory system that signals the speaker (or the model) whether or not a presented picture was shown in the preceding experimental trial. It is fully connected to the concept (connection delay = 20). As reported in Chapter 3, during the pronoun experiments two major utterance conditions were involved: The noun naming and the pronoun generation. These two utterance formats were elicited by a sequence of two pictures. If the pictures contained two different objects, a noun naming had to be executed by the participant, such as *The sun is green...The flower is red*, i.e. when the first depicted object was a green sun and the second object was a red flower. If the pictures contained the same object, a pronoun generation for the second picture description was executed, such as *The flower is green....It turns red.* In this case, the first picture's object changes its color during its second presentation.

As it was discussed in Chapter 3, pronoun generation might be controlled by conceptual information, discourse information, and syntactic information. The combination of conceptual and discourse information could be formulated in an abstract way as an IF/THEN rule: IF the picture has been presented before and has been described, THEN use the pronoun in order to refer to the object at the second picture. This rule can be seen as a process that takes place during "microplanning" of the generation of the message. It should signal that the focused item is accessible in the current discourse (for a review see Levelt, 1989, p. 144ff). When the IF/THEN condition is met, the model should use the pronoun instead of the noun. This condition is implemented in the model by two different kinds of input representations, one for each naming condition.

If a noun naming has to be executed, the picture input is activated as usual, and the discourse module has no input. This empty discourse module is seen as simplification of having no "old" or "given" information available at the moment. In this case the picture activates its concept, its lemma, and its phonemes, and the model names the picture as usual.

If a pronoun has to be generated, the picture input is active simultaneously with its corresponding discourse unit. This unit signals that the item was presented before. The model should now activate the concept of the depicted object, then its lemma - and via the lemma the corresponding pronoun, which in turn activates its phonemes that lead to a pronoun naming. The pronoun unit represents one abstract pronoun only, for matters of simplification. All lemma units are connected to this unit (connection delay = 20). The links between lemma units and the pronoun unit reflect the assumption that the access of the pronoun is a syntactic process, which is governed by the lemma (Levelt, 1989, see also figure 3.2 in Chapter 3). A delay of these connections means that it takes some time to access the pronoun. Because no empirical data exists about the amount of time it takes to access the pronoun, the delay was set to the same size as all other delays in the speech system. The pronoun unit is connected to all phoneme units (connection delay = 20).

The training of pronoun generation

The model was trained in the same way as the preceding versions, in a modular way, and separately for the three tasks it had to execute. First, it was trained on the acoustic input in order to carry out a lexical decision. After this step the weights of connections from the acoustic input to lemmas, from the phonological level for naming to lemmas, and from lemmas to all conceptual units were frozen. Second, it was trained on single noun naming. After this training, the weights of connections from the concept level to the lemma level to the phonological level to the naming output for noun units were fixed. Third, it was trained to generate pronoun naming by learning the weights of connections from the discourse input to the concept, from the lemma level to the pronoun unit, and from the pronoun unit to its phonemes, in this case the phonemes the Dutch pronoun /h/i//j/ (he).[6] This stepwise learning roughly represents the stages during language acquisition: First, a child learns to use nouns, and later on it starts to produce pronouns (Mills, 1985; Deutsch et al., 1994).
After training, the pattern of connection weights for the lexical decision pathway and the noun naming pathway were the same as in the preceding model versions. The pronoun pathway is represented by positive connections form target units at the discourse unit to the corresponding conceptual units. In addition, all lemma units stimulate the pronoun unit via connections with a positive weight in the range of 96 to 99. The pronoun unit has positive connections to its corresponding phoneme units (weights approx. 90), while it inhibits other phoneme units by negative connections (weights approximately -90).

A pronoun gate

Given the trained connection weights after the last learning step, an activated lemma would always automatically activate a pronoun due to the positive connections from the lemmas to the pronoun units. However, this is not what a child or the participant does, and this is also not a desired model performance. Each lemma should only activate its pronoun in a specific discourse situation, by following the above mentioned IF/THEN rule. This selective access to pronouns led to the idea that the connections from lemmas to pronouns should only be open, if a pronoun has to be selected. Otherwise, these connections should be

---

[6] In fact, the computational model served as an explanation for the pronoun results obtained for German and not for Dutch. However, because the model is unaware of possible language specific differences (at the moment) the Dutch material should not be a problem now.

closed. This gating should be discourse input driven, because depending on the discourse information a pronoun should be generated or not.

Such a gating mechanism was already introduced with respect to aspects of attention. It is implemented for pronoun gating by adding a gate unit to the model. This gate unit becomes stimulated by the discourse representation (as it was already proposed in figure 3.2, Chapter 3). If the discourse is "empty", the gate unit is inactive and closes the connections from the lemma level to the pronoun unit. If the discourse signals that a picture re-appeared by activating the picture's discourse unit, it activates the gate unit, which in turn opens the connections between all lemmas and the pronoun unit. Technically, this gate unit is a simple threshold unit with a bias of 0.1. It gets input from all discourse units ($w = +1$). If one discourse unit is active in order to signal that a pronoun condition is met, the gate unit becomes active. If the discourse units are not active, the gate unit is inactive. This gating mechanism results in the flexible performance of the model to produce noun or pronoun naming interchangeably.

4.9.2 The testing phase

During testing the same procedure was applied as in the preceeding versions. The dual task was simulated for noun and pronoun naming, respectively, using the same picture and acoustic probe material, and the same SOAs as during Simulation 1 and 2. In addition, the discourse input could either be de-activated, leading to a noun naming, or it could involve one active unit, leading to a pronoun generation. First, baseline lexical decision latencies were collected. They did not differ from those of the first simulation. Second, dual task lexical decision latencies were collected and subtracted from the baseline latencies.

4.9.3 Results

During the pronoun experiment described in Chapter 3 a phonological condition was compared to an unrelated condition for both the pronoun and the noun naming condition. No semantic or mediated priming conditions were tested empirically. Therefore, only the phonological and unrelated condition will be reported here. The remaining probe conditions behaved approximately as in Simulation 1. Figure 4.10 and Table 4.5 show the results. As can be seen in the table the unrelated condition behaves identically for the noun naming and the pronoun generation. Therefore, Figure 4.10 depicts the unrelated condition for noun naming only.

Table 4.5 Mean differential scores (baseline - main experiment) of lexical decision cycles for each SOA during noun and pronoun generation in the phonologically related and unrelated probe condition. Negative values show inhibition during the dual task performance in comparison to the single task performance.

	Noun generation		**Pronoun generation**	
SOA (in cycles)	**phon. related probe**	**unrelated probe**	**phon. related probe**	**unrelated probe**
10	-52.8	-4.0	-53.3	-4.0
30	-63.8	-3.5	-31.8	-3.8
50	-67.5	-3.3	-14.5	-3.2

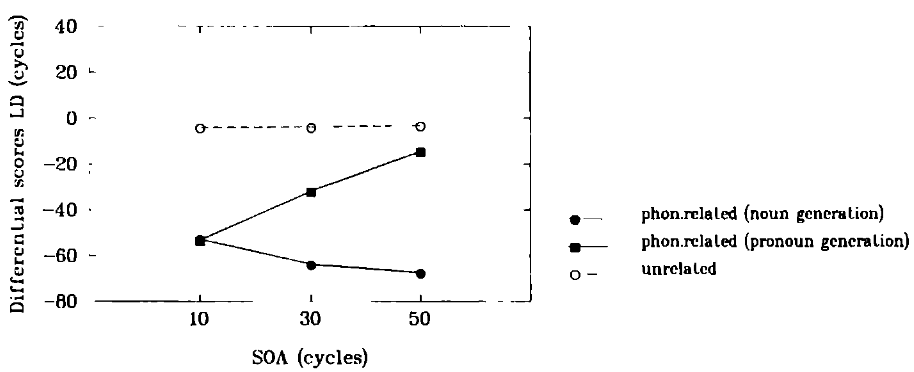


Figure 4.10 Mean differential scores (baseline - main experiment) of lexical decision cycles for unrelated and related probes at each SOA during pronoun and noun generation (Simulation 3). Negative values show inhibition during the dual task performance in comparison to the single task performance.

The *unrelated condition* for the noun naming and the pronoun generation condition are basically identical. They revealed the same pattern of results as obtained in the first simulation, and can be seen as a baseline for lexical decision during picture naming.

The *phonological condition during noun naming* is also identical to that of the first simulation, showing relatively constant phonological inhibition.

The *phonological condition during pronoun generation* showed the same amount of early inhibition as during noun naming (the difference between the unrelated and the related condition at SOA = 10 is 64 cycles). However, across SOA this inhibition constantly decreases. The difference between unrelated and phonological condition at SOA = 50 is 11.3.

4.9.4 Discussion

The *phonological inhibition during noun naming* can be explained in the same way as for the first simulation results, because basically the same activation pattern and connection weights are involved. It developed due to the interaction of moderate early conceptual inhibition and phonological feedback inhibition.

The *time course of phonological activation during pronoun generation* can be explained by considering three components: Early moderate conceptual inhibition, and as a consequence of this, phonological mismatch feedback - as during noun naming. In addition, the different pathway lengths of phonological activation of pronouns and nouns play an important role.

During pronoun naming, the naming path is as follows: The picture activates its concept, and then its lemma. The lemma activates the pronoun, but at the same time it spreads activation towards the form level. Note that the activation of the picture's noun lemma does not differ between pronoun and noun naming. Due to the delay of connections between lemma level and pronoun unit, the phonological activation of the pronoun arises approximately 20 cycles later than that of the picture's noun. The activated pronoun stimulates its corresponding phonemes and inhibits the phonemes of the picture's noun.

The *early phonological inhibition* during pronoun generation was initiated due to moderate conceptual inhibition. As shown in Figure 4.5, the fully activated concept of the picture gives moderate inhibition to unrelated lemmas. This moderate inhibition leads to an early suppression of the appropriate selection of the target lemma for lexical decision. This in turn gives the phonological mismatch units more time to become active. These phonological mismatch units of the picture's noun can become active due to spread of activation of the picture's noun lemma towards the phonological level, even during pronoun generation.

The *lack of phonological inhibition at long SOA* can be explained by two components. First, at long SOA the conceptual activation of the picture's noun is already decayed to some extent, leading to nearly no inhibition of the target lemma. Second, at long SOA the pronoun's phonological activation already starts to inhibit the noun's phonological activation. The pronoun therefore, inhibits also the mismatching units. This leads to a clearly smaller amount of phonological mismatch inhibition of the lexical decision.

In summary, the model simulated the early phonological inhibition effect for both noun and pronoun generation. It, therefore, covers the empirical results of the pronoun study, described in Chapter 3. During these experiments, phonological inhibition was found for the two naming conditions at an SOA = 100 ms. Furthermore, the model gives a prediction for a dual task outcome at longer SOAs that has not been tested yet empirically. Future experiments can easily falsify or support this model by investigating lexical decision effects during pronoun generation at longer SOAs.

4.10 Conclusion

As for all simulations, not much in general can be concluded from a specific simulation result, unlike a mathematical proof. However, the present computational model in its different variants was meant as a tool that could help to gain a better understanding of the ongoing processes during the dual task situation. The analysis of its reaction performance with regard to a combined effect of weight patterns and activation spreading outlined one explanation for the observed empirical findings.

The described Simulations 1 and 2 compared two alternative accounts for the observed phonological inhibition during pronoun naming: A phonological competition effect (O'Seaghdha et al., 1992) vs. a phonological cohort effect at the lemma level (Colombo, 1986; see also Roelofs et al., 1996). According to the simulations for noun naming the phonological competition account was favored due to a better match with the empirical data of Levelt et al. (1991a). Basically in line with a phonological competition account of O'Seaghdha et al., the present model explains the observed inhibition in a different way. O'Seaghdha et al. located the competition directly at the phonological level due to segmental mismatch competition during segment-to-frame placement of target and prime. In contrast, the present model does not locate the competition at a particular level. According to the present model, phonological inhibition occurs because of the interaction of activated phonemes and learned weight patterns. Activated phonemes of the picture naming process send inhibition via self-organized negative feedback connections to the phonologically related lemma. As a result, the selection of the target lemma for lexical decision is delayed.

Limitations of the model: As every computational model, the current one simplifies matters. Its architecture shows its limitations: First, although the model was designed in order to simulate activation processes during pronoun generation it can only process one item at a time for naming. It is not capable of producing connected speech. Second, its so-called discourse representation is limited to signal the re-appearance of only one potential item, and it can only process a very simple IF/THEN rule. The discourse module has not been trained to handle additional discourse processing rules. Third, the conceptual representation involving only two units for a concept is far from modeling the complex structure of human conceptual memory. Fourth, the model is able to handle only 16 words. For a realistic test of its performance, the lexicon should be increased by thousands of different words. Fifth, the pronoun unit can be seen as an abstract pronoun representation. In languages with several grammatical genders, there should be several corresponding pronoun representations. Sixth, the phonological level for naming does not include sequential processing yet, because the phonemes are not coded for word initial or word final positions (as it is proposed by other models, for example, by that of O'Seaghdha et al. (1992). It also lacks more detailed representations such as morphemes and syllables (as proposed in other models, for example, Roelofs, 1996). Seventh, the model was not designed to handle speech error data. Eighth, it has no built-in mechanism for simulating negative lexical decision on pseudo-word presentations.

In spite of its limitations - and there are surely many more than those listed - the model incorporates various empirically supported insights in activation processes during single task and dual task performance of picture naming and lexical decision. In addition, the model gives room for future and more detailed modeling. For example, the delay function, that was applied to delay activation spreading from one level to the next could be used to model word frequency effects by varying the delay of connections between lemma and corresponding phonemes per word. In addition, the gating devices, used to model shifts of attention and variations in discourse processing, provide room for a more flexible performance in a connectionist network that has a trained and fixed pattern of connection weights. The use of gate functions, as implemented for the discourse module, presents the opportunity of making a model able to perform multiple tasks within the same architecture, such as picture naming, lexical decision, reading, categorization. Such flexibility is needed if a model should ultimately account for empirical findings of different experimental paradigms. The present version is only a beginning.

