Ambiguous Pronoun Resolution in L1 and L2 German and Dutch

ISBN: 978-90-76203-37-9
Cover Illustration: Tilman Harpe
Printed and bound by Ponsen & Looijen bv, Wageningen

Ambiguous Pronoun Resolution in L1 and L2 German and Dutch

een wetenschappelijke proeve
op het gebied van de Letteren

PROEFSCHRIFT

te verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann
volgens besluit van het college van decanen
in het openbaar te verdedigen
op vrijdag 7 januari 2011
om 10.30 uur precies

door

Miriam Ellert

geboren op 27 juli 1979
te Bonn, Duitsland

PROMOTOR:	Prof. dr. W. Klein
COPROMOTORES:	Dr. L. Roberts (MPI, Nijmegen)
	Dr. J. Järvikivi (MPI Nijmegen)
MANUSCRIPTCOMMISSIE:	Prof. dr. J. van Hell
	Prof. dr. C. Dimroth (Universität Osnabrück)
	Dr. R. van Gompel (University of Dundee)

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

Acknowledgments

I would like to thank a number of people who supported me in one way or another, and who always believed that this day would come at some point.

Thank you to

- Wolfgang Klein, my promoter, whose knowledge in linguistics as well as various matters of life always impressed and inspired me
- Leah Roberts, my day-to-day supervisor, especially for her theoretical and applied advice on L2 acquisition, and her sense of humor
- Juhani Järvikivi, my second supervisor, who helped me to "finnish" my thesis and who let me see clearer in the visual world
- the members of the Acquisition Group at the MPI who helped me with their helpful suggestions and comments: Aoju Chen, Bhuvana Narasimhan, Christine Dimroth, Giusy Turco, Josje Verhagen, Juhani Järvikivi, Laura de Ruiter, Leah Roberts, Peter Jordens, Sarah Schimke, Wolfgang Klein
- Sarah Schimke, especially for giving a talk in my place when I was ill
- Christine Dimroth, especially for her advice to use comparatives in my experimental stimuli
- Laura de Ruiter, my office mate and friend, and on top of that somebody who shared my addiction for caffeine (unfortunately, the office plant did not)
- Eva Reinisch for being my paranimf and friend, especially for saving my life with her programming skills
- Holger Mitterer for lending me his scripts and Eva
- Sonja Gipper – my paranimf and friend- for dragging me to the sports lessons, especially the "bomming", and who did not stop even when we were home (I will never forget the *politieagenten* ringing at our door saying the volume of the music was fine, if we could just stop the bouncing.)
- Susanne Brower – my very Dutch friend with whom you just do not stay unnoticed. I remember great times in Barcelona and speaking my first sentences of Dutch to you when we needed a secret language
- Nicole Altvater-Mackensen whose life path I seem to cross for the last five years – you gave and still give me lots of input and I am proud to have a "Nijmegen" friend in Göttingen
- Nanjo Bogdanowicz who is one of the greatest organizational talents I have ever met
- Agnes de Boer at the Reception of the MPI for having my day begin with a wonderful smile

- Ad Verbunt for never losing his patience and good humor when I approached him with yet another "technical" question
- Alexander König for helping me to archive my statistical appendix
- Nina Zuurveen, Markus van Ackeren, Johanna Tewes and Kerstin Zimmermann for helping me to prepare and run experiments, and especially for lending me their voices
- Tilman Harpe for drawing probably about 300 pictures so that I could set up visual world displays, and for designing the cover of this book
- Joost Kremers, Josje Verhagen, Susanne Brouwer and Jan-Peter de Ruiter for helping me with the Dutch "Samenvatting" of this thesis
- Steffi Winkler at the VU Amsterdam who was so kind as to distribute my questionnaire to the control group of Dutch natives
- Heidi Rösch at the Educational University Karlsruhe, Albrecht Aichelin at the Goethe-Gymnasium Karlsruhe, and Jochem Reinkens at the Freiherr-vom-Stein-Gynmasium Kleve for making it possible for me to test their students
- Bernt Ahrenholz at the University of Jena and Frances Wilson at the University of Delaware for giving me the permission to reprint their figures and for discussing research matters with me
- the English Linguistics Department at the University of Mannheim: Rosemarie Tracy, Holger Hopp and Dieter Thoma for welcoming me as a guest researcher at their institute
- the German Linguistics Department at the University of Göttingen for welcoming me so open-heartedly and making me feel at home in Göttingen (and for the German version of a pre-proefpromotie): Anke Holler, Annika Herrmann, Annika Hübl, Augustin Speyer, Helena Trompelt, Jana Hosemann, Joost Kremers, Mailin Antomo, Markus Steinbach, Petra Bonitz, and Thomas Weskott

And last but not least I would like to thank my family and friends who always supported me and excused my sometimes turbulent schedule, who accept me for who I am and encouraged me along the way.

- my family, especially my parents and stepparents, Eckard Ellert and Ulrike Kalász-Ellert, Dagmar and Hans Sallge, my siblings and stepsibling Maren and Frank Ellert and Martin Sallge, and my grandmother Gertrud Ellert. *Et es wie et es, et kütt wie et kütt un et hätt noch emmer joot jejange.*
- my friends, especially: Alice Welck, Anne Filorizzo, Henning Vespermann, Jochen Müller, and Julia Breuer. Ihr seid wunderbar.

To all the participants who so patiently lent me their eyes and time, and especially to the L2 learners of this thesis who dealt with the sometimes "Awful Germanic Languages"

Mark Twain

"It is a bleak Day. Hear the Rain, how *he* pours, and the Hail, how *he* rattles [...]. Ah the poor Fishwife, *it* is stuck fast in the Mire [...]. *It* opens *its* Mouth to cry for Help; but if any Sound comes out of *him*, alas *he* is drowned by the raging of the Storm. And now the Tomcat has got one of the Fishes and *she* will surely escape with *him* [...].

There, now, the reader can see for himself that this pronoun-business is a very akward thing for the unaccustomed tongue."[1]

[1] Excerpt taken from *The Awful German Language* by Mark Twain where he retells the *Tale of the Fishwife and Its sad Fate* in English, capitalizing the nouns and marking the gender of the pronouns in the German fashion. (See last page of this thesis for the German translation.)

Contents

Introduction

Chapter 1

Coherent discourse often entails repeated reference to the same discourse entity, and this is frequently achieved by the use of personal pronouns such as *he, she, it.* Unlike in English, German and Dutch have two pronominal forms which can refer to the same singular, masculine entity, as in example (1): the personal pronoun (German *er* and Dutch *hij*) and the d-pronoun[2] (German *der* and Dutch *die*).

(1) German: **Peter**$_i$ wollte Tennis spielen. Doch **er**$_i$/**der**$_i$ war krank.
Dutch: **Peter**$_i$ wilde gaan tennissen. Maar **hij**$_i$/**die**$_i$ was ziek.
English: ***Peter***$_i$ *wanted to play tennis. But* ***he*** *[P*$_i$*/D*$_i$*]* [3] *was sick.*

The two forms do not have exactly the same distribution, however, as can be seen when more than one potential antecedent is available in the discourse (2) (example adopted from Bosch & Umbach, 2007b; p.2). In this case, the personal pronoun is arguably resolved towards the topical entity (*Peter*), while the d-pronoun prefers the non-topical entity (*Paul*) (Bosch, Rozario, & Zhao, 2003; Comrie, 1994; Diessel, 1999; Lambrecht, 1994); or, as some researchers have claimed (at least for German), the d-pronoun is marked for non-topical reference, whereas the personal pronoun is neutral in this regard (Ahrenholz, 2007; Bosch & Umbach, 2007b; Zifonun, Hoffmann, Strecker, & al., 1997).

(2) German: **Peter**$_i$ wollte mit **Paul**$_j$ Tennis spielen. Doch **er**$_i$/**der**$_j$ war krank.
Dutch: **Peter**$_i$ wilde met **Paul**$_j$ gaan tennissen. Maar **hij**$_i$/**die**$_j$ was ziek.
English: ***Peter***$_i$ *wanted to play tennis with* ***Paul***$_j$*. But* ***he*** *[P*$_i$*/D*$_j$*] was sick.*

Therefore, the two pronominal forms in German and Dutch can refer to the same entity in certain contexts (1), but at the same time, they differ in anaphoric function in a way that is not yet fully understood, as Bosch et al. (2007b) state, commenting on examples like (2) above:

> "even though the difference between the interpretations of the Dutch and German personal and demonstrative pronouns [d-pronouns] is clear in the case at hand, it has proven difficult to get a clear picture of what exactly the difference in the constraints on the interpretation of demonstrative [d-pronouns] and personal

[2] Following Ahrenholz's (2007) and Van Kampen's and Pinto's (2007) terminology, we will call these forms *d-pronouns* to highlight their different use compared to demonstratives. A more detailed discussion will be provided in chapter 2.

[3] P = personal pronoun; D = d-pronoun

pronouns is (p.2)."

The overall aim of this thesis is to get a clearer picture of the different functions of personal and d-pronouns in German and Dutch. The constraints which guide the interpretation of the two pronominal forms will be investigated in more detail by studying the resolution of the forms in ambiguous contexts, i.e. when two potentially matching antecedent candidates are available to co-refer with the pronoun, as in (2). Is it the case that in these ambiguous contexts the two forms are resolved in a complementary fashion, i.e. that the personal pronoun preferentially resolves towards the topical entity *Peter* while the d-pronoun is resolved towards the non-topical entity *Paul* (Bosch, et al., 2003; Comrie, 1994; Lambrecht, 1994)? Or is the difference in function between the two pronouns more complex, going beyond such a simple asymmetric distribution. For example, perhaps the d-pronoun takes over a disambiguating function in that it is resolved towards the non-topical entity *Paul*, while the personal pronoun, despite a certain preference for the topical entity *Peter*, may be more neutral in this regard (Ahrenholz, 2007; Bosch & Umbach, 2007b; Zifonun, Hoffmann, Strecker, & al., 1997)?

Previous psycholinguistic studies which have investigated the resolution preferences of personal and d-pronouns in German or Dutch (Bittner, 2007; Bosch, Katz, & Umbach, 2007a; Bosch, et al., 2003; Bosch & Umbach, 2007b; Kaiser & Trueswell, 2004; Wilson, 2009) have proposed to differentiate between the functions of the two pronominal forms in terms of a number of different phenomena: order of mention (first- vs. second-mentioned), grammatical role (subject vs. object), case marking (nominative vs. non-nominative), and topicality (topic vs. non-topic). While some have found that the two forms work in relatively complementary distribution, others report an asymmetry in function. In sum, although we have an idea of the constraints which may contribute to the interpretation preferences of the two pronominal forms, there is still no clear picture. This might be due to the fact that the studies differ with regard to the following criteria:

- Tasks:
    - production vs. comprehension (e.g. corpus analyses vs. self-paced reading)
    - on- versus off-line interpretation (e.g. eye-tracking vs. completion tasks)
    - written vs. spoken language
- Materials:
    - canonical vs. non-canonical antecedent structures (e.g. SVO vs. OVS structures)
    - pronoun appears sentence-initially vs. not (e.g. *He was sick* vs. *But he was sick*)

The experiments reported in this thesis focus on the interpretation of personal and d-pronouns in discourse comprehension. They were designed to measure the on- and off-line resolution pattern for the two pronouns in spoken and in written language comprehension. While one set of experiments investigates these resolution patterns for pronouns following canonical antecedent structures, another set of experiments presented the pronoun after non-canonical antecedent structures. The pronouns always appeared sentence-initially. This set-up aimed at controlling as far as possible for the above-mentioned variables, while at the same time, ensuring maximal comparability to the results of previous studies. Furthermore, to investigate whether the functions of the two pronominal forms generalize cross-linguistically, this thesis presents experiments conducted in both German and Dutch, whereas most previous studies have mainly addressed these questions focusing on German (Bittner, 2007; Bosch, et al., 2007a; Bosch, et al., 2003; Bosch & Umbach, 2007b; Wilson, 2009), or Dutch alone (Kaiser & Trueswell, 2004).

Many psycholinguistic theories of reference predict that there should be complimentary distribution in the interpretation of personal and d-pronouns (Ariel, 1990, 2001; Gundel, 2003; Gundel, Hedberg, & Zacharski, 1993; Levinson, 1987, 1991). These theories state that the most reduced referring expressions, in terms of lexical or prosodic weight (in this case the personal pronoun), should resolve towards the most highly accessible, or cognitively salient, referent in the mind of the speaker/hearer. This raises the question of what determines the relative accessibility of one potential antecedent over another, a question that has been the subject of debate in many psycholinguistic studies on pronoun resolution. In particular, is it being the subject of the sentence (Crawley, Stevenson, & Kleinman, 1990; Frederiksen, 1981) or is it being the first-mentioned antecedent that makes a referent more accessible (Gernsbacher, 1989)? As this is difficult to disentangle in English, because the first-mentioned entity is usually also the syntactic subject, researchers have turned to flexible word-order languages such as German and Finnish. By investigating SVO and OVS antecedent structures, they have attempted to identify the effects of order of mention and grammatical role (Bosch, et al., 2007a; Bosch & Umbach, 2007b; Bouma & Hopp, 2007; Crawley, et al., 1990; Järvikivi, Van Gompel, Hyönä, & Bertram, 2005; Kaiser & Trueswell, 2008; Wilson, 2009). However, the results are inconsistent, perhaps because of the possible influence of the different word orders on information structure.

To address the possible confound of grammatical role and order of mention, the experiments in this thesis were constructed to investigate the relative influence of the order of mention of the antecedent candidates by controlling for grammatical role and the information structure of the antecedent clause. This was achieved by using comparative double-nominative antecedent structures, such as in (3). Using this type of construction

allows us to disentangle order of mention from grammatical role information, since grammatical information cannot be used for disambiguation purposes, given that both nouns are marked for nominative case. Information structure is operationalized in the current set of studies by manipulating the word order of the antecedent clause. The antecedent clause was presented in non-canonical order (4) to investigate the potential effect of this change of information structure on the resolution preferences of the two pronouns.

(3)	German:	**Der Schrank** ist schwerer als **der Tisch.** ***Er/Der*** stammt aus einem Möbelgeschäft in Belgien…
	Dutch:	**De kast** is zwaarder dan **de tafel.** ***Hij/Die*** is afkomstig uit een meubelwinkel in België...
	English:	***The cupboard*** *is heavier than* ***the table.*** ***It*** *[P/D] originates from a furniture store in Belgium...*
(4)	German:	Schwerer als **der Tisch** ist **der Schrank.** ***Er/Der*** stammt aus einem Möbelgeschäft in Belgien…
	Dutch:	Zwaarder dan **de tafel** is **de kast.** ***Hij/Die*** is afkomstig uit een meubelwinkel in België...
	English:	*Heavier than* ***the table*** *is* ***the cupboard.*** ***It*** *[P/D] originates from a furniture store in Belgium...*

In contrast to English where the personal pronoun *he* refers to masculine animate entities only, the German and Dutch masculine personal and d-pronouns[4] refer to animate (biological gender) as well as inanimate entities (grammatical gender). Previous studies have controlled for animacy effects (most often, with all antecedents being animate), but in the current thesis, as well as order of mention and information structure, the effect of animacy on pronoun resolution was also explicitly tested (see (3) and (4)).