CHAPTER 5

5.1 Summary of the main findings

The main question of my dissertation research was how lexical access takes place during the generation of reduced forms, such as ellipses and pronouns. Because it is assumed that lexical access consists of at least two stages, semantic/syntactic encoding and phonological encoding, the present thesis investigated whether the meaning and the phonological form of a word are available if an anaphoric construction is used that refers back to the referent. According to a strict two stage model (Levelt et al., 1991a), the meaning of the referent should be active because the anaphoric utterance refers to the meaning of the referent. However, the phonological form of the referent should not be activated while generating anaphora, because the referent's phonological form of the referent is not needed: The referent is not overtly spoken. From the perspective of theories that posit cascading spreading of activation (Dell, 1986; Dell & O'Seaghdha, 1992) a different assumption might be made. Here, the meaning of the referent should be active, which activates the lemma. But in contrast to the two-stage assumption, the lemma of the referent automatically spreads activation towards the phonological form, even if the referent is not articulated.

The dissertation consists of three parts. First, I investigated in a picture-word interference paradigm the semantic and phonological activation of a referent during the production of ellipses. Second, I moved from ellipses to pronouns and did research on the phonological activation of the referent during the production of pronouns. The paradigm I used here was a dual task paradigm. Participants had to describe pictures and, in a secondary task, made lexical decisions on acoustic probes. Third, I used a PDP approach to model activation processes in the dual task situation in order to explain the phonological effect I found in the pronoun studies.

5.1.1 Lexical access during the generation of ellipsis

The aim of the experiments in Chapter 2 was to investigate lexical access during the generation of ellipses. Is a word semantically and phonologically activated if it is elided in the second of two adjacent sentences? In a series of experiments, which were carried out in Dutch, I used partial correction as the utterance format, such as *Kiss Toon...Paul!*. There are reasons to suppose that the target verb ‘kiss’ is semantically active in the gap between 'Toon' and 'Paul'. The complete structure of the utterance is *Kiss Toon...no...Kiss Paul!* However, because the verb is not articulated overtly in the second position of the elliptic utterance, it is possible that no phonological activation might be found at the gap-position. This hypothesis was tested with a picture-word interference paradigm. The participants had to name a sequence of two pictures. The description of the first picture was, for example, *Kiss Toon.* On the second picture, the patient Toon changed to Paul. The participants were asked to interpret this change as a kind of correction of a visual scene and to name it aloud, leading to the description of the two-picture sequence as *Kiss Toon...Paul!* Time-locked to the onset of the second picture, an acoustic distractor stimulus was presented. This stimulus was either semantically or phonologically related, or unrelated to the referent verb. For this task, a strict two-stage model predicts semantic inhibition and no phonological effect relative to the unrelated distractor. The results of the first ellipses experiment showed semantic inhibition, but it revealed no phonological effect at all for the elided verb during the generation of ellipsis. In a control condition, where the verb was not elided, as in *Feed Toon....Kiss Paul,* semantic inhibition and phonological facilitation were found for the target verb ‘kiss’. The results of the control condition indicated that the observed phonological null effect during ellipsis generation could not be due to methodological aspects of the experiment, but should be interpreted in terms of the linguistic utterance format used. The semantic inhibition during the generation of ellipsis was interpreted as stemming from activation of the lemma of the elided verb. The phonological null effect was seen as evidence that the lemma does not send information further to its corresponding phonemes at the phonological level of speech production.

In a follow-up experiment involving a change of speakers, the results of the first study were basically replicated. Taken together, the findings of the experiments supported Klein’s (1993) assumption of *p-reduction* and *speaker-independence* during the generation of ellipses. In addition, these results provided support for the two-stage theory: An active lemma that is not selected for overt naming spreads no activation towards its corresponding phonemes.

5.1.2 Lexical access during the generation of pronouns

In Chapter 3, a second form of reduction is discussed with regard to processes of lexical access: The generation of pronouns. How are pronouns accessed? One idea is as follows (see figures 3.2, and 5.1): The picture name activates its concept. The concept in turn activates its lemma and via the lemma the corresponding syntactic information, such as grammatical gender. If the discourse carries the information that the current entity is in focus, the discourse mechanism allows access from syntactic gender nodes to the pronoun nodes by opening the connections from the gender nodes to the pronoun nodes. The grammatical gender information is needed in order to generate the appropriate pronoun (the masculine, feminine, or neuter pronoun in German). The pronoun becomes selected and activates its phonemes.

The experiments described in Chapter 3 investigated the lemma's activation by looking at whether its phonemes might become available during pronoun generation. The question whether the speaker activates the phonological form of a noun when it is pronominally referred to was investigated for utterance formats, such as *The flower is red...It turns blue*. Here, the pronoun 'it' refers to the noun 'flower'. A theory that assumes cascadant spreading of activation (Dell, 1986; Dell & O'Seaghdha, 1991, 192) would predict phonological activation, if the lemma of the referent noun becomes activated. In contrast, according to a two-stage theory of lexical access (Levelt et al., 1991a; Roelofs, 1992a, b) one may expect no phonological activation of the referent noun at the pronoun position, because the noun is not selected for articulation. This hypothesis was tested in experiments in German, where participants made lexical decisions during picture descriptions. They described a sequence of two pictures. The first picture description, such as *The flower is red*, included reference to an object. In the second picture this object re-appeared in a different color in some trials. In these cases the participants used a pronoun, such as *It turns blue*. In 50% of all trials an acoustic lexical decision probe was presented, time-locked to the second picture onset. The probe could either be a word or a pseudoword. In these trials the task was to postpone the picture description, and to make a lexical decision on the probe first. The word probe was phonologically related or unrelated to the noun of the second picture. The results showed phonological inhibition, suggesting that during pronoun generation the word form of the noun is active.

To find out whether this phonological activation was due to residual activation of the first picture description, a follow-up experiment was carried out. Again, participants had to describe a sequence of two pictures, such as *The flower is red....The sun is blue*. As in the preceeding experiment, probe-presentation was time-locked to the second picture onset. But now the probe was phonologically related or unrelated to the noun of the first picture. No phonological effect was found in the pronoun condition. This result suggested that the phonological form of the first picture's referent is not available anymore during the presentation of the probe shortly after the onset of the second picture. The phonological effect found during pronoun generation, therefore, was interpreted as not being due to residual activation of the overt noun naming of the first picture, but as a consequence of re-accessing the lemma.

The observed phonological inhibition supported a cascading view of activation spreading during lexical access (see Figure 3.3), which locates the phonological effect at the phonological level, and assumes feedback from delayed phonological encoding to the lemma level.

The phonological inhibition during pronoun generation contradicts the two-stage model as it was stated by Levelt et al. (1991a). This is the case, because in the model of Levelt et al. the phonological effect found during single noun naming was explained due to competition of overlapping phonological stages. According to the model, during pronoun generation this competition should not take place because the referent noun is not selected for naming. The prediction according to the 1991 theory, therefore, is not compatible with the findings. But, as discussed in the conclusion of Chapter 3, according to more recent ideas the two-stage theory might explain the phonological effect in terms of cohort-processes at the lemma level.

The experiments, therefore, could not clearly distinguish between the two different theories of language production (cascading vs. two-stage lexical access). Nevertheless the two theories agree upon the idea that the observed phonological effect involves activation of the lemma of the referent noun. The obtained effect, therefore, empirically supported the lexical access view of pronoun generation (as depicted in figures 3.2, and 5.1).

5.1.3 Phonological inhibition on lexical decision during noun and pronoun generation: A PDP approach

In Chapter 4, a PDP approach is described that was used to simulate activation processes in a dual task situation. The computational model was used as a tool to locate the phonological interference effect at either the phonological level (as proposed by a cascaded processing view), or at the lemma level (as proposed by a two-stage view). The model consists of the levels of speech processing assumed by current theories of language production. The computational model was first trained to carry out lexical decisions on acoustic word probes, and then to name pictures. Later on, its performance on the dual task was investigated.

In a first simulation, the model should test the assumption of the cascaded processing view. This view explains phonological inhibition that was observed in the pronoun experiments in terms of segmental competition at the phonological level. This in turn means, that phonological inhibition can only be observed if the noun lemma of the corresponding pronoun becomes phonologically active and interferes with segmental encoding of the related probe word. In addition, in order to see this segmental competition in a lexical decision task, the delayed phonological encoding of a related probe activates its lemma later, leading to a delayed lexical decision. In order to cover this feedback from phonological level to lemma level the computational model included feedback connections between the phonological level and the lemma level.

In the first simulation the behavioral data of the Levelt et al. (1991a) study of single noun naming were replicated, except for the identical condition. Here, the computational model showed not inhibition but facilitation (as observed in my own replication of Levelt et al., Chapter 3). Most important, the simulation replicated the phonological inhibition during lexical decision observed by Levelt et al. (and in the replication study). A descriptive analysis of weight pattern and time course of activation explained the observed effect by inhibitory feedback of phonological mismatching segments towards the lemma level. The first simulation, therefore, supports the cascaded processing view.

In a second simulation, no feedback between the phonological level of naming and the lemma level was implemented in the architecture of the computational model. This simulation should test the cohort account in order to explain phonological inhibition at the lemma level, as assumed by the two-stage model. According to the cohort account, the acoustic probe initially activates a cohort of lemmas that share the same onset. The more acoustic information reaches the lemma level the more the cohort becomes reduced towards a single element. A phonologically related picture name is also a member of this cohort and, because it is highly activated during the naming process, it is a strong competitor to the phonologically related lemma. This competition leads to a delayed cohort reduction process, which results in elongated lexical decision latencies.

The results of the second simulation showed the phonological inhibition at short SOA observed by Levelt et al., and in the replication study. However, the simulation showed a decrease of phonological inhibition at long SOA, whereas the empirical data of Levelt et al. indicated no such decrease of inhibition across SOAs.

The comparison of the two simulation studies led to the decision to advance the feedback variant, and to simulate pronoun generation with this version. This is described in Simulation 3 of Chapter 4. The network architecture was extended by a discourse module, and by a pronoun node. It should simulate the proposed lexical access view of pronoun generation, as discussed in Chapter 3. In addition to learning the lexical decision and noun naming task, the computational model was trained to generate nouns or pronouns, depending on the discourse information. This discourse-dependent performance was realized by implementing a gate function. The gate allows pronoun access, if the discourse module signals that the current entity is in focus. The gate is closed if the discourse module is not active, which leads to an overt noun naming.

The computational model successfully simulated the empirically observed phonological inhibition at short SOA for both noun and pronoun generation. The performance of the computational model with regard to the time course of phonological activation during pronoun generation can be explained by an interplay of three components: An early conceptual inhibition, inhibitory feedback of phonological mismatch, and different pathways of phonological activation of pronouns and nouns. The performance of the PDP model, therefore, supports the lexical access view of pronoun generation.

In addition, the computational model yielded predictions for phonological inhibition at long SOAs during pronoun generation. These predictions can be tested empirically. The empirical data in turn will then support or falsify the model.

Furthermore, the self-organizing nature of the PDP model gave a solution for a co-activation problem: If phonemes of the noun and the pronoun are activated at the same time, as may be the case in a cascading process, how does it happen that only the pronoun becomes articulated? The model gives one possible explanation: The generation of pronouns is acquired later than that of nouns. In the model, this later training leads to strong inhibitory connections between the pronoun node at the lemma level and mismatching (noun) segments at the phonological level. A slightly activated pronoun node, therefore, can easily suppress the already available phonological information of the noun.

5.2 Differences in effects between ellipsis and pronouns

This thesis addressed lexical access of two different kinds of reduced forms in speech production, ellipsis and pronouns. The present experiments tested whether the lemma and the phonological information of a referent becomes activated when a speaker is generating a reduced form of this referent.

The *activation of the lemma* was predicted by current theories of speech production. These theories assume that the lemma carries the syntactic information of the referent. This syntactic information is available during unreduced speech in terms of case marking or gender information. However, it is also available during the generation of reduced forms. An elided verb, for example, still yields syntactic information about case, as in the German example *Ich begegne*$_{\mathit{dative}}$ *dem*$_{\mathit{dative}}$ *Mann und .. der*$_{\mathit{dative}}$ *Frau* (I meet the man and the woman). A pronoun still carries the information of the syntactic gender of the referent noun, as in *Er sah das Buch*$_{\mathit{neuter}}$ *und kaufte es*$_{\mathit{neuter}}$ (He saw the book and bought it). According to theories of speech production, this syntactic information can only become available via the lemma. The results showed that the lemma becomes activated during the generation of ellipsis and pronouns. The present experiments, therefore, support the theoretical assumption that lemma access is equally involved during the generation of pronouns and ellipsis.

The *activation of the phonological form* during the generation of reductions was the central issue in this thesis. It was central because different theories predicted different outcomes. A cascaded processing view assumes that an activated lemma automatically spreads activation towards the form, and therefore predicts a phonological effect for the referent during the overt generation of the referent as well as during the generation of its reduced form. In contrast, a two-stage view does not predict a phonological effect during the generation of reduced forms, because the referent word lemma does not become selected for naming (the word is elided or pronominalized). Therefore, the lemma does not activate its form. The basic idea at the beginning of the dissertation project was that the reduced forms could be used as a tool to distinguish between the two theoretical assumptions.

However, the results that were obtained for ellipsis and pronoun generation differ. Whereas during the production of ellipsis the phonological form of the referent is not active, phonological effects were observed during pronoun production. The ellipsis experiments, therefore, support the two-stage theory, the pronoun studies support the cascaded processing view. *How can this difference of phonological activation be reconciled?*

5.2.1 Methodological differences

There are *methodological reasons* that might explain the difference of a phonological null effect for ellipsis and the phonological inhibition effect for pronouns. The two experimental series differ in many respects. I will only discuss the most obvious ones, the language and the paradigm.

First, the experiments used different target languages: Dutch in the ellipsis experiments, and German in the pronoun study. It might be the case that cross-linguistic differences caused the difference in experimental outcome. For example, Dutch native speakers might not activate phonological information during reduced speech at all, whereas German native speakers do. If the observed difference is a language specific phenomenon, Dutch participants should produce a null effect during pronoun generation, and German participants should reveal a phonological effect in elliptic utterances. However, at the moment there is no theoretical reason why this should be the case.

Second, the experiments used different paradigms. The ellipsis generation was investigated in a picture-word interference paradigm. Participants had to describe pictures, and heard distractor words that they were asked to ignore. They only had to carry out one task. In contrast, the generation of pronouns were studied in a dual task paradigm. Participants described pictures, but when they heard an acoustic stimulus, they postponed naming, and carried out an active lexical decision on the acoustic probe. Ignoring a distractor word on the one hand, or making an active decision on it on the other might lead to different processing of the acoustic probe. For example, the to be ignored stimulus in the ellipsis experiments might activate phonological information to a lesser extent than the lexical decision probe in the pronoun study, leading to the observed null effect for ellipsis in contrast to the phonological inhibition for pronouns. However, each experiment had its built-in control condition, i.e., the overt generation of the target word. These control conditions revealed phonological effects in both experimental series, i.e., in different paradigms. I therefore conclude, that the processing of stimuli in the two paradigms does not differ to such an extent that it could explain the different phonological effects observed in the reduced utterance format (null effect for ellipsis, inhibition for pronouns). Of course, in order to be on the safe side, one should switch the paradigms.

Although one should not compare empirical data that result from different paradigms in different languages directly, the commonalities are substantial. The common aspects do justify a preliminary comparison of the results. For example, they both used acoustic stimuli that should interfere with the ongoing speech planning process. The probe conditions are comparable because the critical experimental variable ‘phonological relatedness’ was defined in the same way. In addition, in both experimental series, the probe words were presented at a comparable SOA during the presentation of a sequence of two pictures.

5.2.2 A linguistic theory about differences of ellipsis and pronouns

If there exist no obvious methodological reasons for the observed differences of phonological effects between ellipsis and pronouns, they may have been caused by *linguistic differences in the two kinds of reduction.* To my knowledge so far, there are no studies that investigated the generation of reductions by using picture descriptions. Therefore, I would like to discuss a linguistic approach that might explain the difference by postulating different types of reductions.

Hankamer and Sag (1976) proposed that anaphoric expressions can be divided into two main classes. They labeled the classes *deep* and *surface* anaphors. An example of deep anaphors are *pronouns*, one example of surface anaphors are *ellipses*. The authors claimed that these classes differ in terms of the level of representation that must be accessed to determine their referent in comprehension, but the approach can be applied to production as well. Deep

anaphors, such as pronouns, directly access a level of representation in a discourse model or mental model (Johnson-Laird, 1983), but can also access the syntactic level. In contrast, surface anaphors *must* first access a purely linguistic (syntactic) level of representation.
That pronouns can occur with non-linguistic control, Hankamer and Sag (1976, p. 407) exemplify by:

(1) Hankamer [observing Sag successfully ripping a phone book in half]
I don't believe *it*
Sag [same circumstance]·
It's not easy

The generation of the pronoun *it* is controlled only by aspects of the non-linguistic situation. It has no linguistic antecedent in this case. Following the account of pronoun generation of Chapter 3, the action concept RIPPING is in the focus of the two men's discourse situation. It is available in Sag's discourse record, because he just performed the action, and in Hankamer's because he just saw it. The feature *+(in focus)* leads to an appropriate reduction of the action concept.

But, as Hankamer and Sag (1976) pointed out, pronouns may also be controlled linguistically, as could be the case in our experimental example:

(2) The flower is red *It* turns blue.

Here, *it* has a clear linguistic antecedent, *flower*. The examples should demonstrate that the generation of pronouns may be controlled in two different ways. First, pronouns may be accessed by using discourse information only, as in (1): The *flower* may get an +(*in focus*) feature in the discourse record, leading to the generation of the reduced form. Second, pronouns may be accessed by using information of the preceding linguistic format, as in (2).

In contrast, ellipsis, for example in the case of 'gapping',[1] always requires syntactic control, as Hankamer and Sag (1976, p. 410) showed in the following example:

(3) [Hankamer produces an orange, proceeds to peel it, and just as Sag

---

[1] According to Hankamer and Sag (1976) "gapping" is an ellipsis rule that applies in coordinate structures to delete all but two major constituents from the right conjunct under identity with corresponding parts of the left conjunct" (p 410) See also the example in Chapter 2 about syntactic reduction of identical utterance elements, and Klein (1993) for a discussion.

produces an apple, says:]
#[2] And Ivan, an apple.

This discourse seems to be bizarre, whereas the following one seems to be fine:

(4) Hankamer: Ivan is now going to peel an apple.
Sag: And Jorge, an orange.

Applied to our ellipsis case, Sag could also generate a correction, such as:

No, Jorge, an orange.

The authors concluded that the strangeness of (3) is due to the attempt to "gap under pragmatic control" (p. 411), which means to delete what is not available in the discourse context. The bizarreness of (3) in contrast to the normality of (1) clearly shows different generation rules for pronouns and ellipses.

These different processing rules are outlined in Sag and Hankamer (1984). The article is a revision of the former theory. The authors maintained the dichotomy of anaphoric processes proposed in 1976, but renamed the two classes because they also modified their theory. Former deep anaphors became so-called *model-interpretive anaphors*, and former surface anaphors became *ellipsis*.

Model-interpretive anaphors, such as pronouns, refer to an element of the *discourse model*. This element represents an entity from the physical environment (a concrete object, event, or state of affairs). This element could either be available by non-linguistic discourse context, as in example (1), or it could become linguistically available in terms of a preceding utterance, as in (2). But in contrast to their previous assumption (1976), Sag and Hankamer (1984) postulated that there is no evidence for such a dichotomy. Therefore, they argued that the relevant module that helps accessing the pronoun in example (1) as well as in (2) is the discourse model.

In contrast to pronouns, ellipsis refers to an antecedent that was linguistically expressed in the preceding utterance. Sag and Hankamer assume that the preceding utterance creates a *propositional construct*, in terms of a 'logical form' (after Williams, 1977).

The distinction between reference to the discourse model and propositions is based on Johnson-Laird's (1983, 1989) suggestion about discourse processing. He assumed that a new piece of discourse information is encoded first in terms of propositional representations. Thus, at any given point in a discourse at least

---

[2] The cross-hatch (#) indicates a sentence as incompatible with the context.

two mechanisms are active in the mind of the speaker and the listener: On the one hand, the representation of the proposition, and on the other hand, the discourse model. Sag and Hankamer assumed that as the discourse proceeds, the content of the propositional representations is integrated into the discourse model (for a review of the construction of discourse models see Oakhill, Garnham, and Vonk, 1989).

Furthermore, Sag and Hankamer assumed that the propositional information is available only momentarily in short-term memory, after which it decays in order to make room for new propositional frames. Propositional information, therefore, differs from meaning representations in the discourse model, which are supposed to available for longer. Empirical support for this assumption came from sentence recall experiments. Participants show a recency effect in verbatim recall tasks: Immediate verbatim recall is highly correct, whereas delayed verbatim recall is not. In contrast, gross meaning recall is not affected by the temporal distance between sentence presentation and recall (Sachs, 1967; for a review see Garnham & Oakhill, 1987; Tanenhaus & Carlson, 1990).

5.2.3 Empirical investigations of the linguistic theory

Sag and Hankamer's classification of anaphors into model-interpretive anaphors (for example, pronouns) and ellipsis has been challenged in the comprehension literature.

Murphy (1985) found evidence against the dichotomy of anaphors. He investigated the effect of antecedent length on reading times for model-interpretive anaphors and ellipses. Participants read paragraphs of text including sentence pairs like:

Johanna swept the wooden floor [behind the chairs free of toys].
a Later, her sister *did* too. (elliptic construction)
b. Later, her sister *did it* too. (pronoun construction)

The first sentence could either include a short or a long antecedent (in brackets). The second sentence could either involve ellipsis or pronouns. Murphy argued, following Sag and Hankamer, that antecedent length should matter for ellipsis due to different underlying copy processes during propositional encoding: Copying a long phrase should take longer. In contrast, antecedent length should play no role during pronoun comprehension, because access to discourse information is less affected by intervening material. Murphy observed no differences in reading times for ellipsis and pronoun constructions. Reading times were longer in the long antecedent condition for both types of anaphors, suggesting that participants copied both forms of anaphors in the same way.

However, the longer reading times in the long antecedent condition in the pronoun case could also be due to more complex discourse processing, or a general increase of cognitive load.

Tanenhaus and Carlson (1990) supported the distinction between pronouns and ellipsis by means of a "make sense" judgment task. Participants were asked to read pairs of sentences, such as

1a. Someone had to take out the garbage. (syntactic parallel)
1b. The garbage had to be taken out. (syntactic non-parallel)
2a. But Rabea refused to do it. (pronoun construction)
2b. But Rabea refused to. (elliptic construction)

Elliptic constructions were judged to make sense more often when the antecedent was introduced by a phrase that was syntactically parallel (active sentence) to the anaphor, than by one that was non-parallel (passive sentence). This was predicted from Sag and Hankamer's postulation of a copy process for ellipsis. A copy of a parallel structure is more easily recovered than a copy of non-parallel structure. The remapping might increase the cognitive load for interpretation because in passive structures the conceptual elements may be ranked in a different order due to a different status of accessibility (Bock & Warren, 1985; Bock, 1986, 1987). In contrast, the pronoun constructions were judged to make sense equally often in both syntactic forms of the antecedent. This was also predicted by Sag and Hankamer: A "make sense" judgment of pronouns should not be sensitive to aspects of linguistic/syntactic formats because it is the non-linguistic discourse information that matters. However, Tanenhaus and Carlson observed a difference in judgment latencies. Reaction times were prolonged for both the ellipsis and the pronoun constructions in the syntactic non-parallel condition in contrast to the parallel condition. This elongation might be due to a necessary re-ranking of conceptual information if the antecedent was given in a passive format instead of an active format where no re-ranking of conceptual information is needed.

More recently, Garnham et al. (1995) challenged the hypothesis of pronouns accessing discourse information only. The authors investigated whether a mere grammatical cue could affect the interpretation of pronouns in reading experiments. In many languages other than English, nouns have grammatical gender with no semantic reflex. The gender for *table*, for example, is feminine in French (la table) and Spanish (la mesa), but masculine in German (der Tisch). But tables themselves are neither male nor female. Garnham et al. argued that if pronoun resolution was speeded by grammatical gender information, such an effect could not readily be explained on the assumption that only a representation of discourse was important in resolving the pronoun. The authors asked French and Spanish speaking participants to read pairs of sentences that were about people (biological gender) or things (grammatical gender). The second sentence included a pronoun that referred back to one of the people or one of the things, for example:

> Richard/Alice arrested Paul because...
> he/she found him stealing a car.
> The truck$_{\text{masculine}}$/breakdown truck$_{\text{feminine}}$ towed the bus$_{\text{masculine}}$ because...
> it$_{\text{masculine}}$ was stuck in the snow.

In one version of the first sentence both people or things were of the same gender, and in the other they were not. Thus, in the version in which the people or things were of different gender, the pronoun could be resolved from its gender alone. The main concern of the study was whether there would be an effect of gender cueing in the sentences about things. This cueing effect would indicate that pronouns could be accessed by means of grammatical information. Garnham et al. observed this gender cueing effect. Reading times were speeded up in the cued condition in sentences about things. The authors interpreted the results as evidence for a 'superficial' encoding, i.e., a linguistic/syntactic encoding of pronouns. They argued that the finding contradicts the discourse/meaning-driven interpretation of pronouns, proposed by Sag and Hankamer. However, the result was expected by the proposed lexical access hypothesis (see Figure 5.1). In agreement with Sag and Hankamer, pronoun access is discourse-driven. But in addition, the observed phonological effect in the pronoun experiments was interpreted in terms of having the lemma of the referent noun active. This in turn means that the access to grammatical gender is available. A gender cue, therefore, does not contradict the Sag and Hankamer hypothesis. A gender cue may help to access the pronoun in comprehension (as observed by Garnham et al., 1995) as well as in production, which has to be shown in future research.

5.2.4 Fitting the linguistic theory into a psycholinguistic model

The proposed distinction between accessing discourse information for the generation of pronouns, and accessing propositional information for the generation of ellipsis, can now be applied to the speech production system.

Lexical access during ellipsis generation

With regard to ellipsis, as has been addressed in Chapter 2, Klein (1993) postulated that elements can be elided if they were marked as the *topic* in the preceding utterance and remain the topic in the present one. According to Levelt (1989, p. 98) an element is marked as sentence topic in the propositional structure. Therefore, this information may be taken over in the copied propositional frame for ellipses. The propositional frame of an utterance, such as *Kiss Pien*, marks the verb as topic. In the partial correction *Kiss Pien..(No)...Paul* the topicalized verb can, therefore, be deleted at the second occurrence. This deletion was assumed to be a phonological reduction (Klein, 1993). This assumption was supported by the empirical data during the generation of ellipses. No phonological effect was observed. The observed semantic inhibition effect was interpreted as having the lemma active, which in turn means that the syntactic information of the elided verb can get accessed. A lemma however can only be activated by its corresponding concept. The observed semantic effect, therefore, is in line with Sag and Hankamer's theory about propositional access of ellipses. If propositional information is relevant for generating ellipses then the semantic information should be available, because propositions are part of the conceptual encoding process (Levelt, 1989). However, activated propositional frames might be more short-lived than semantic information of concepts, which indicates that they are processed differently in comparison to concepts. They could, for example, be a linguistic representation of the message, and not a preverbal one. In any case, propositions should be located somewhere between conceptual encoding and lemma access.