When investigating the above-mentioned questions, some methodological issues arise. For example, how much inter-individual variability in resolving personal and d-pronouns is there between participants? Do all participants show the same resolution preferences or are there individual participant specific preferences? To what extent are the findings obtained by different (off- and on-line) tasks comparable? In this dissertation, all experiments comprised an on-line visual-world eye-tracking task and an off-line forced-choice questionnaire task. The resolution similarities and differences across participants and tasks

[4] The Dutch d-pronoun *die* actually encodes common gender, i.e. masculine and feminine gender.

will be addressed in more detail in chapter 3.4.

Taken together, the overall aim of this dissertation is to get a better picture of the different functions of personal and d-pronouns by investigating how they are resolved in discourse by native adult speakers of German and Dutch. More specifically, the first experimental chapter of this thesis (chapter 3) addresses the following research questions:

1. How does the *order of mention* of the antecedent candidates influence the resolution of personal and d-pronouns? (chapter 3.1)
2. How does the *information structure* of the antecedent clause influence the resolution of personal and d-pronouns, i.e. how do non-canonical antecedent structures affect the resolution preferences of personal and d-pronouns? (chapter 3.2)
3. How does the inherent semantic factor *animacy* influence the resolution of personal and d-pronouns? (chapter 3.3)
4. Do pronoun resolution preferences vary inter-individually across participants or tasks? (chapter 3.4)
5. Is there cross-linguistic evidence for general resolution preferences for personal and d-pronouns? (chapters 3.1, 3.2, 3.3, 3.4)

The second experimental chapter (chapter 4) deals with the question of how Dutch learners of German and German learners of Dutch resolve personal and d-pronouns. As outlined above, although the exact constraints which operate on the distinct resolution behaviors of personal and d-pronouns still needs to be investigated in more detail, there seems to be consensus on the fact that they occupy two different functions namely, in that the personal pronoun is preferentially resolved towards the topical and the d-pronoun towards the non-topical entity. Critically, this is assumed to be the case in both German and Dutch and so the study of how second language (L2) learners resolve personal and d-pronouns in their L2 can give us more insight into whether the processing strategies reflected in one's native language (L1) can be applied to a second language. In other words, it is assumed that the learner groups under investigation could highly benefit from positive L1 transfer in this domain. On the other hand, previous studies on the resolution of personal and d-pronouns in L2 German by English and Finnish learners have reported non-native-like resolution behaviors (Roberts, Järvikivi, Ellert, & Schumacher, in prog.; Wilson, 2009). Thus, it is possible that L2 pronoun resolution might in general be different from L1 pronoun resolution. Moreover, since there are only two on-line studies on the L2 resolution of personal and d-pronouns so far (both on L2 German) more empirical evidence is needed to draw any general conclusions. Furthermore, although the question of inter-individual variability with regard to pronoun resolution preferences is important for all

comprehenders, it may be even more so for L2 learners, whose performance is notoriously more variable than that of native speakers, often because of differences in general proficiency. This issue will be addressed in more detail in chapter 4.4.

The second part of this thesis aims to better understand how L2 learners resolve personal and d-pronouns in discourse, taking into account the typological similarity of the L1-L2 language pairings German-Dutch and Dutch-German. Specifically, it will be asked if the learners are able to differentiate between the functions of the two pronominal forms. In addition to the above outlined research questions, the following acquisition questions will be addressed:

6. How do the factors *order of mention, information structure of the antecedent clause* and *animacy* influence L2 pronoun resolution? (chapters 4.1, 4.2, 4.3)
7. Do L2 pronoun resolution preferences vary inter-individually across participants according to differences in proficiency? (chapter 4.4)
8. Do L2 learners show the same resolution preferences for personal and d-pronouns as native speakers? (And if not, in what ways do they differ?) (chapter 4)
9. Do Dutch L2 learners of German and German L2 learners of Dutch resolve personal and d-pronouns in their second languages in the same way? (chapter 4)

The thesis is organized as follows. Chapter 2 provides the theoretical background for the dissertation. I discuss theoretical concepts of discourse and information structure in relation to pronominal reference and cognitive and pragmatic theories of reference. I give an overview of the German and Dutch pronominal systems and review previous psycholinguistic work on pronominal reference. In chapter 3, I present four experiments on the resolution of personal and d-pronouns in native German and Dutch, and provide additional analyses on the effect of animacy and inter-individual variability. Chapter 4 reports on four experiments on the resolution of personal and d-pronouns by Dutch L2 learners of German and German L2 learners of Dutch, and provides additional analyses on the influence of animacy on L2 pronoun resolution and L2 proficiency effects. In chapter 5, I summarize the findings, answer the above outlined research questions and discuss future work.

Background

Chapter 2

2.1 Discourse and Information Structure

When two interlocutors engage in a conversation, the information that is represented in the speaker's as well as in the addressee's mind forms their common ground. The common ground changes dynamically as the conversation moves on: it is gradually updated across utterances as new pieces of information are conveyed. In addition to the content of the utterance, the packaging of the information (Chafe, 1976), its "information structure", optimizes the updating of the common ground in that it provides some extra clues about how it should be understood. As can be seen in example (5) (adapted from Krifka & Féry, 2008), while the lexical content and the truth conditions are the same, only the answer in (5)a is an acceptable answer to the question.

(5) What did you see on the road?
   a. I saw a TIGER on the road.
   b. # I saw a tiger on the ROAD.

In (5), both answers differ with respect to which constituent carries the pitch accent, i.e. which constituent is in *focus*. The focus of a sentence is the part of the sentence which asserts new information to an open proposition such as *I saw X on the road* (Lambrecht, 1994). Some definitions of focus entail that the focused constituent is a member of a set of possible alternatives (Krifka & Féry, 2008; Rooth, 1992). While the example in (5)a implies that the *tiger* was seen out of a set of animals, in sentence (5)b, the tiger—which could have been seen in a variety of locations—was seen on the *road*. Besides the information that the speaker *had seen a tiger on the road*, by means of accentuation *the tiger* is signaled to be the new piece of information, thus the addressee is to create a mental representation of *a tiger* in the discourse (see also *Identifiability* in the next section).

Another pragmatic distinction is that of the *topic* and the *comment* of an utterance. The *topic* is the element of the sentence that the sentence is about (Lambrecht, 1994; Reinhart, 1982)[5]. To clarify the relationship between a topic and an utterance, Reinhart (1982) uses the metaphor of a file card system. In this imaginary file card system, a sentence such as

[5] The notion of *topic* has wide-spread use. I will adopt the view which limits the topic to the sentence or clause level (Lambrecht, 1994; Reinhart, 1982), as opposed to operating beyond the sentence level as e.g. in Givón's (1983) terminology of *topic chains* and *discourse topic*.

(6)a is perceived as an utterance about the topic *Ulrike's new sunhat*. The topic constitutes the heading of the file card which contains the entry *Resembles the flowerpot on the table.* This entry is the *comment* of the utterance; thus the part which asserts information about the topic. But a non-canonical version of the sentence as in (6)b would classify the comment as information about the *flowerpot on the table,* which by means of sentence inversion is the topic here. The element of a sentence which constitutes the *comment* is often the same element which qualifies as the *focus* of the sentence. Whether researchers differentiate between *topic-comment* or *topic-focus* structures depends on their information structural definitions[6]. The *focus* definition often entails that the element is one from a set of alternatives which needs not be the case for the *comment*.