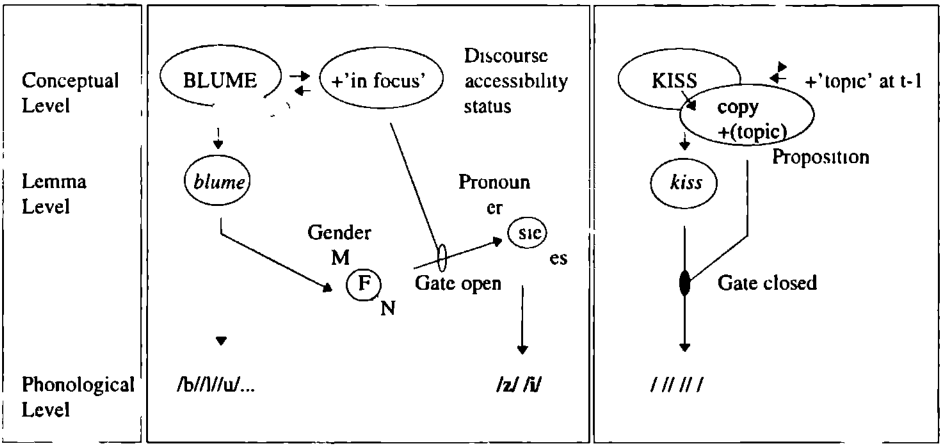


Figure 5.1 A lexical access view of the generation of pronouns and ellipsis. *Left panel:* The discourse-driven access of the pronoun *sie* in *Die Blume ist rot. Sie wird blau.* (The flower is red. It turns blue.) *Right panel:* The proposition-driven phonological deletion of the verb *kiss* in *Kiss Pien.(No)...Paul.*

The propositions became an extra element in Figure 5.1, which shows the proposed access of pronouns (left panel) and ellipsis (right panel). Propositional encoding is depicted in gray as a submodule of conceptual encoding. The content of the proposition module is kept empty in the pronoun case, because it is not relevant to Sag and Hankamer's theory of pronominalization.

As can be seen in the right panel of Figure 5.1, in the ellipsis case the proposition for KISS has the feature *copy +(topic).* "Copy" refers to the taking over of the propositional format from the preceding sentence, according to Sag and Hankamer. "+(topic)" refers to the marking of the element as topic in the current proposition, following Klein (1993). If the element was marked as topic in the preceding utterance, depicted in the figure as *+(topic at t-1)* in the discourse model, it will be assigned as topic in the copied proposition as well. Following Klein, *kiss* can then be phonologically deleted. Using a comparable gating mechanism as for pronouns, the phonological deletion could be realized in terms of a proposition-driven closing of the connections from the lemma level towards the phonological level.

Lexical access during pronoun generation

As has been discussed in Chapter 3, the discourse information needed for pronoun generation may be the *+(in focus)* feature of accessibility. This idea is in line with Sag and Hankamer's assumption that discourse information is relevant for pronoun generation. As postulated in Chapter 3, a procedural rule could open the connections between lemma and gender information towards pronoun access (see Figure 5.1, left panel). The figure is a copy of Figure 3.2, and shows the access of the pronoun that corresponds to the concept of a flower (Blume). In addition, the figure depicts temporary phonological co-activation of the noun lemma, as predicted by cascaded processing theories. This co-activation may become suppressed the more the pronoun node becomes activated. The suppressing may be due to inhibitory connections between the pronoun node and mismatching phoneme units at the phonological level, as discovered during Simulation 3 in Chapter 4.

5.2.5 Conclusions

In conclusion, (psycho)linguistic theories and empirical evidence support the idea that pronouns and ellipsis differ with regard to lexical access. Both reduced forms involve the activation of their referent's lemma. But only during the generation of pronouns may the phonological form of the referent become activated. The theoretical assumption of phonological deletion during ellipsis generation (depicted in Figure 5.1, right panel, as a gate that closes connections between lemma level and phonological form), in turn, means that the generation of ellipsis is not an appropriate tool to investigate the distinction between cascading and two-stage lexical access. This is the case because phonological deletion is located exactly at the position where the two theories make different assumptions, namely with respect to the time course of lemma access and phonological activation. In contrast, pronouns may be a useful tool to tap into the time course of lexical access. In this thesis, a cascaded processing view is advanced, and supported by the simulation results of the computational network. Of course, further empirical work is needed to validate this interpretation.

REFERENCES

Allport, A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives of perception and action* (pp. 395-419). Hillsdale, NJ: Erlbaum.

Allport, A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In *Attention and Performance 14: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 183-218). Cambridge, MA: MIT Press.

Appelt, E. A. (1985). *Planning English sentences.* Cambridge University Press.

Baddeley, A. (1986). *Working memory.* Oxford: Clarendon Press.

Balota, D. A., & Abrams, R. A. (1995) Mental chronometry. Beyond onset latencies in the lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21* (5), 1289-1302.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3-73). Cambridge, UK: Cambridge University Press.

Bauman, M. (1995). *The production of syllables in connected speech.* Doctoral dissertation. Nijmegen University.

Beattie, G. (1983). *Talk: An analysis of speech and non-verbal behavior in conversation.* Milton Keyes: Open University Press.

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

Bock, J. K. (1986). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 575-586.

Bock, J. K. (1987). An effect of accessibility of word forms on sentence structures. *Journal of Memory and Language, 26,* 119-137.

Bock, J. K. (1995). Sentence production: From mind to mouth. In J. L. Miller, & P. D. Eimas (Eds.), *Handbook of perception and cognition:* Vol. 11. *Speech, language, and communication* (pp. 181-216). San Diego: Academic Press.

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition, 21,* 47-67.

Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language, 31,* 99-127.

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945-984). San Diego: Academic Press.

Bolinger, D. L. (1961). Verbal evocation. *Lingua, 10,* 113-127.

Booij, G. (1995). *The phonology of Dutch.* Oxford: Clarendon Press.

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon Press.

Brown, P., & Levinson, S. C. (1993). Linguistic and nonlinguistic coding of spatial arrays: Explorations in Mayan cognition. *Working paper No. 24,* Cognitive Anthropology Research Group, Max Planck Institute for Psycholinguistics.

Brown, R. & McNeill, D. (1966). The "tip of the tongue phenomenon". *Journal of Verbal Learning and Verbal Behaviour, 5,* 325-337.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109,* 204-223.

Bühler, K. (1934). *Sprachtheorie.* Jena: Gustav Fischer.

Burnage, G. (1990). *CELEX: A guide for users.* Nijmegen, The Netherlands: Centre for Lexical Information, University of Nijmegen.

Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language production:* Vol. 1. *Speech and talk* (pp. 155-176). London: Academic Press.

Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 108-135). Cambridge, MA: MIT Press.

Caplan, D. (1994). Language and the brain. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 1023-1053). San Diego: Acadamic Press.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25-56). New York: Academic Press.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge: MIT.

Clark, H. H. (1994). Discourse in production. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 985-1021). San Diego: Acadamic Press.

Cloitre, M., & Bever, T. G. (1988). Linguistic anaphors, levels of representation, and discourse. *Language and Cognitive Processes, 3* (4), 293-322.

Cohen, J. (1988). *Power analysis for the behavioral sciences.* Hillsdale: Erlbaum.

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

Colombo, L. (1986). Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 226-234.

Corbett, G. G. (1991). *Gender.* Cambridge: Cambridge University Press.

Cutler, A. (1995). Spoken word recognition and production. In J. L. Miller, & P. D. Eimas (Eds.), *Handbook of perception and cognition:* Vol. 11: *Speech, language, and communication* (pp. 97-136). San Diego: Academic Press.

Deese, J. (1984). *Thought into speech: The psychology of a language.* Englewood Cliffs, NJ: Prentice-Hall.

De Groot, A. M. B. (1985). Word-context effects in word naming and lexical decision. *Quarterly Journal of Experimental Psychology, 37A,* 281-297.

De Smedt, K. (1996). Computational models of incremental grammatical encoding. In T. Dijkstra & K. de Smedt (Eds.), *Computational Psycholinguistics* (pp. 279-307). London: Taylor & Francis.

De Smedt, K., & Kempen, G. (1987). Incremental sentence production, self-correction, and coordination. In G. Kempen (Ed.), *Natural language generation: Recent advances in artificial intelligence, psychology and linguistics* (pp. 365-376). Dordrecht: Kluwer.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283-321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language, 27,* 124-142.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5,* 313-349.

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review, 98,* 604-614.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42,* 287-314.

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behaviour,* 20, 611-629.

Deutsch, W., & Pechmann, T. (1978). Ihr, dir, or mir? On the acquisition of pronouns in German children. *Cognition, 6,* 155-168.

Deutsch, W., Wagner, A., & Masche, G. (1994). Vom Nomen zum Pronomen: Eine experimentelle Untersuchung zur Entwicklung der Personreferenz bei deutschen Kindern während des zweiten bis vierten Lebensjahres. *Sprache und Kognition, 13* (1), 41-51.

Eberhard, K. M., Bock, J. K., & Griffin, Z. (1994). Interference from phonological primes in spoken word production. Paper presented at the 35th meeting of the Psychonomic Society, St. Louis, MO, USA.

Ericsson, K. A. & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102* (2), 211-245.

Eriksen, B. A., & Eriksen, C.W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16,* 1, 143-149.

Fahlmann, S. E. (1988). Faster learning variations on back-propagation: An empirical study. In D. S. Touretzky, G. E. Hinton, & T. J. Sejnowski (Eds.), *Proceedings of the 1988 connectionist models summer school.* San Mateo: Morgan Kaufman.

Fay, D., & Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry, 8,* 505-520.

Fodor, J. D. (1989). Empty categories in sentence processing. *Language and Cognitive Processes, 4* (3/4), 155-209.

Fowler, C. A., & Housum, J. (1987). Talkers' signalling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language, 26,* 489-504.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading* (pp. 559-586). Hillsdale, NJ: Erlbaum.

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

Fromkin, V. A. (Ed.) (1973). *Speech errors as linguistic evidence.* The Hague: Mouton.

Garnham, A., & Oakhill, J. (1987). Interpreting elliptical verb phrases. *Quarterly Journal of Experimental Psychology, 39A,* 611-627.

Garnham, A., Oakhill, J., Ehrlich, M.-F., & Carreiras, M. (1995). Representations and processes in the interpretation of pronouns: New evidence from Spanish and French. *Journal of Memory and Language, 34,* 41-62.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory,* Vol. 9 (pp. 133-177). New York: Academic Press.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production,* Vol. 1 (pp. 177-220). London: Academic Press.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19-76). London: Academic Press.

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey,* Vol. 3: *Language: Biological and psychological aspects* (pp. 69-96). Cambridge, MA: Cambridge University Press.

Garrett, M. (1992). Disorders of lexical selection. *Cognition, 42,* 143-180.

Glaser, W. R. (1992). Picture naming. *Cognition, 42,* 61-105.

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 640-654.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General, 118,* 13-42.

Goebel, R. (1995). Neurolator - Ein Programm zur Simulation neuronaler Netzwerke. In T. Plesser, P. Wittenburg (Hrsg.), *Forschung und wissenschaftliches Rechnen. Beiträge zum Heinz-Billing-Preis 1994.* GWDG-Bericht Nr. 40. Göttingen: Gesellschaft für wissenschaftliche Datenverarbeitung mbH.

Goebel, R. (1996). *Visuelle Aufmerksamkeit, perzeptive Organisation und invariante Objekterkennung: Eine Integration neurobiologischer und psychologischer Befunde in einem neuronalen Netzwerk-Modell.* Dissertation. Technische Universität Carolo-Wilhelmina, Braunschweig, Germany.

Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18* (6), 1211-1238.

Goldman, N. (1975). Conceptual generation. In R. C. Schank (Ed.), *Conceptual information processing* (pp. 289-374). Amsterdam: North-Holland.

Gordon, P. C., & Scearce, K. A. (1995). Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory and Cognition, 23* (3), 313-323.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17,* 311-347.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech acts,* Vol. 3 (pp. 41-58). New York: Academic Press.

Grosz, B. J., & Sidner, C. L. (1985). Discourse structure and the proper treatment of interruptions. *Proceedings of the International Joint Conference on Artificial Intelligence.* Los Altos, CA: Morgan Kaufman.

Grosz, B. J., Joshi, A., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. *Proceedings of the 21st Annual Meeting of the Association of Computational Linguistics* (pp. 41-50). Cambridge, MA: ACL.

Hankamer, J., & Sag, I. A. (1976). Deep and surface anaphora. *Linguistic Inquiry, 7,* 391-426.

Harley, T. A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes, 8,* 291-309.

Heeschen, C., & Kolk, H. (1988). Agrammatism and paragrammatism. *Aphasiology, 2,* 299-302.

Henderson, A., Goldman-Eisler, F., & Skarbek, A. (1966). Sequential temporal patterns on spontaneous speech. *Language and Speech, 9,* 207-216.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review, 98,* 1, 74-95.

Hinton, G. E., McClelland, J. L. & Rumelhart, D. E. (1986). Distributed representations. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition.* Vol. 1: *Foundations* (pp. 77-109). Cambridge: MIT Press/Bradford Books.

Hotopf, W. H. N. (1980). Semantic similarity as a factor in whole-word slips of the tongue. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 97-109). New York: Academic Press.

Humphreys, G. W., Riddock, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5,* 67-104.

Indefrey, P., & Goebel, R. (1993). The learning of weak noun declensions in German children vs. artificial neural networks. In *Proceedings of the fifteenth Annual Conference of the Cognitive Science Society* (pp. 575-580). Hillsdale, NJ: Erlbaum.

Jackendoff, R. (1983). *Semantics and cognition.* Cambridge, MA: MIT Press.

Jackendoff, R. (1987). *Consciousness and the computational mind.* Cambridge, MA: MIT Press.

James, W. (1890/1950). *The principles of psychology (Vol. 1).* Authorized edition, Dover Publications, Inc.

Jarvella, R. J., & Klein, W. (Eds.) (1982). *Speech, place and action. Studies in deixis and related topics.* Chichester: John Wiley.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 824-843.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge: University Press.

Johnson-Laird, P. N. (1989). Mental models. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 469-499). Cambridge, MA: MIT Press.

Jordan, M. I. (1986). Attractor dynamics and parallelism in a connectionist sequential machine. *Proceedings of the 8th Annual Conference of the Cognitive Science Society* (pp. 531-546). Hilsdale, NJ: Erlbaum.

Joshi, A., & Weinstein, S. (1981). Control of inference: Role of some aspects of discourse-structured centering. *Proceedings of the International Joint Conference on Artificial Intelligence* (pp. 385-387).

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.

Kintsch, W. (1994). The psychology of discourse processing. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 721-739). San Diego: Acadamic Press.

Kirby, K. N. (1993). *Advanced data analysis with SYSTAT.* Multiscience Press, Inc.

Klein, W. (1984). Bühlers Ellipse. In C. F. Graumann, & Th. Herrmann (Eds.), *Karl Bühlers Axiomatik. Fünfzig Jahre Axiomatik der Sprachwissenschaften.* Frankfurt: Klostermann.

Klein, W. (1993). Ellipse. In J. Jacobs, A. von Stechow, W. Sternefeld, & Th. Vennemann (Eds.), *Syntax. Ein internationales Handbuch zeitgenössischer Forschung,* Vol. 1 (pp. 763-799). Berlin: De Gruyter.

La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory and Cognition, 16,* 400-410.

La Heij, W., Happel, B., & Mulder M. (1990). Components of Stroop-like interference in word reading. *Acta Psychologica, 73,* 115-129.

La Heij, W., Dirkx, J., & Kramer, P. (1990). Categorical interference and associative priming in picture naming. *British Journal of Psychology, 81,* 511-525.

Levelt, W. J. M. (1982). Cognitive styles in the use of spatial direction terms. In R. J. Jarvella & W. Klein (Eds.), *Speech, place, and action: Studies in deixis and related topics* (pp. 251-268). Chichester: John Wiley.

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. (1992). Accessing words in speech production: Stages, processes, and representations. *Cognition, 42,* 1-22.

Levelt, W. J. M. (1993). The architecture of normal spoken language use. In G. Blanken, J. Dittmann, H. Grimm, J. C. Marschall, & C.-W. Wallesh (Eds.), *Linguistic disorders and pathologies: An international handbook* (pp. 1-15). Berlin: De Gruyter.

Levelt, W. J. M. (1996). Perspective taking and ellipsis in spatial development. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 77-108). Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (accepted). A theory of lexical access in speech production.

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

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991b). Normal and deviant lexical processing: Reply to Dell and O'Seaghdha (1991). *Psychological Review, 98,* 615-618.

Levinson, S. C. (1983). *Pragmatics.* Cambridge University Press.

Levinson, S. C. (1996). Frames of reference and Molyneux's question. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp.109-169). Cambridge, MA: MIT Press.

Lively, S. E., Pisoni, D. B., & Goldinger, S. D. (1994). Spoken word recognition: Research and theory. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 265-301). San Diego: Academic Press.

Love, T., & Swinney, D. (1996). Coreference processing and levels of analysis in object-relative constructions: Demonstration of antecedent reactivation with the cross-modal priming paradigm. *Journal of Psycholinguistic Research, 25,* 5-24.

Luce, R. D. (1959). *Individual choice behavior.* New York: Wiley.

Lupker, S. J., & Katz, A. N. (1981). Input, decision, and response factors in picture-word interference. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 269-282.

MacDonald, M. C., & MacWhinney, B. (1990). Measuring inhibition and facilitation from pronouns. *Journal of Memory and Language, 29,* 469-492.

MacKay, D. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology, 3,* 210-227.

MacKay, D. (1987). *The organization of perception and action. A theory for language and other cognitive skills.* New York: Springer.

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word, 15,* 19-44.

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

MacWhinney, B., & Bates, E. (1978). Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behaviour, 17,* 539-558.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25,* 71-102.

Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing* (pp. 148-172). Cambridge, MA: MIT Press.

Marslen-Wilson, W., Levy, E., & Tyler, L. K. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In R. Jarvella & W. Klein (Eds.), *Speech, place and action. Studies in deixis and related topics* (pp. 339-378). Chichester: John Wiley.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8,* 1-71.

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

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 101*(4), 653-675.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 576-585

Martin, N., Gagnon, D. A., Schwartz, M. F., Dell, G. S., & Saffran, E. M. (1996). Phonological facilitation of semantic errors in normal and aphasic speakers. *Language and Cognitive Processes, 11,* 257-282.

Martin, N., Weisberg, R. W., & Saffran, E. M. (1989). Variables influencing the occurrence of naming errors: Implications for models of lexical retrieval. *Journal of Memory and Language, 28,* 462-485.

McCauley, C., Parmelee, C. M., Sperber, R. D., & Carr, T. H. (1980). Early extraction of meaning from pictures and its relation to conscious identification. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 265-276.

McClelland, J. L., & Elman, J. L. (1986). Interactive processes in speech perception The TRACE Model. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition.* Vol. 2: *Psychological and biological models* (pp. 58-121). Cambridge: MIT Press/Bradford Books.

McDonald, J. L., & MacWhinney, B. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language, 34,* 543-566.

McEvoy, C. L. (1988). Automatic and strategic processes in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14* (4), 618-626.

McKoon, G., Ratcliff, R., & Ward, G. (1994). Testing theories of language processing: An empirical investigation of the on-line lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20* (5), 1219-1228.

McQueen, J. M., & Cutler, A. (1997). Cognitive processes in speech perception. In W. J. Hardcastle & J. Lawer (Eds.), *The handbook of phonetic sciences* (pp. 566-585). Oxford: Blackwell.

McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20* (3), 621-638.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language, 29,* 524-545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language, 30,* 69-89.

Meyer, A. S. (1992). Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives. *Cognition, 42,* 181-211.

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.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception.* Cambridge, MA: Harvard University Press.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Miller, G. A. (1969). A psychological method to investigate verbal concepts. *Journal of Mathematical Psychology, 6,* 169-191.

Miller, G. A., & Fellbaum, C. (1991). Semantic networks of English. *Cognition, 41,* 197-229.

Mills, A. E. (1985). The acquisition of German. In D. I. Slobin (Ed.), *The cross-linguistic study of language acquisition.* Vol. 1: *The data* (pp. 141-254). Hillsdale, NJ: Erlbaum.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review, 76,* 165-178.

Murphy, G. L. (1985). Processes of understanding anaphora. *Journal of Memory and Language, 24,* 290-303.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives of perception and action* (pp. 361-394). Hillsdale, NJ: Erlbaum,.

Neumann, O. (1992). Theorien der Aufmerksamkeit: Von Metaphern zu Mechanismen. *Psychologische Rundschau, 43,* 83-101.

Nicol, J., & Swinney, D. (1989). The role of structure and coreference assignment during sentence comprehension. Sentence Processing [Special issue]. *Journal of Psycholinguistic Research, 18,* 5-19.

Nicol, J., Fodor, J. D., & Swinney, D. (1994). Using cross modal lexical decision tasks to investigate sentence processing. *Journal of Experimental Psychology: Language, Memory, and Cognition, 20,* 1229-1238.

Norris, D. (1992). Connectionism: A new breed of bottom-up model? In R. G. Reilly & N. E. Sharkey (Eds.), *Connectionist approaches to natural language processing* (pp. 351-371). Hillsdale, NJ: Erlbaum.

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

Norris, D., & Cutler, A. (1988). The relative accessibility of phonemes and syllables. *Perception and Psychophysics, 43,* 541-550.

Oakhill, J., Garnham, A., & Vonk, W. (1989). The on-line construction of discourse models. *Language and Cognitive Processes, 4,* 263-286.

O'Seaghdha, P. G., Dell, G. S., Peterson, R. R., & Juliano, C. (1992). Models of from-related priming in comprehension and production. In R. G. Reilly & N. E. Sharkey (Eds.), *Connectionist approaches to natural language processing* (pp. 373-408). Hillsdale, NJ: Erlbaum.

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

O'Seaghdha, P. G., & Marin, J. W. (submitted). Phonological competition and coorperation in form-related priming. Sequential and nonsequential processes in perception and production.

Osterhout, L., & Swinney, D. (1993). On the temporal course of gap-filling during comprehension of verbal passives. *Journal of Psycholinguistic Research, 22,* 273-286.

Paivio, A. (1966). Latency of verbal associations and imagery to noun stimuli as a function of abstractness and generality. *Canadian Journal of Psychology, 20,* 378-387.

Paivio, A., Yuille, J. C., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76,* (1, Part 2).

Peterson, R. R., Dell, G. S., & O'Seaghdha, P. G. (1989). A connectionist model of form-related priming effects. In *Proceedings of the 11th Annual Conference of the Cognitive Science Society* (pp. 196-203). Hillsdale, NJ: Erlbaum.

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

Peterson, R. R., Shim, M., & Savoy, P. (1993). Time course of lexicalization in language production. Paper presented at the 34th annual meeting of the Psychonomic Society. Washington, D. C.

Pineda, F.J. (1987). Generalisation of back-propagation to recurrent neural networks. *Physical Review Letters, 59* (19), 2229-2232.

Plunkett, K., & Marchmann, V. (1991). U-shaped learning and frequency effects in a multi-layered perceptron: Implications of child language acquisition. *Cognition, 38,* 43-102.

Plunkett, K., & Marchmann, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition, 48,* 21-69.

Pollock, J.-Y. (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry, 20,* 365-424.

Posner, M I., & Snyder, C R. R. (1975) Attention and cognitive control. In R. L. Solso (Ed), *Information processing and cognition: The Loyola Symposion* (pp. 55-86). Hillsdale, NJ: Erlbaum.

Power, M J. (1983). Are there cognitive rhythms in speech? *Language and Speech, 26,* 253-261.

Prince, E. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed), *Radical pragmatics* (pp. 223-256) New York· Academic Press

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

Roelofs, A. (1992b). *Lemma retrieval in speaking· A theory, computer simulations, and empirical data.* Doctoral dissertation, NICI Technical Report 92-08, University of Nijmegen.

Roelofs, A. (1996) Serial order in planning the production of successive morphemes of a word *Journal of Memory and Language, 35,* 854-876

Roelofs, A. (1997). A case for nondecomposition in conceptually driven word retrieval. *Journal of Psycholinguistic Research, 26 (1),* 33-67

Roelofs, A. (in press). The WEAVER model of word-form encoding in speech production. *Cognition.*

Roelofs, A., Meyer, A. S., & Levelt, W. J M (1996). Interaction between semantic and orthographic factors in conceptually driven naming. Comment on Starreveld and La Heij (1995). *Journal of Experimental Psychology. Learning, Memory, and Cognition, 22,* 1, 246-251.

Rumelhart, D E , Hinton, G. E , & McClelland, J. L (1986). A general framework for parallel distributed processing. In D. E. Rumelhart & J. L. McClelland (Eds), *Parallel distributed processing. Explorations in the microstructure of cognition.* Vol. 1: *Foundations* (pp 45-76). Cambridge MIT Press

Rumelhart, D E , Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation In D E Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing. Explorations in the microstructure of cognition.* Vol. 1: *Foundations* (pp 318-362) Cambridge MIT Press.

Rumelhart, D E., & McClelland, J L. (1986). On learning the past tense of English verbs. In J L. McClelland & D E Rumelhart (Eds), *Parallel distributed processing: Explorations in the microstructure of cognition* Vol 2· *Psychological and biological models* (pp 216-271) Cambridge: MIT Press.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics, 2,* 437-442.

Sag, I A., & Hankamer, J. (1984). Toward a theory of anaphoric processing *Linguistics and Philosophy, 7,* 325-345.

Schiller, N. O., Meyer, A. S., Baayen, R. H , & Levelt, W. J. M. (1996). A comparison of lexeme and speech syllables in Dutch. *Journal of Quantitative Linguistics, 3,* 8-28.

Schriefers, H. (1990). Lexical and conceptual factors in the naming of relations *Cognitive Psychology, 22,* 11-142.

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

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

Segui, J., Frauenfelder, U. H., & Mehler, J. (1981). Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology, 72,* 471-477.

Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 65-76.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed developmental model of visual word recognition and naming. *Psychological Review, 96,* 523-568.

Selkirk, E. (1984). Phonology and syntax: The relation between sound and structure. Cambridge, MA: MIT Press.

Shallice, T., McLeod, P., & Lewis, K. (1985). Isolating cognitive modules with the dual-task paradigm: Are speech perception and production separate processes? *Quarterly Journal of Experimental Psychology, 37A,* 507-532.

Shattuck-Huffnagel, S. (1979). Speech errors as evidence for a serial-order mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 295-342). Hillsdale, NJ: Erlbaum.

Shattuck-Huffnagel, S. (1987). The role of word onset consonants in speech production planning: New evidence from speech error patterns. In E. Keller & M. Gopnik (Eds.), *Motor and sensory processes of language* (pp. 17-51). Hillsdale, NJ: Erlbaum.

Slowiaczek, L. M., & Hamburger, M. B. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18* (6), 1239-1250.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 686-698.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 896-918.

Stemberger, J. P. (1985). An interactive activation model of language production. In W. W. Ellis (Ed.), *Progress in the psychology of language,* Vol. 1 (pp. 143-186). Hillsdale, NJ: Erlbaum.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Tanenhaus, M. K., & Carlson, G. N. (1990). Comprehension of deep and surface verbphrase anaphors. *Language and Cognitive Processes, 5,* 257-280.

Terken, J. M. B. (1984). The distribution of pitch accents in instructions as a function of discourse structure. *Language and Speech, 27,* 269-289.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology, 37A,* 571-590.

Tipper, S. P., & Driver, J. (1988). Negative priming between pictures and words in a selective attention task: Evidence for semantic processing of ignored stimuli. *Memory and Cognition, 16,* 1, 64-70.

Tipper, S. P., MacQueen, G. M., & Brehaut, J. C. (1988). Negative priming between response modalities: Evidence for the central locus of inhibition in selective attention. *Perception and Psychophysics, 43,* 45-52.

Tukey, J. W. (1977). *Exploratory data analysis.* Addison-Wesley Publishing Company, Inc.

Van Berkum, J. J. A. (1996). *The psycholinguistics of grammatical gender.* Doctoral dissertation, University of Nijmegen.

Van Berkum, J. J. A. (in press). Syntactic Processes in Speech Production: The retrieval of Grammatical Gender. *Cognition.*

Van der Heijden, A. H. C. (1981). *Short-term visual information forgetting.* London: Routledge and Kegan Paul.