(6)

a. Ulrikes neuer Sommerhut ähnelt dem Blumentopf auf dem Tisch.
   Ulrike's new sunhat$_{NOM}$ *resembles the flowerpot*$_{DAT}$ *on the table.*
b. Dem Blumentopf auf dem Tisch ähnelt Ulrikes neuer Sommerhut.
   The flowerpot$_{DAT}$ *on the table resembles Ulrike's new sunhat*$_{NOM}$.

Example from Musan (2010; p.27)

Information structural notions, such as *topic* and *focus*, are problematic. It is often difficult to determine what is the topic and what is the focus of an utterance. One way that has been used to identify them in a sentence is to form question-answer pairs. The topic is then the part of the answer which also forms part of the question. Thus, in (7)a *Peter played* is the topic, while in (7)b it is *played tennis*. The elements which form the topic thus depend on the context. The elements in focus are the elements which answer the question.

(7)

a. What did Peter play?
   [Peter played]$_{TOP}$ [tennis]$_{FOC}$.
b. Who played tennis?
   [Peter]$_{FOC}$ [played tennis]$_{TOP}$.

The distinction between topic and focus is sometimes made in terms of *givennenss* and *newness* of the information. While the topic represents the part of the sentence information which is given, the focus is the part which contains or entails the new information.

[6] Note that in information structure, there is no terminology that linguists agree on; rather parallel terms have developed which are used to denote the same or resembling criteria (see Musan, 2010 for an overview in German). From here on, I will concentrate on the terms topic and focus.

However, this distinction does not fully capture the nature of topics; there are also topics in focus, i.e. *contrastive topics* as *HANS* in (8).

(8) Did Peter want to play tennis with Paul?
- No, Peter wanted to play tennis with HANS.

The above examples have outlined the two pragmatic functions of topic and focus. We have seen that intonation can signal the difference between the two. Like intonation, referring expressions and word order are devices marking information structure. They will be outlined in the next sections.

2.1.1 The Information Status of a Referent: Identifiability and Activation

Before going into the subtle differences between personal and d-pronouns, I will explain how the use of these pronouns can be differentiated from the use of other types of referring expressions such as full NPs.

When a speaker makes reference to a certain entity in discourse, she can do so with a variety of devices, such as indefinite and definite lexical NPs, proper names, d-pronouns, personal pronouns and zero pronouns. The choice of one form over another is not arbitrary, but rather depends on what the speaker assumes the hearer to know about this entity and how available its mental representation may be in the current discourse. In this section, I present two information-structure categories by which Lambrecht (1994) classifies the preference for one form over another: *Identifiability* and *Activation*. Note, that both concepts elaborate on the mental representation of a referent, rather than the referent itself.

When a speaker assumes[7] that an entity is identifiable to the hearer, she assumes that the hearer has already stored a mental representation of that entity. Thus, at the time of the utterance, a representation exists in the speaker's as well as in the addressee's mind. An entity is considered to be unidentifiable, when it is not yet represented in the addressee's mind. Thus, at the time of the utterance a representation is only available in the speaker's mind, and the addressee needs to create such a representation. This is done via a linguistic[8] description which subsequently allows for anaphoric reference. In languages which mark definiteness e.g. by the use of indefinite and definite determiners (such as English, German and Dutch), the two grammatical distinctions definite-indefinite mostly overlap with the two cognitive distinctions identifiable – unidentifiable. Thus, in (9)a (Lambrecht, 1994:

[7] Note, that the word "assume" does not necessarily refer to conscious processes.
[8] Note, that also non-linguistic means can be used to achieve identifiability. However, as this thesis is on anaphora resolution, I focus on the linguistic means.

p.89), the speaker uses an indefinite NP to signal that the addressee has not yet created a representation of *the meeting* which the speaker needs to attend, since it was not mentioned in the previous discourse. Once mentioned, the addressee can build a mental representation of it, and from now on, it is identifiable, and must therefore be referred to with a definite NP or a pronoun, as in (9)b[9]. Sentence (9)c would be unacceptable under this interpretation.

(9)

a. I'm going to **a meeting** tonight.
b. How long is **the meeting/your meeting/it** supposed to last?
c. # How long is **a meeting** supposed to last?

This classification as being identifiable persists throughout the discourse, and even from one discourse to another, except for when the speaker assumes that the addressee has forgotten the existence of the referent. Thus, it entails a binary distinction: either the addressee knows the referent or she does not. For a pronoun to be used, the mental representation of its referent always needs to be identifiable.

A discourse entity can be further classified by the second information-structural category, namely *Activation.* This cognitive criterion is a graded distinction between the current cognitive activation of the referent in the speaker's and addressee's memory. Lambrecht relies mostly on Chafe's distinction (1974, 1976), according to which a referent can either be *active, semi-active/accessible,* or *inactive.* While active referents are in the "current focus of consciousness" in a person, semi-active referents are in the "peripheral consciousness" (one has a "background awareness" of them). Inactive referents are located in the long-term memory. They can be activated after long time periods, while active referents keep their information status only as long as they are among *the current topics of discussion.* An identifiable referent is thus always either active, or accessible, or inactive. Lambrecht (1994; p.105) mentions the following formal properties of referential expressions to distinguish between these different cognitive states of the representations of referents: *presence or absence of an accent, pronominal vs. lexical coding; and in some languages definite vs. indefinite marking.*

Typically, an active referent is encoded by an unaccented expression. In terms of markedness, Lambrecht states that unaccented expressions are marked and accented expressions are unmarked for activeness. An active referent can be encoded by a lexical NP as well as by a pronoun. However, a pronoun always has an active referent. While pronouns

[9] There are exceptions, in particular non-specific and generic NPs (see Lambrecht, 1994; p.80-85).

are marked for activeness, lexical NPs are neutral in this regard, thus they are unmarked. An inactive referent is typically coded as a definite lexical noun phrase and it is prosodically relatively prominent. Chafe understands the concept of activation as a continuum, rather than a binary distinction. This is represented by the fact that he describes an intermediate category *accessible referents*, between the two extremes of active and inactive. While active and inactive referents have phonological and morphological correlates, the coding of accessible referents may sometimes overlap with active and sometimes with inactive codings. A referent typically becomes accessible by means of an active referent being deactivated during discourse, or due to belonging to a schema which has been activated.

With regard to different types of pronouns, such as German and Dutch personal and d-pronouns, both information-structure categories *Identifiability* and *Activation* would predict their referents to be activated, which entails them to be also identifiable. As Lambrecht (1994; p.96) puts it:

> "The formal category 'pronoun' is no doubt the best evidence for the grammatical reality of the information-structure category 'activeness'."

So far, we have seen how the use of pronouns can be differentiated from the use of other types of referring expressions, in that their referents are always active and identifiable. In the next section, I will outline Lambrecht's general rule (1994) which differentiates between the use of personal and d-pronouns on the basis of the topic-focus parameter.

2.1.2. Pronouns Referring to Sentence Topics

The categories of topic and focus indicate relations between referents and propositions and, according to Lambrecht's terminology, act on the sentence level. Consider the following two examples, where in (10), the pronouns *her* and *she* referring to *Pat* form a topic relation with their propositions. The information, namely *X was called twice*, is provided about *she* (= *Pat*), and this is the case for both sentence structures (10)a and (10)b. In (11), however, the pronouns *HER* and *SHE* stand in a focus relation to their propositions. Here, a relevant new piece of information is added to the open proposition *they called X*, namely that it was *Pat (HER, SHE)* who was called. Although in both examples the pronouns can be said to denote an identifiable and active referent, the difference between them is the relation towards their propositions, namely being in a topic or in a focus relation.