Van Kuijk, D., Wittenburg, P., & Dijkstra, T. (1996). RAW: A real-speech model for human word recognition. In H. T. Bunnell and W. Idsardi (Eds.), *Proceedings of the International Conference on Spoken Language Processing '96* (pp. 70-73). New Castle, Delaware: Citation Delaware.

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 Wijk, C., & Kempen, G. (1987). A dual system for producing self-repairs in spontaneous speech: Evidence from experimentally elicited corrections. *Cognitive Psychology, 19,* 403-440.

Visser, J. (1992) Imperatieven. Een syntaktische analyse van imperatief-constructies in het Nederlands en het Engels. *Unpublished master thesis.* University of Tilburg.

Vorberg, D. (1985). Unerwartete Folgen von zufälliger Variabilität: Wettlauf-Modelle für den Stroop-Versuch. *Zeitschrift für experimentelle und angewandte Psychologie, Band XXXII, Heft 3,* 494-521.

Westermann, G., & Goebel, R. (1995). Connectionist rules of language. In J. D. Moore & J. F. Lehmann (Eds.), *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 236-241). Hillsdale, NJ: Erlbaum.

Wheeldon L. R., & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language, 33,* 332-356.

Wheeldon, L. R., & Monsell, S. (1992). The locus of repetition priming of spoken word production. *Quarterly Journal of Experimental Psychology, 44A* (4), 723-761.

Williams, E. S. (1977). Discourse and logical form. *Linguistic Inquiry, 8,* 102-139.

Winer, J. B. (1971). *Statistical principles in experimental design.* New York McGraw-Hill.

Zubin, D. A., & Köpke, K.-M. (1981). Gender: A less than arbitrary grammatical category. In R. A. Hendrick, C. S. Masek, & M. F. Miller (Eds.), *Papers from the seventh regional meeting of the Chicago Linguistic Society* (pp. 439-449). Chicago, IL: University of Chicago Press.

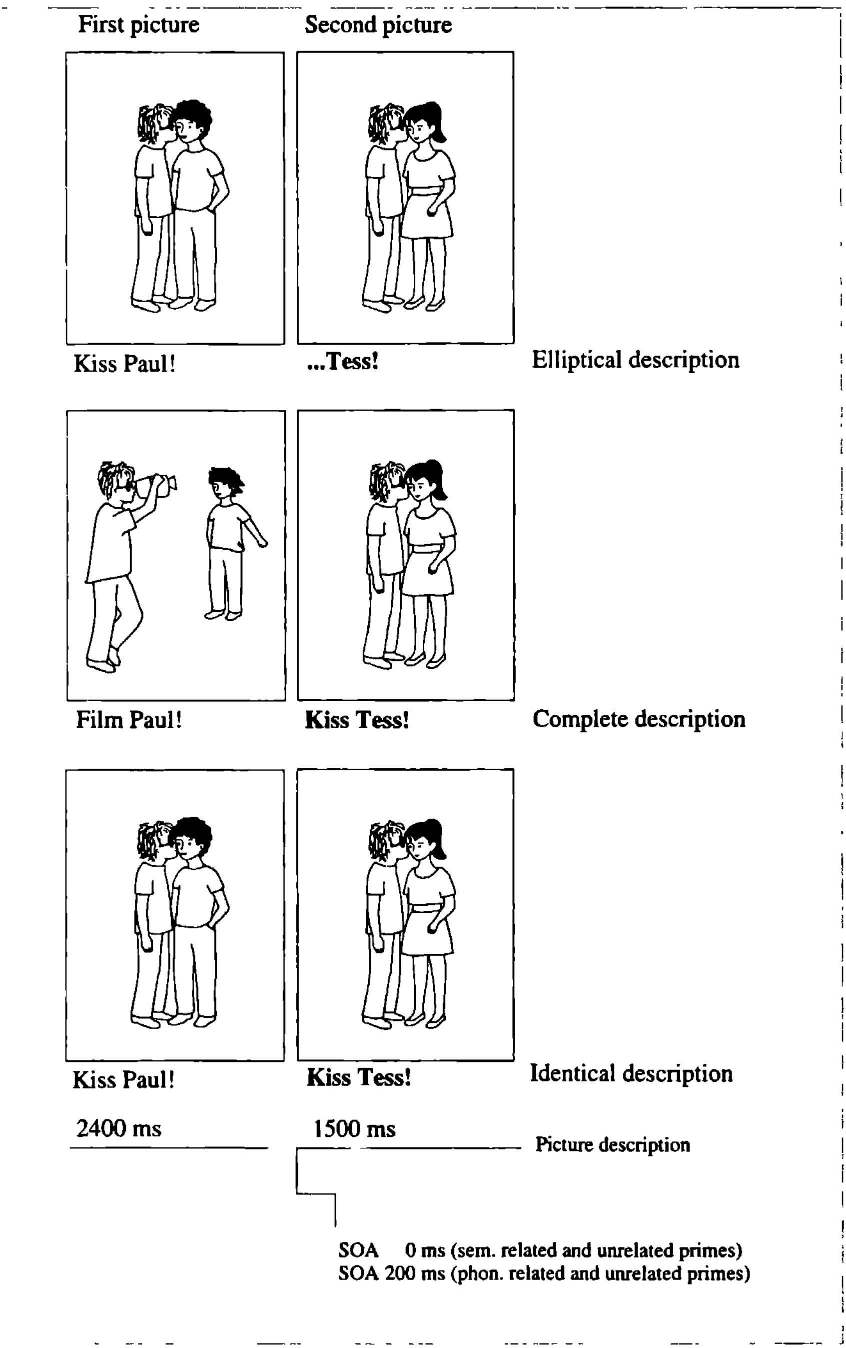
Zurif, E., & Swinney, D. (1994). The neuropsychology of language. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 1055-1074). San Diego: Academic Press.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-workd processing. *Cognition, 32,* 25-64.

Zwitserlood, P. (1994). Access to phonological form representations in language comprehension and production. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp.83-106). Hillsdale, NJ: Erlbaum.

Zwitserlood, P. (1996). Form Priming. *Language and Cognitive Processes, 11* (6), 589-596.

APPENDIX A
Illustration of picture description conditions and trial structure used in the ellipsis experiments (Chapter 2)



APPENDIX B

Materials used in the ellipsis experiments (Chapter 2)

Target verb (action depicted at the second picture)	**Prime**			
	Semantically related	**Semantically unrelated**	**Phonologically related**	**Phonologically unrelated**
was (wash)	douche (shower)	bijt	wacht (wait)	droom
droog (dry)	baad (bath)	wenk	droom (dream)	kaft
kam (comb)	knip (cut)	wek	kaft (cover)	groei
groet (greet)	wenk (beckon)	sla	groei (grow)	wacht
roep (call)	wek (wake)	baad	roer (stir)	berg
bel (call)	meld (announce)	schets	berg (salvage)	schok
schop (kick)	bijt (bite)	douche	schok (shock)	print
prik (inject)	sla (beat)	meet	print (print)	wurm
wurg (strangle)	steek (stick)	knip	wurm (squeeze)	troon
draai (turn)	schud (shake)	gips	draal (hesitate)	week
draag (carry)	sleep (drag)	kroon	draaf (canter)	volg
duw (push)	trek (pull)	zoog	duur (last)	span
weeg (weigh)	meet (measure)	spuit	week (soak)	draal
spalk (splint)	gips (plaster)	vind	span (tighten)	voel
troost (comfort)	steun (support)	bouw	troon (parade)	duur
voer (feed)	zoog (wean)	schmink	voel (feel)	lach
lak (varnish)	spuit (spray)	trek	lach (laugh)	vel
verf (paint)	schmink (make up)	boei	vel (fell)	toon
tooi (clothe)	kroon (crown)	steun	toon (show)	fiks
film (film)	schets (sketch)	meld	fiks (organize)	draaf
vorm (shape)	bouw (build)	steek	volg (follow)	zoem
zoek (seek)	vind (find)	schud	zoem (buzz)	val
vang (catch)	boei (handcuff)	sleep	val (fall)	roer

APPENDIX C

Materials used in the first main experiment (Chapter 3)

Picture/ identical Probe		**Phonologically related probe**		**Semantically related probe**		**unrelated probe**	
Apfel	*appel*	Abfall	*rubbish*	Pfirsich	*peach*	Tresor	*safe*
Auge	*eye*	Auto	*car*	Ohr	*ear*	Bad	*bathroom*
Ballon	*balloon*	Balkon	*balcony*	Drachen	*dragon*	Wecker	*alarmclock*
Banane	*banana*	Bank	*bank*	Zitrone	*lemon*	Kelle	*ladle*
Berg	*mountain*	Becher	*cup*	Wald	*forest*	Laden	*shop*
Blitz	*lightning*	Blick	*look*	Orkan	*hurricane*	Reis	*rice*
Buerste	*brush*	Buechse	*tin*	Spange	*hairpin*	Polizei	*police*
Bus	*bus*	Bulle	*bull*	Zug	*train*	Dolch	*dagger*
Eimer	*bucket*	Eiter	*pus*	Topf	*pot*	Kanal	*canal*
Ellipse	*ellipse*	Elle	*yardstick*	Raute	*rhomb*	Palme	*palm*
Fass	*barrel*	Fach	*drawer*	Sieb	*sieve*	Brot	*bread*
Fenster	*window*	Fett	*fat*	Sofa	*sofa*	Schnitzel	*escalope*
Flasche	*bottle*	Flamme	*flame*	Dose	*can*	Scheune	*barn*
Frosch	*frog*	Frost	*frost*	Lurch	*batrachian*	Deckel	*lid*
Gans	*goose*	Gasse	*lane*	Ente	*dug*	Hutte	*hut*
Gewehr	*rifle*	Geweih	*horns*	Messer	*knife*	Dia	*dia*
Gras	*gras*	Grab	*grave*	Blatt	*leaf*	Bier	*beer*
Hammer	*hammer*	Hammel	*wether*	Spaten	*spade*	Vogel	*bird*
Harke	*rake*	Halle	*hall*	Schaufel	*shovel*	Pille	*pill*
Hose	*trousers*	Hochzeit	*wedding*	Jacke	*jacket*	Alge	*alga*
Hund	*dog*	Husten	*cough*	Fisch	*fish*	Teller	*board*
Hut	*hat*	Huf	*hoof*	Schal	*shawl*	Block	*block*
Kabel	*cable*	Kanu	*canoe*	Seil	*rope*	Moped	*motor bike*
Kaktus	*cactus*	Kasten	*chast*	Flieder	*elder*	Napf	*bowl*
Kanone	*cannon*	Kapelle	*chapel*	Bombe	*bomb*	Nudel	*noodle*
Katze	*cat*	Kappe	*cap*	Maus	*mous*	Saege	*saw*
Kirche	*church*	Kiste	*box*	Moschee	*mosque*	Zahl	*number*
Kleid	*dress*	Klavier	*piano*	Tuch	*cloth*	Horn	*horn*
Kreis	*circle*	Krug	*jug*	Wuerfel	*die*	Brief	*letter*
Kutsche	*coach*	Kuppel	*cupola*	Gondel	*gondola*	Torte	*pie*
Lampe	*lamp*	Lanze	*lance*	Gardine	*curtain*	Seife	*soap*
Lineal	*ruler*	Lied	*song*	Band	*ribbon*	Nest	*nest*
Nabel	*navel*	Name	*name*	Hals	*neck*	Kaese	*cheese*
Nase	*nose*	Nadel	*needle*	Hand	*hand*	Weste	*vest*
Parfuem	*parfume*	Papier	*paper*	Shampoo	*shampoo*	Feuer	*fire*
Pfeil	*arrow*	Pfau	*peacock*	Speer	*spear*	Bart	*beard*
Pinsel	*paint brush*	Pilz	*mushroom*	Gips	*plaster*	Fels	*rock*
Pinzette	*tweezers*	Pistole	*gun*	Feile	*file*	Kamera	*camera*
Puzzle	*puzzle*	Pulver	*powder*	Domino	*domino*	Fell	*hide*
Quadrat	*square*	Quartett	*quartett*	Dreieck	*triangle*	Benzin	*gasoline*
Schiff	*boat*	Schild	*board*	Floss	*float*	Rohr	*pipe*
Schlange	*snake*	Schlacht	*battle*	Kroete	*paddock*	Waage	*scale*
Schwein	*pig*	Schloss	*castle*	Pferd	*horse*	Buch	*book*
Sonne	*sun*	Sohle	*sole*	Erde	*soil*	Bremse	*brake*
Teelicht	*candle*	Telefon	*telephone*	Streichholz	*match*	Atom	*atom*
Teppich	*carpet*	Tempel	*temple*	Lauefer	*drugget*	Bugel	*hanger*
Zebra	*zebra*	Zehntel	*tenth(part)*	Kamel	*camel*	Radio	*radio*
Zelt	*tent*	Zentrum	*center*	Iglo	*iglo*	Lamm	*lamb*

APPENDIX D

Materials used in the pronoun experiments (Chapter 3)

second picture in 1 pronoun experiment/ **first picture** in 2 pronoun experiment	**phonologically related probe**	**unrelated probe**	**Gender**
Apfel (apple)	Abfall (rubbish)	Tresor (safe)	masculine
Burste (brush)	Buchse (tin)	Polizei (police)	feminine
Eimer (bucket)	Eiter (pus)	Kanal (canal)	masculine
Ellipse (ellipse)	Elle (yardstick)	Palme (palm)	feminine
Fass (barrel)	Fach (drawer)	Brot (bread)	neuter
Fenster (window)	Fett (fat)	Schnitzel (escalope)	neuter
Flasche (bottle)	Flamme (flame)	Scheune (barn)	feminine
Frosch (frog)	Frost (frost)	Deckel (lid)	masculine
Gans (goose)	Gasse (lane)	Huette (hut)	feminine
Gras (gras)	Grab (grave)	Bier (beer)	neuter
Hammer (hammer)	Hammel (wether)	Vogel (bird)	masculine
Kaktus (cactus)	Kasten (box)	Napf (bowl)	masculine
Lineal (ruler)	Lied (song)	Nest (nest)	neuter
Nase (nose)	Nadel (needle)	Weste (vest)	feminine
Quadrat (square)	Quartett (quartet)	Benzin (gasoline)	neuter
Zelt (tent)	Zentrum (center)	Lamm (lamb)	neuter

List of 16 Pseudowords used in the pronoun experiments

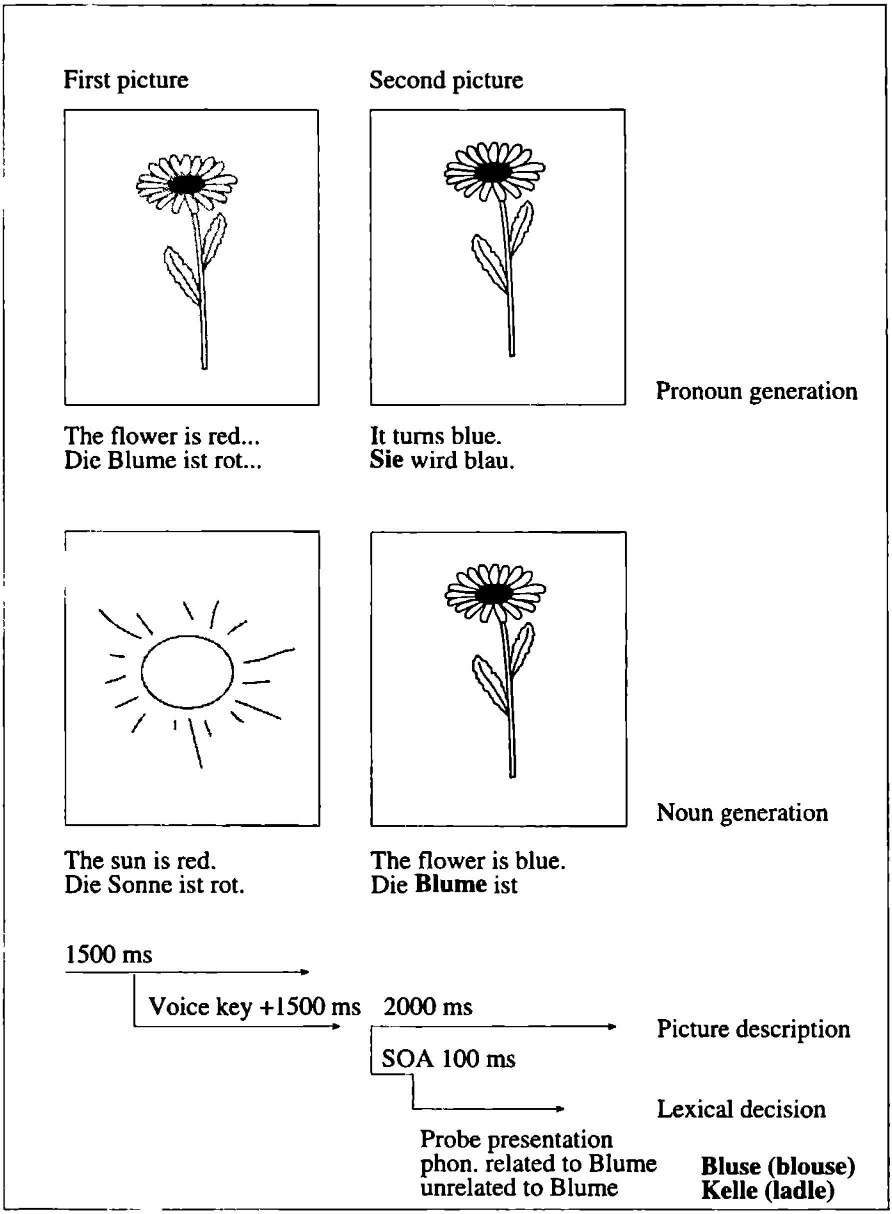
Banako, Eip, Holk, Humpf, Kator, Mosak, Natis, Nepf, Pidda, Pufke, Quesse, Schill, Stulp, Trekal, Vodil, Zars

List of 48 Pseudowords used in the first main experiments

Bannul, Beusel, Done, Dos, Ellt, Erke, Festa, Fliemc, Flon, Foka, Fronk, Galp, Gapf, Garluse, Gewoht, Gilk, Gopfel, Haal, Himbule, Honer, Humf, Igra, Kanafe, Kaspe, Kerst, Klassir, Kna, Knolk, Kreuk, Krose, Limior, Mimpf, Mopfo, Nepf, Oem, Pfellor, Pfum, Pinralle, Pufke, Schlaste, Schweil, Seik, Sepf, Siem, Soka, Tolk, Tuke, Zuf

APPENDIX E

Illustration of picture description conditions and trial structure of the first pronoun experiment (Chapter 3)



The flower is red...
Die Blume ist rot...

It turns blue.
Sie wird blau.

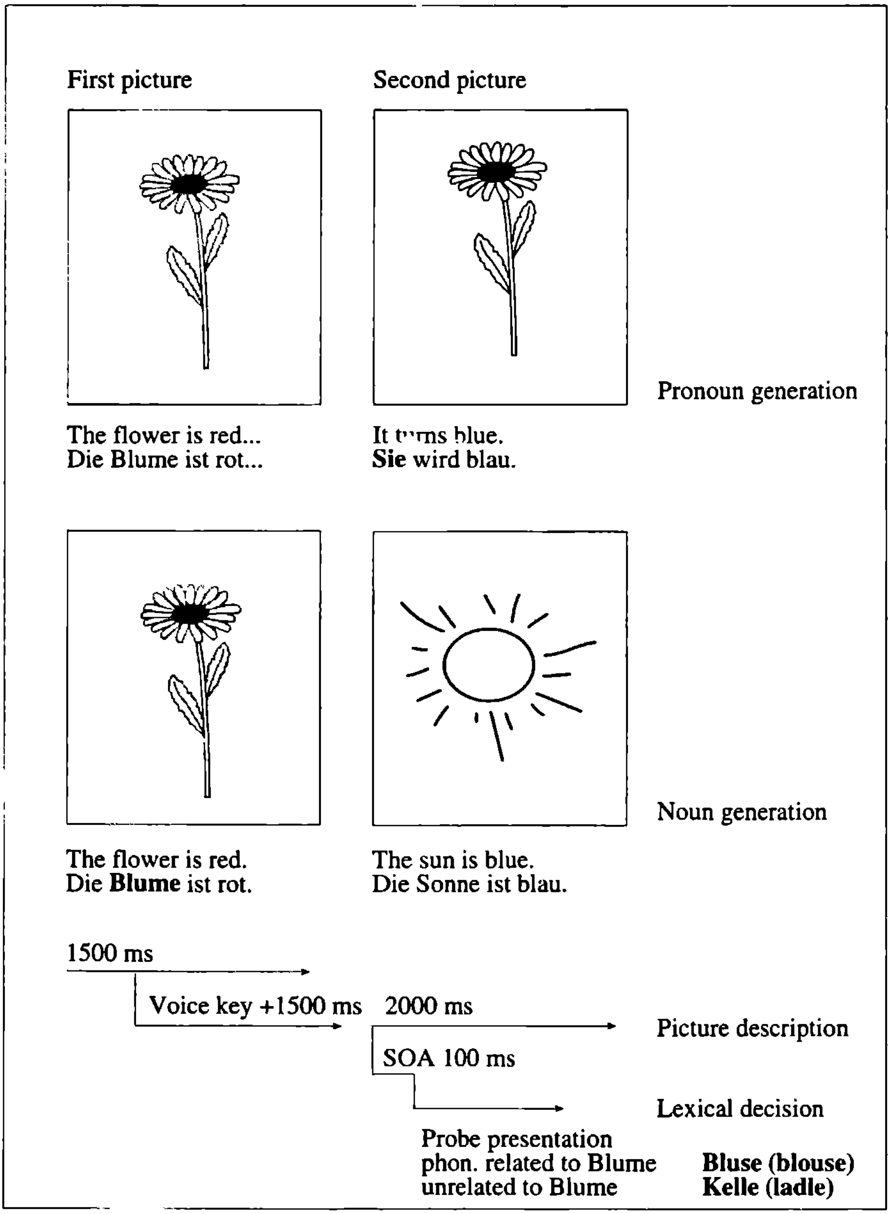
The sun is red.
Die Sonne ist rot.

The flower is blue.
Die **Blume** ist

Bluse (blouse)
Kelle (ladle)

APPENDIX F

Illustration of picture description conditions and trial structure of the second pronoun experiment (Chapter 3)



The flower is red...
Die Blume ist rot...

It turns blue.
Sie wird blau.

The flower is red.
Die **Blume** ist rot.

The sun is blue.
Die Sonne ist blau.

Probe presentation

phon. related to Blume	**Bluse (blouse)**
unrelated to Blume	**Kelle (ladle)**

SAMENVATTING

De hoofdvraag van mijn dissertatie-onderzoek is de vraag hoe lexicale toegang plaatsvindt tijdens de productie van gereduceerde vormen zoals ellipsen en voornaamwoorden. In modellen van taalproductie wordt verondersteld dat lexicale toegang uit tenminste twee stadia bestaat, te weten semantische/ syntactische codering en fonologische codering. In deze dissertatie wordt onderzocht of zowel de betekenis als de fonologische vorm van een woord beschikbaar zijn als een anaforische constructie wordt gebruikt die verwijst naar een eerdere referent. Volgens het stricte-twee-stadia-model (Levelt et al., 1991a) zou de betekenis van een referent actief worden omdat de anafoor verwijst naar de betekenis van de referent. De fonologische vorm van de referent zou echter niet actief worden, omdat de fonologische vorm van de referent niet nodig is: de referent wordt namelijk niet expliciet uitgesproken. Vanuit het perspectief van theorieën die veronderstellen dat er een cascade van verspreidende activatie is (Dell, 1986; Dell & O'Seaghdha, 1992), kan een andere assumptie gemaakt worden. In dit geval zou de betekenis van de referent actief worden, die op haar beurt het lemma activeert. In tegenstelling tot in het stricte-twee-stadia-model wordt in een cascade-model ook de fonologische vorm actief, omdat het lemma automatisch activatie verspreidt naar de fonologische vorm, ook al wordt de referent niet uitgesproken.

Het onderzoek dat in deze dissertatie wordt beschreven bestaat uit drie delen. In het eerste deel onderzocht ik met het plaatje-woord-interferentieparadigma de semantische en fonologische activatie van een referent tijdens de productie van ellipsen. In het tweede deel ging ik van ellipsen over naar voornaamwoorden en deed onderzoek naar de fonologische activatie van de referent tijdens de productie van voornaamwoorden. Het paradigma dat ik hierbij gebruikte was een dubbeltaakparadigma. Deelnemers moesten als eerste taak plaatjes beschrijven en, als tweede taak, lexicale decisies maken op akoestische stimuli. In het derde deel gebruikte ik de PDP-benadering voor modelactivatieprocessen in de dubbeltaaksituatie om het in de voornaamwoord-experimenten gevonden fonologische effect te verklaren. Afsluitend vergelijk ik de resultaten die in de ellipsen- en voornaamwoord-experimenten werden gevonden.

A. Lexicale toegang tijdens de productie van ellipsen

Het doel van de experimenten in hoofdstuk 2 was te onderzoeken hoe lexicale toegang verloopt tijdens de productie van ellipsen. Wordt een woord semantisch en fonologisch geactiveerd als het weggelaten wordt in de tweede zin van twee aaneengesloten zinnen? In een serie in het Nederlands uitgevoerde experimenten gebruikte ik gedeeltelijke correctie als vorm van de uiting, *zoals "Kus Toon...Paul!"*. Er zijn redenen om te veronderstellen dat het werkwoord *"kussen"* actief is in de opening tussen Toon en Paul. De complete structuur van de uiting is *"Kus Toon, ... nee, ... kus Paul!"*. Omdat het werkwoord niet wordt gearticuleerd in de elliptische uitdrukking, is het mogelijk dat er geen fonologische activatie wordt gevonden op de openingspositie. Deze hypothese werd getest met een plaatje-woord-interferentieparadigma. De deelnemers moesten een sequentie van twee plaatjes benoemen. De beschrijving van het eerste plaatje was bijvoorbeeld *"Kus Toon"*. Op het tweede plaatje veranderde de patients van Toon naar Paul. De deelnemers werd gevraagd om deze verandering te interpreteren als een soort correctie van de visuele omgeving en deze hardop te benoemen. Dit leidde tot een beschrijving van de twee plaatjes in een volgorde als *"Kus Toon...Paul!"*. Tijdgebonden aan het begin van de presentatie van het tweede plaatje werd een akoestische stimulus gepresenteerd. Deze stimulus was ofwel semantisch of fonologisch gerelateerd, ofwel ongerelateerd aan het gerefereerde werkwoord. In deze taak voorspelt een strict-twee-stadia-model semantische inhibitie, maar geen fonologisch effect voor gerelateerde stimuli ten opzichte van een ongerelateerde stimuli.

De resultaten van het eerste ellipsen-experiment toonden een semantisch inhibitie-effect aan, maar lieten geen fonologisch effect zien voor het weggelaten werkwoord tijdens de productie van een ellips. In een controle-conditie, waar het werkwoord niet was weggelaten zoals in *"Voed Toon ... Kus Paul"*, werd semantische inhibitie en fonologische facilitatie gevonden voor het werkwoord *"kus"*. De resultaten van de controle-conditie tonen aan dat het geobserveerde fonologische nuleffect tijdens de productie van een ellips niet te wijten kan zijn aan de methodologische aspecten van het experiment, maar moet worden geïnterpreteerd als een linguïstisch effect. De semantische inhibitie tijdens de productie van een ellips werd geïnterpreteerd als afkomstig van de activatie van het lemma van het weggelaten werkwoord. Het fonologische nuleffect laat zien dat het lemma geen activatie verspreidt naar zijn corresponderende fonemen op het fonologische niveau van de spraakproductie.

In het tweede ellipsen-experiment, dat een verandering van sprekers betrof, werden de resultaten van de ellipsen-condities in het eerste experiment gerepliceerd. Samen bevestigen de resultaten van de experimenten Kleins (1993) assumptie van p-reductie en sprekeronafhankelijkheid tijdens de productie van ellipsen. Bovendien bevestigen deze resultaten de stricte-twee-stadia-theorie: een actief lemma dat niet wordt geselecteerd voor overte benoeming zal geen activatie verspreiden naar zijn corresponderende fonemen.