(10) *Topic*
Has Pat been called yet?
a. - Pat said they called her TWICE.
b. - Pat said she was called TWICE.

(11) *Focus*
Who did they call?
a. - Pat said SHE was called.
b. - Pat said they called HER.

Examples from Gundel (1980); in Lambrecht (1994; p.115)

As Lambrecht (1994; p.204) points out for German, these categories are of special interest when differentiating between personal and d-pronouns, the latter of which he sees as a second set of personal pronouns. "As a general rule, pronouns of the *er*-series [personal pronouns] are used when a referent is active AND already topical, while those of the *der*-series [d-pronouns] are used when a referent is active but not yet an established topic." Example (12) shows that the use of a d-pronoun in the main clause makes a co-referential reading towards the subject of the subordinate clause impossible; indicating that only the personal pronoun may be used to refer to the already topical entity.

(12) Wenn *er/der* isst, macht *er/*der* so komische Geräusche.
*When he [P/D] eats, he [P/*D] makes kind of funny noises.*

Example from Lambrecht (1994; p.204)

However, example (13) taken from Klein (2008) is a typical topic drop example. In German, topic drop is possible when the topic appears before the finite verb. This example shows that it is not a personal pronoun which is dropped, but rather a d-pronoun; thus, there are at least some exceptions to Lambrecht's general rule.

(13) (Den) kenne ich nicht.
(Him [D]), I don't know.

In the empirical part of this thesis, Lambrecht's general rule will be tested experimentally. More specifically, I address the question whether listeners indeed resolve the personal pronoun towards the topical and the d-pronoun towards the non-topical entity. Furthermore, the question of how listeners resolve both pronouns when the non-topical entity is in focus

will be addressed. This different information status will be operationalized by changing the word order of the antecedent sentences from canonical to non-canonical. The influence of word order changes on information structure will be outlined in the next section.

2.1.3. Word Order and Information Structure

Word order is yet another means to signal information structure like referring expressions or intonation. As Klein (2008) has pointed out there exists a general tendency across languages to place *topical elements*[10], or in his words *Topic-Situation-Identification*-elements (TSI), first. However, he further explained that the topic-first principle can be overruled if explicitly marked as in (14) where *HANS*, although being placed first, is in focus, as marked by accentuation.

(14) Who called?
- HANS_{FOC} called.

Furthermore, the topical part of the utterance can include more than one constituent as example (15) shows.

(15) What did Peter and Hans do?
- $[\text{Peter and Hans/They}]_{\text{TOP}}$ played tennis.

Thus, unless marked otherwise, the word order of the sentence usually encodes topical elements first. Non-canonical word orders in English, allows one to place elements other than the syntactic subject in the sentence-initial position. Among the English word orders which present their constituents in a non-canonical order are passivizations and inversions. Ward and Birner (2001) studied these two word orders and their functions, and concluded that by inversing the canonical order of two constituents, a pragmatic constraint is posed upon them. The preposed element must not transfer newer information than the postposed element. Ward and Birner thus see this reordering of the constituents as a grammatical possibility to order the constituents in terms of discourse-familiarity. They found that the preposed element might represent discourse-old information (previously mentioned or inferable information) and the postposed element discourse-new information. However, it could also be the case that both elements are discourse-old. In that case, they can be differentiated in terms of recency of mention, in that the preposed element has been more recently mentioned than the postposed element. According to Ward and Birner, to have been evoked, inferred or recently mentioned defines the degree of discourse-familiarity.

[10] Klein (2008) actually calls the topical element the *Topic Situation Identification* element. He distinguishes between different topic situations depending on the truth value which might differ across time, space etc.

Since they investigated discourse, recency of mention can be understood as topic continuity. Example (16) illustrates that the subject-entity of the passive clause should not be less familiar than the *by*-element (Ward & Birner, 2001; p.28-29).

(16)

a. The mayor's present term of office expires Jan. 1. *He* will be succeeded by *Ivan Alleen Jr.*
b. Ivan Alleen Jr. will take office Jan. 1. # *The mayor* will be succeeded by *him.*

This also applies for inverted sentences, as in example (17) from Ward and Birner (2001; p.25). In this example the preposed *also complimentary* is discourse-old while *red and white wine* is discourse-new. Thus, *red and white wine* constitutes the focus. *Also complimentary* qualifies as the sentence topic (situation). However, it implies contrastiveness in that the speaker could go on explaining what is *not complimentary*.

(17) We have complimentary soft drinks, coffee, Sanka, tea, and milk. *Also complimentary is red and white wine.* We have cocktails available for $2.00.

Flexible word order languages, such as German and Dutch make much use of non-canonical OVS structures. In such languages, the resolution of personal and d-pronouns has been studied after SVO and OVS sentence structures to measure the relative contribution of grammatical role cues and order of mention of the antecedent candidates. Example (18) is a German SVO-OVS example. Both answers (18)a and (18)b are equivalent to the English sentence *Yes, the doctor saw the cook.* With regard to Klein's (2008) topic identification principles, the first constituent represents the topic of the sentence, unless marked otherwise. However, sentence (18)b evokes the impression that *the cook* was seen, as opposed to somebody else, while in (18)a (if *the cook* is not prosodically marked) such an impression is not created. Thus, although both sentences have the same lexical content, they differ with regard to the contrastive impression they create.

(18) Hat der Arzt den Koch gesehen?
Did the doctor see the cook?

a. Ja, der Arzt hat den Koch gesehen.
Yes, the doctor$_{\text{NOM}}$ *has seen the cook*$_{\text{ACC}}$.
b. Ja, den Koch hat der Arzt gesehen.
Yes, the cook$_{\text{ACC}}$ *has seen the doctor*$_{\text{NOM}}$.

Example (19) is an English example which corresponds to the German and Dutch

experimental materials used in this thesis. The example is similar to the above mentioned example. While sentence (19)a represents a canonical comparative structure, sentence (19)b is an inversion of it, thus representing a non-canonical structure. It is quite similar to Ward and Birners (2001) inversion example of *Also complimentary are red and white wine* and its use is to topicalize the first part of the sentence, namely *Heavier than the table is X*. This "topic situation" (following Klein, 2008) creates the impression that there is something else that is lighter or is as heavy as the table. Interestingly, *the cupboard* which in (19)a has a topical function, has the function of focus in (19)b. This distinction and its influence on the resolution of personal and d-pronouns will be experimentally addressed in this dissertation in chapters 3 and 4.

(19)

a. [The cupboard]$_{TOP}$ is heavier than the table.
b. [Heavier than the table is]$_{TOP}$ the cupboard.

In the following section, I outline some theories of reference which have been explicitly concerned with explaining the relationship between referential forms and their functions in discourse, and explained from different perspectives, namely from a pragmatic perspective and a cognitive perspective.

2.2 Theories of Reference

Theories of Reference propose to anchor anaphora within a more general theory of communication. While the first account to be presented in this section explains the use of anaphoric expressions from a pragmatic perspective, the cognitive approaches to be presented afterwards are concerned with the mapping between the activation of a mental representation of a discourse referent and the type of referential expression used to denote it. Although the accounts do not explicitly address the issue of personal and d-pronouns, they make general predictions about the use of different referring expressions and their functions. In this section, both approaches will be reviewed, and special attention will be given to the following question:

- What predictions can be derived for the resolution of personal and d-pronouns from theories of reference?