B. Lexicale toegang tijdens de productie van voornaamwoorden

In hoofdstuk 3 werd een tweede vorm van reductie onderzocht met betrekking tot het proces van lexicale toegang: de productie van voornaamwoorden. Hoe verloopt de lexicale toegang voor voornaamwoorden? En idee is als volgt (zie figuur 3.2. en 5.1): het plaatje activeert zijn concept. Het concept activeert op zijn beurt het lemma en via het lemma wordt de corresponderende syntactische informatie geactiveerd, zoals het grammaticale woordgeslacht. Als de discourse aangeeft dat de huidige eenheid in focus is, dan zal het discoursemechanisme toegang vanuit de grammaticale-geslachtsknopen toestaan naar de voornaamwoordknopen door de verbinding tussen grammaticale geslachtsknopen en de voornaamwoorden te openen. De grammaticale geslachtsinformatie is nodig om het juiste voornaamwoord te produceren (het mannelijke, het vrouwelijke en het onzijdige voornaamwoord in het Duits). Het voornaamwoord wordt dan geselecteerd en activeert zijn fonemen.

In de experimenten onderzocht ik de activatie van het lemma en de beschikbaarheid van zijn fonemen tijdens de voornaamwoordproductie. De vraag of de spreker de fonologische vorm activeert van een zelfstandig naamwoord wanneer eraan gerefereerd wordt met een voornaamwoord werd onderzocht voor uitingen zoals *"De bloem is rood...zij wordt blauw"*. Hier verwijst het voornaamwoord *"zij"* naar het zelfstandige naamwoord "bloem". Een theorie die cascade-verspreidende-activatie veronderstelt (Dell, 1986; Dell & O'Seaghdha, 1991, 1992) zou voorspellen dat het lemma van het gerefereerde zelfstandige naamwoord fonologisch actief wordt. In tegenstelling hiermee zou men volgens het stricte-twee-stadia-model van lexicale toegang (Levelt et al., 1991a; Roelofs, 1992a, b) geen fonologische activatie verwachten van het gerefereerde zelfstandige naamwoord op de voornaamwoordpositie, omdat het lemma niet wordt geselecteerd voor articulatie. Deze hypothese werd getest in experimenten in het Duits, waarin de deelnemers lexicale decisies maakten tijdens plaatjesbeschrijvingen. Ze beschreven een opeenvolging van twee plaatjes. De eerste plaatjesbeschrijving zoals *"De bloem is rood"* bevatte een referentie naar een object. In het tweede plaatje verscheen dit object opnieuw maar in een deel van de gevallen in een andere kleur. In vijftig procent van alle gevallen werd tijdgebonden aan het begin van het tweede plaatje een akoestische stimulus aangeboden. De stimulus kon ofwel een woord ofwel een pseudo-woord zijn. In deze gevallen was de taak voor de deelnemer de plaatjesbeschrijving uit te stellen en eerst een lexicale decisie te maken op de stimulus. De woordstimulus was fonologisch gerelateerd of ongerelateerd aan het zelfstandige naamwoord van het plaatje. De resultaten lieten fonologische inhibitie zien, wat suggereert dat tijdens de productie van het voornaamwoord de fonologische vorm van het gerefereerde zelfstandige naamwoord actief is.

Om te onderzoeken of de fonologische activatie te wijten is aan residue activatie van de eerste plaatjesbeschrijving werd een vervolgexperiment uitgevoerd. Wederom hadden de deelnemers als taak een sequentie van twee plaatjes te beschrijven, zoals *"De bloem is rood... De zon is blauw"*. Zoals in het voorgaande experiment was de stimuluspresentatie tijdgebonden aan het begin van het tweede plaatje. Maar nu was de stimulus fonologisch gerelateerd of ongerelateerd aan het zelfstandige naamwoord van het eerste plaatje. Er werd geen fonologisch effect gevonden in de voornaamwoordconditie.
Dit resultaat suggereert dat de fonologische vorm van de referent van het eerste plaatje niet meer beschikbaar is tijdens de presentatie van de akoestische stimulus net na het verschijnen van het tweede plaatje. Het gevonden fonologische effect tijdens de voornaamwoord-productie kan daarom niet geïnterpreteerd worden als afkomstig van residu activatie van de overte benoeming van het zelfstandige naamwoord voor het eerste plaatje, maar moet geïnterpreteerd worden als een gevolg van herhaalde toegang tot het lemma.

De geobserveerde fonologische inhibitie steunt de cascade-verklaring van activatieverspreiding tijdens lexicale toegang (zie figuur 3.3), die het fonologische effect lokaliseert op het fonologische niveau, en een feedback veronderstelt van de vertraagde fonologische codering naar het lemmaniveau. De fonologische inhibitie tijdens de voornaamwoordproductie is in tegenspraak met het stricte-twee-stadia-model zoals opgesteld door Levelt et al. (1991a). In dit model werd het fonologische effect tijdens de benoeming van een enkel zelfstandig naamwoord verklaard door een competitie van overlappende fonologische stadia. Volgens het model mag deze competitie niet plaats vinden tijdens de productie van voornaamwoorden omdat het gerefereerde zelfstandige naamwoord niet wordt geselecteerd wordt voor benoeming. De bevindingen zijn dus in strijd met de voorspellingen volgens de Levelt et al. theorie van 1991. Maar, zoals bediscussieerd is in de conclusie van hoofdstuk 3, kan de stricte-twee-stadia-theorie volgens meer recente ideeën, het fonologische effect verklaren in termen van cohortprocessen op het lemmaniveau. De experimenten kunnen daarom niet eenduidig onderscheid maken tussen de twee verschillende taalproductie-theorieën (lexicale toegang als cascade versus als strict-twee-stadia). Maar volgens beide theorieën betekent het geobserveerde fonologische effect dat het lemma van het gerefereerde zelfstandige naamwoord actief is tijdens voornaamwoordproductie. Het verkregen effect gaf daarom empirische steun aan de lexicale toegang-verklaring van voornaamwoordproductie (zoals te zien is in figuur 3.2., en 5.1.).

C. Fonologische inhibitie op lexicale decisie tijdens de productie van zelfstandige naamwoorden en voornaamwoorden: een PDP-benadering.

In hoofdstuk 4 werd een PDP-benadering gebruikt om activatieprocessen in een dubbeltaaksituatie te simuleren. Het computationele model werd gebruikt als een gereedschap om het fonologische interferentie-effect te lokaliseren op ofwel het fonologische niveau of het lemmaniveau (zoals wordt voorgesteld in het stricte-twee-stadia-model). Het model bestaat uit niveaus van spraakverwerking zoals verondersteld in huidige taalproductie-theorieën. Het computationele model werd eerst getraind om lexicale decisies uit te voeren op 'akoestische' stimuli, en vervolgens om plaatjes te benoemen. Na de trainingsfase werd de prestatie van het model in een dubbeltaak onderzocht.

In een eerste simulatie werd de assumptie van de cascade-verwerkingsopvatting getest. Deze opvatting verklaart de fonologische inhibitie die werd geobserveerd in de voornaamwoordexperimenten in termen van segmentele competitie op het fonologische niveau. Dit betekent dat de fonologische inhibitie alleen dan geobserveerd kan worden als het lemma van het zelfstandige naamwoord waarnaar het voornaamwoord verwijst, fonologisch actief wordt en interfereert met de segmentele codering van het gerelateerde stimuluswoord. Om deze segmentele competitie te modelleren in een lexicale decisietaak activeert de vertraagde fonologische codering van een gerelateerde akoestische stimulus zijn lemma later, wat leidt tot een vertraagde lexicale decisie. Om de feedback van het fonologische niveau naar het lemmaniveau te modelleren, had het computationele model feedback-verbindingen tussen het fonologische niveau naar het lemmaniveau.

De simulatieresulaten repliceerden de gedragsdata van Levelt et al. (1991a), met uitzondering van de identieke conditie. Deze conditie liet geen inhibitie maar facilitatie zien, zoals ook werd geobserveerd in mijn eigen replicatie van Levelt et al. (zie hoofdstuk 3, het eerste hoofdexperiment). Het belangrijkste is dat het model de fonologische inhibitie repliceerde tijdens de lexicale decisie zoals werd geobserveerd door Levelt et al. en de replicatie daarvan. Een descriptieve analyse van de patronen van gewichten en het tijdsverloop van activatie verklaarde het geobserveerde effect door inhibitoire feedback van fonologische "mismatching" segmenten naar het lemmaniveau. De simulatie ondersteunt daarom de opvatting van cascade-verwerking.

In een tweede simulatie werd in de architectuur van het computationele model geen feedback geïmplementeerd tussen het fonologische niveau en het

lemmaniveau van benoeming. De simulatie moest de cohort-verklaring testen waarbij de fonologische inhibitie op het lemmaniveau gesitueerd is zoals wordt verondersteld in het twee-stadia-model. Volgens de cohort-verklaring activeert de akoestische stimulus in eerste instantie een cohort van lemma's met hetzelfde woordbegin. Hoe meer akoestische informatie het lemmaniveau bereikt, des te kleiner wordt het cohort, tot een enkel element overblijft. Een fonologisch gerelateerde benoeming voor een plaatje maakt ook deel uit van dit cohort en omdat ze sterk geactiveerd is tijdens het benoemingsproces is ze een sterke tegenstander voor een fonologisch gerelateerd lemma. Deze competitie leidt tot een vertraagde verkleining van het cohort, wat resulteert in langere lexicale decisietijden.
De simulatieresultaten repliceerden de fonologische inhibitie bij korte SOA's zoals werd geobserveerd in het onderzoek van Levelt et al., en in het replicatie-onderzoek. De simulatie liet echter een afname van de fonologische inhibitie zien bij lange SOA's, terwijl de empirische data van Levelt et al. een dergelijke afname niet lieten zien.

De vergelijking tussen de twee simulatie-onderzoeken leidde tot de beslissing om de voorkeur te geven aan de feedback-variant en de voornaamwoord-productie te simuleren met deze versie van het model (zie simulatie 3 in hoofdstuk 4). De netwerkarchitectuur werd uitgebreid met een discourse-module, en met een voornaamwoordknoop. Het moest de in hoofdstuk 3 voorgestelde verklaring van lexicale toegang simuleren. Naast het leren van de lexicale decisie en het benoemen van zelfstandige naamwoorden, werd het computationele model getraind om afhankelijk van de discourse-context zelfstandige naamwoorden of voornaamwoorden te produceren. Deze discourse-afhankelijke prestatie werd gerealiseerd door een poortfunctie in te bouwen. De poort staat toegang van lemmas tot de voornaamwoorden toe als de discourse-module aangeeft dat de huidige entiteit in focus is. De poort is gesloten als de discourse-module niet actief is, wat leidt tot een overte benoeming van het zelfstandige naamwoord.
Het model simuleerde succesvol de empirisch geobserveerde fonologische inhibitie bij korte SOA's zowel voor de productie van zelfstandige naamwoorden als voornaamwoorden. De prestatie van het computationele model met betrekking tot het tijdsverloop van fonologische activatie tijdens de productie van voornaamwoorden kan verklaard worden door een samenspel van drie componenten: een vroege conceptuele inhibitie; inhibitoire feedback van een fonologische mismatch, en verschillende paden van fonologische activatie van voornaamwoorden en zelfstandige naamwoorden. Het model bevestigt daarom de lexicale toegangsverklaring van voornaamwoordproductie.

Het model geeft tevens voorspellingen voor resultaten voor voornaamwoorden bij lange SOA's die nog empirisch moeten worden getoetst. Bovendien bood het zelf-organiserende vermogen een oplossing voor een co-activatieprobleem; als fonemen van het zelfstandig naamwoord en het voornaamwoord op hetzelfde moment actief worden zoals het geval kan zijn in een cascade-proces, hoe kan het dan dat alleen het voornaamwoord wordt uitgesproken ? Het model geeft een mogelijke verklaring: de productie van voornaamwoorden wordt later verworven dan die van zelfstandige naamwoorden. In het model leidt deze latere training tot sterke inhibitoire verbindingen tussen het lemmaniveau en de mismatching segmenten (van het zelfstandige naamwoord) op het fonologische niveau. Een licht geactiveerde voornaamwoordknoop kan daarom gemakkelijk de al beschikbare fonologische informatie van het zelfstandige naamwoord onderdrukken.

D. Ellipsen vs. voornaamwoorden

In hoofdstuk 5 bespreek ik een aantal mogelijke oorzaken voor het feit dat er bij de productie van ellipsen geen fonologische effecten zijn gevonden, terwijl er bij de productie van voornaamwoorden fonologische inhibitie optreedt. Als eerste beschrijf ik een aantal methodologische verschillen tussen de ellipsis-experimenten en het voornaamwoord- experiment die dit patroon van resultaten zouden kunnen verklaren. Daarna bespreek ik de linguïstische verschillen tussen de twee soorten anaforen (Hankamer & Sag, 1976). Deze linguïstische verschillen komen in figuur 5.1 tot uitdrukking als twee verschillende manieren van lexicale toegang voor de productie van ellipsen en van voornaamwoorden.

CURRICULUM VITAE

Bernadette Schmitt studeerde psychologie aan de Carolo-Wilhelmina Universiteit te Braunschweig, Duitsland, met als specialisatie functieleer. Na het behalen van haar doctoraaldiploma in 1993 werkte zij als wetenschappelijk assistent bij de onderzoeksgroep *Cognitieve Anthropologie* op het Max Planck Instituut voor Psycholinguïstiek te Nijmegen. In september 1993 werd haar een stipendium toegekend door de Max Planck Gesellschaft zur Förderung der Wissenschaften om promotieonderzoek te verrichten binnen de projectgroep *Lexical access in speech production.* Vanaf najaar 1997 is zij als postdoctoral fellow verbonden aan het *Department of Cognitive Science* van de Universiteit van California te San Diego, USA.

ISBN 90-76203-03-2