2.2.1. Levinson's Pragmatic Approach

Levinson (1987, 1991) developed a pragmatic theory of anaphoric reference based on Gricean principles, i.e. the distribution of anaphoric expressions follows some general

pragmatic strategies: Levinson's Q-, I-, and M-principles.

- The Q-principle
  Do not say less than is required.
- The I-principle
  Do not say more than is required.
- The M-principle
  Do not use a marked expression without reason.
- The interaction
  Level of genus: Q > M > I

Whenever a referential form is chosen there is a trade-off between the Q-and the I-principle. For example, when introducing a referent into the discourse, following the Q-principle a full NP or a proper name with a long description is preferred; however, according to the I-principle a relatively light expression is favored, i.e. a pronoun or a zero pronoun. Since both principles are operating, a trade-off takes place, and a sufficient yet recognizable form, such as a full NP or a proper name, is chosen.

With regard to anaphora, the Q-, I- and M-principles follow an order of application (the interaction) where Q-implicatures take precedence over I-implicatures, i.e. where the use of a reflexive is permitted for a direct encoding of co-referentiality, the application of a less full form, such as a personal pronoun, will generate a non-co-referential interpretation as in (20).

(20) Sally$_i$ accidently called herself$_i$/her$_j$.

According to the I-principle, minimally informative referential forms, such as personal pronouns, will favor co-referential interpretations, producing stereotypical interpretations as in (21).

(21) Did Hans$_i$ call? – He$_i$ called.

According to the M-principle, if a marked form, such as an overt pronoun, is used where an unmarked form, such as a zero pronoun, could have been used, a non-co-referential interpretation will be generated as in (22).

(22) Llamó Hans$_i$? – Ø$_i$/él$_j$ llamó.
Did Hans$_i$ call? -Ø$_i$/he$_j$ called.

However, the principles above mentioned do not only apply to anaphora, but also to a more general discourse level, such that the I-principle, *Do not say more than is required*, also instructs the addressee to favor co-reference interpretations over non-co-reference interpretations. In other words, minimizing the number of discourse entities referred to enriches the informativity of the speaker's statements. Thus, both brevity and semantic content are taken as criteria of minimization. Furthermore, the addressee's default is to prefer reference maintenance interpretations over interpretations of referential shift.

With regard to anaphora, the *semantic content hierarchy* predicts that the forms to the left prefer locally non-co-referential interpretations, while the forms to the right prefer locally co-referential interpretations. Thus, the more minimal a form, the more it prefers a co-referential interpretation.

The semantic content hierarchy
lexical NP > pronoun > Ø

In cases where a zero pronoun is grammatically allowed, and an overt pronoun is used instead, Levinson argues that the overt pronoun will favor a disjoint interpretation. Thus, a pronoun is not acceptable under a co-reference reading, but it is acceptable under a disjoint reading.

Levinson does not make direct predictions with regard to personal and d-pronouns. According to the semantic content hierarchy, the two types of pronouns carry an equal amount of semantic content, in that they encode less semantic content than a lexical NP, but more than a zero pronoun. Taking form brevity as a minimization criterion, the German d-pronoun can be classified as a more marked form than the personal pronoun, leading for a speaker preference to use a personal rather than a d-pronoun in example (23) (preference appears in bold letters).

(23) Der Arzt$_i$ wollte Tennis spielen. Doch ***er*** */der*$_i$ war krank.
The doctor$_i$ *wanted to play tennis. But he* ***[P]*** */[D]*$_i$ *was sick.*

However, this would also imply that when a d-pronoun is used in this context, it would not be interpreted co-referentially, but with a disjoint reading, leading to a non-co-reference interpretation where *der* does not refer to *the doctor*. However, as will be treated in more detail in chapter 2.3 (e.g. Ahrenholz, 2007; Bosch, Katz, & Umbach, 2007a; Bosch, et al., 2003; Bosch & Umbach, 2007b), a co-reference interpretation is perfectly possible with the

d-pronoun. Thus, it is not clear how to differentiate between both forms[11].

A further problem arises when we take the two pronominal Dutch forms into consideration. Both minimizing criteria, semantic content and form brevity, would predict the personal pronoun to be more marked than the d-pronoun, since the personal pronoun differentiates between three types of gender, while the d-pronoun only distinguishes between two types of gender. Thus, for Dutch we would obtain the exact opposite predictions than for German.

With regard to co-referential and non-co-referential interpretations, the question arises as to what happens in an ambiguous context where two potential antecedents are provided for the anaphor. Since the M-principle states that when a marked form is used instead of an unmarked form, a non-co-referential interpretation is generated, this would mean that if a d-pronoun was used instead of a personal pronoun, a disjoint reading would be obtained. With regard to the I-principle stating that referential maintenance interpretations are preferred over referential shift interpretations, in example (24) the unmarked pronoun would be interpreted as co-referential with the topical first-mentioned entity, while the marked pronoun would initiate a disjoint interpretation, leading to a non-topical second mentioned antecedent preference[12].

(24) **Der Arzt**$_i$ wollte mit dem Koch$_j$ Tennis spielen. Doch ***er***$_i$ war krank.
The doctor$_i$ *wanted to play tennis with the cook*$_j$. *But* ***he*** *[P]*$_i$ *was sick.*

Der Arzt$_i$ wollte mit **dem Koch**$_j$ Tennis spielen. Doch ***der***$_j$ war krank.
The doctor$_i$ *wanted to play tennis with* ***the cook***$_j$. *But* ***he*** *[D]*$_j$ *was sick.*

Turning to a non-canonical sentence structure containing two potential antecedents as in (25), the question arises as to which form would be preferentially resolved towards which entity? According to Levinson's principles, a preferred reference-maintaining co-referential interpretation of the personal pronoun would be made towards the second-mentioned antecedent in focus position. The addressee assumes that the discourse will elaborate on that second entity. However, when a d-pronoun is used, the disjoint reading will be

[11] Under the assumption that the d-pronoun reading, although co-referential, introduces a contrast or a shift from the stereotypical semantic development of the discourse rather than a referential shift, this view can be saved. However, in the case of two potential antecedents it is not clear what to predict for the d-pronoun: co-reference and unstereotypical semantic uptake in discourse, or disjoint referential interpretation.

[12] We understand the *disjoint interpretation* after sentence structures presenting two potential antecedents as co-referential with the non-prominent entity, thus here with the non-topical entity. However, it could also be understood as referring to a discourse-external entity.

preferred, thus it should be resolved towards the first-mentioned entity.

(25) Freundlicher als der Koch$_i$ ist **der Arzt**$_j$. **Er**$_j$ ist heute Morgen recht früh aufgestanden.
*Friendlier than the cook$_i$, is **the doctor**$_j$. **He** [P]$_j$ has gotten up quite early this morning.*

Freundlicher als **der Koch**$_i$ ist der Arzt$_j$. **Der**$_i$ ist heute Morgen recht früh aufgestanden.
*Friendlier than **the cook**$_i$, is the doctor$_j$. **He** [D]$_i$ has gotten up quite early this morning.*

Note, that the application of the above-mentioned principles leads to complementary predictions for personal and d-pronouns: either co-referential or disjoint. This is different from the Cognitive Approach which predicts gradual rather than complementary behavior (as memory nodes receive more or less activation), as we will see in the next section.

2.2.2. Ariel's Cognitive Approach

Cognitive Theories of Reference commonly postulate a scalar relationship between the form of the referring expression and the cognitive status of its referent. More specifically, based on Givón (1983), all theories rely on the basic assumption that more reduced forms (e.g. personal pronouns) refer to more salient (or accessible) antecedents, and less reduced forms (e.g. full noun phrases) refer to less salient (or accessible) antecedents. The first account to be reviewed here is Ariel's Accessibility Theory (1990, 2001). Then Gundel et al.'s Givenness Hierarchy (2003; 1993) will be presented.

Ariel's Accessibility Theory (1990, 2001) has been developed from Chafe's Activation States (1976) and Givón's (1983) reversed mapping principle (fullness of forms depends on amount of antecedent salience). In the center of this theory lies the notion of accessibility which refers to the current level of activation of a referent in the mind of the addressee. In Ariel's words, "The [the] basic idea is that referring expressions instruct the addressee to retrieve a certain piece of Given information from his memory by indicating to him how accessible this piece of information is to him at the current stage of the discourse (Ariel, 2001; p.29)". By indicating its level of activation, **perceptual** information is encoded on the referring expression. Yet, referring expressions also convey **conceptual** information (lexical information), in that *he* not only means "highly accessible", but also "masculine and singular". The degree of conceptual information that different types of pronouns encode, such as personal and demonstrative pronouns (e.g. *it* and *this/that*), i.e. "inanimate

entity", is indistinguishable. Thus, the difference between these two referring forms lies in the amount of perceptual information they mark, i.e. personal pronouns encode a higher degree of accessibility than demonstrative pronouns.

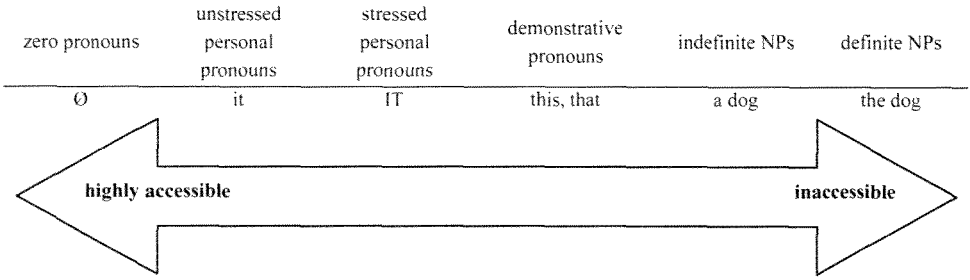

Figure 2.1: Concept of Accessibility in Reference Resolution (e.g., Ariel, 1990; 2001)

The hierarchy is based on three form-function relation criteria, according to which the fuller the form, the more informative, rigid, and unattenuated it is. The **information** criterion refers to the amount of lexical/conceptual information on the form[13]. The **rigidity** criterion concerns whether or not a unique referent can be chosen (thus the degree of disambiguation or concreteness)[14], and the **attenuation** criterion refers to the phonological size of the form, namely length and loudness (more attenuated = smaller phonological size). Table 2.1 shows how these three criteria classify the forms *it* and *that* in terms of their fullness.

	it	*that*
informativity	same	same
rigidity	same	same
phonological size (length of expression and loudness)	less	more
	less full form	**fuller form**

Table 2.1: Classification of *it* and *that* in terms of fullness of form according to Ariel's Theory

According to this table, it becomes clear that it is only one of the three criteria which differentiates between the fullness of the two forms *it* and *that*: phonological size. On the basis of the reverse mapping principle, the personal pronoun thus marks a higher degree of accessibility than the demonstrative. The factors that make an antecedent more or less

[13] e.g. *Mary's husband* encodes more lexical information than *he*.
[14] Rigidity can be achieved in two ways. 1) NP is unequivocal e.g. *the sun*, 2) some NPs are more rigid than others, because the addressee has a specific co-reference interpretation of the expression, e.g. *the teacher* might be understood better than *my neighbour from the 5th floor* (Ariel, 1994; p.30)

accessible are, on the one hand, the properties of the antecedent and on the other hand, the relationship between the antecedent and the anaphor. According to Ariel, e.g. topicality (global discourse topics > local discourse topics > non-topics) is a relevant antecedent property which contributes to the antecedent's accessibility (among other factors). Global discourse topics are associated with a higher degree of accessibility than local discourse topics, and the same for local discourse topics compared to non-topics. However, there is not an exhaustive list of antecedent properties which defines accessibility. Ariel suggests that it is not a single factor but rather a combination of factors which contributes to an antecedents' accessibility; but it is difficult to pin down which factors are involved and thus what saliency actually is.

Turning to German personal and d-pronouns, the question arises as to where they are located on the accessibility scale. Like in the English example of *it* and *that*, the amount of *conceptual information* on *er* and *der* is equal, in that reference is made to a single entity of either biological or syntactic masculine gender. This means that the *informativity* criterion cannot distinguish between the two forms (see Table 2.2). Neither does the *rigidity* criterion differentiate between the forms. The only difference, like for the English example of *it* and *that*, lies in the *phonological size* with *er* being more attenuated than *der*. *Der* is thus a fuller form and according to the reverse mapping principle should refer to less salient antecedents than *er*.

	er	*der*
informativity	same	same
rigidity	same	same
phonological size	less	more
	less full form	**fuller form**

Table 2.2: Classification of the German personal (*er*) and d-pronoun (*der*) in terms fullness of form according to Ariel's Theory

Different from the English and German examples is the case of the Dutch personal pronoun *hij* and the d-pronoun *die*. With regard to *conceptual information*, *hij* encodes reference to a single entity of common (so either masculine or feminine) syntactic gender, or biological gender if used for a person. *Die* encodes reference to a single entity of either biological or syntactic common (masculine or feminine) gender. Therefore, *hij* conveys more conceptual information than *die*, in that it is not possible to use *die* to refer to a woman (or to a feminine animal). Thus, *hij* is more informative than *die*, because it carries more lexical content than *die*. Neither form is more *rigid*[15] than the other. The Dutch d-pronoun *die* is

[15] One could argue that due to the three-way gender distinction with the personal pronouns opposed to the two-way

not less attenuated than the personal pronoun *hij*. Thus, none of the three criteria actually defines the d-pronoun *die* as a fuller form than the personal pronoun (see Table 2.3).

	hij	*die*
informativity	more	less
rigidity	same	same
phonological size	same[16]	same
	fuller form	**less full form**

Table 2.3: Classification of the Dutch personal (*hij*) and d-pronoun (*die*) in terms of fullness of form according to Ariel's Theory

The attempt to position German and Dutch personal and d-pronouns on the accessibility hierarchy is therefore problematic. While the German personal and d-pronoun fit the classification criteria, the criteria cannot account for the properties for the Dutch pronouns[17].

Ahrenholz (Ahrenholz, 2007; p.243) extended the accessibility hierarchy[18] according to his corpus findings on German d-pronouns, and the linguistic contexts in which they appear. His addition appears in Figure 2.2. Interestingly, Ahrenholz positions all uses of d-pronouns between unstressed and stressed personal pronouns. This means that he considers stressed personal pronouns to be fuller forms than d-pronouns, thus while d-pronouns refer to less salient entities than unstressed personal pronouns, stressed personal pronouns refer to even less salient entities. The materials of the experiments reported in this thesis only consisted of unstressed personal and d-pronouns. Therefore, the unstressed personal pronoun is quite high on the accessibility scale, while the d-pronoun occupies the mid-point.

gender distinction with the d-pronoun, the personal pronoun appears to be more rigid. However, we understand this criterion to rather define the degree of disambiguation carried by lexical NPs, such as *the sun*.

[16] In the recordings of the stimulus materials of the Dutch experiment reported in chapters 3.1.4 and 4.1.4, the average duration for the pronunciation of *hij* (mean = 241 ms; std = 33 ms) was longer than the pronunciation for *die* (mean = 202 ms; std = 25 ms), so that at least for the length of the pronoun in our experiment one could argue that the personal pronoun was (if anything) a less attenuated form than the d-pronoun.

[17] Ariel (1994) explicitely mentions the case of the Dutch d-pronouns (demonstrative pronouns) not fitting her informativity criterion. She still classifies them as fuller forms according to their level of expectedness (with Comrie, 1994) which is supposed to be lower than for personal pronouns.

[18] Actually, he followed Givón's Topic Identification Devices (1984), but for presentational reasons I will work his suggestions into the accessibility hierarchy (which is also based on Givón's theory).

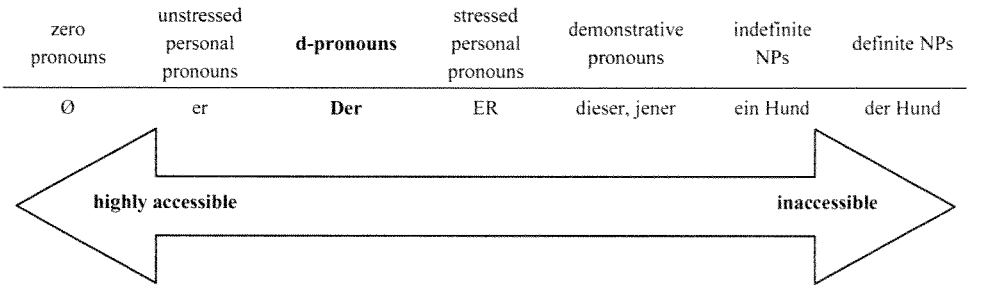

Figure 2.2: German Accessibility Scale adapted from Ahrenholz (2007)

Interestingly, for Dutch personal and d-pronouns, Broeder (1991; p.128) comes to a different conclusion. He analyzed the use of different referring expressions in Dutch to either establish, maintain, or shift reference, and proposes the order exemplified in Table 2.4. Note that Broeder does not differentiate between d- and demonstrative pronouns (they build the same category here). In this table, Broeder presents personal pronouns in Dutch to be more flexible than d-pronouns. They are likely to establish reference to an entity, whereas this is uncommon for d-pronouns. In addition, topic shift commonly occurs with personal pronouns, but it is only likely to appear with d-pronouns. Broeder explains his pro-NPs subdivision into personal and d-/demonstrative pronouns as follows: "[T]he two types of pro-NPs may function differently in marking referential movement in the discourse. For example, in a narrative in which there are only two persons to refer to, a man and a woman, the personal pronouns *hij* ('he') and *zij* ('she') will unambiguously mark referential movement; reference through demonstrative pronouns [d-pronouns] would not be unambiguous (p.128-129)".

Encoding devices		Establishment	Shift	Maintenance
Full-NPs:	indefinite NPs	common	uncommon	uncommon
	definite NPs (incl. proper names)	likely	common	uncommon
Pro-NPs:	**personal pronouns**	**likely**	**common**	**common**
	d-pronouns (demonstrative pronouns)	**uncommon**	**likely**	**common**
Zero-NPs:	Ø	uncommon	uncommon	common

Table 2.4: Broeder's table on the occurrence of Dutch referring expressions in different information structural contexts

Broeder does not differentiate between stressed and unstressed personal pronouns, and it remains unclear whether he treats unstressed and stressed tokens of personal pronouns as

belonging to the same category. This might at least partially account for the flexibility of personal pronouns reported here. Therefore, it cannot be decided whether it is an inherent property of Dutch personal pronouns to be more flexible and thus to encode less salient entities than d-pronouns, or whether this pattern is due to the non-differentiation of stressed and unstressed tokens of the personal pronouns.

2.2.3. Gundel et al.'s Givenness Hierarchy

The Givenness Hierarchy by Gundel, Hedberg and Zacharski (1993) is another cognitive approach which attributes referential expressions to different cognitive statuses of the mental representations of the referents. These cognitive statuses are differentiated by the givenness of their representations which may arise from the linguistic or non-linguistic context. The main difference to Ariel's Accessibility Hierarchy is that the levels of the Givenness Hierarchy on the left (see Figure 2.3) entail the levels to the right, which is not necessarily the case in Ariel's approach.

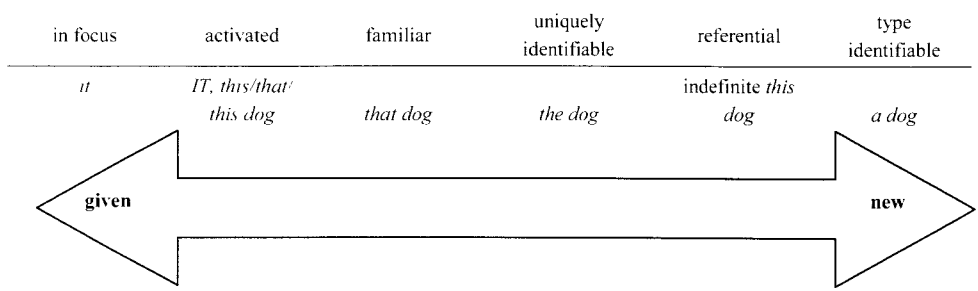

Figure 2.3: Gundel et al.'s Givenness Hierarchy (1993)

Gundel et al. (1993) studied five different languages, namely English, Russian, Japanese, Mandarin, and Spanish, on which the theory was based. Since German and Dutch are most similar to English (as non-pro drop languages), in the following I will present the different activation statuses and their associated English referential expressions, before classifying the German and Dutch pronouns on the hierarchy. According to the hierarchy, there are six different levels of activation: Type identifiable, Referential, Uniquely identifiable, Familiar, Activated and In focus[19]. These levels and their definitions are given in (26).

[19] Note that the terminology *in focus* refers to the cognitive status of the mental representation of the referent, namely to be in the current *focus of attention*, and is independent from the pragmatic *topic-focus* parameters distinction.

(26) **Type identifiable** – identify what kind of thing this is.
Referential – associate a unique representation by the time the sentence is processed
Uniquely identifiable – associate a unique representation by the time the nominal is processed
Familiar – associate a representation already in memory
Activated – associate a representation from working memory
In focus – associate a representation that your attention is currently focused on.

Gundel (2003)

According to this hierarchy, a demonstrative pronoun is associated with an *activation* status, i.e. the representation is currently in working memory. The use of a personal pronoun does not only entail that the mental representation is in working memory, but also that it is in the current focus of attention, i.e. the *in focus* status. The antecedent of the demonstrative pronoun could be, but needs not be, *in focus*, while the antecedent of the unstressed personal pronoun has to be *in focus*. According to the Givenness hierarchy the use of the demonstrative pronoun in (27)a is ambiguous as it could refer to either *the package* or *the table*. The personal pronoun in (27)b however, is more likely to refer to *the package*, since it is in the subject position. This syntactic position is associated with an entity in focus. *The table* is less likely to be in focus, since it occupies a syntactically less prominent position, and is thus less accessible as an antecedent for the personal pronoun. Therefore, the personal pronoun in this English example can be said to be less ambiguous than the demonstrative pronoun.

(27) The package$_i$ was on the table$_j$.
a. **That**$_{i,j}$ looked new.
b. **It**$_i$ looked new.

Gundel (2003) further argues that demonstrative pronouns refer more frequently to non-nominal antecedents than personal pronouns, as shown by numerous studies (Byron & Allen, 1998; Hegarty, Gundel, & Borthen, 2002; Webber, 1991). In example (28), the demonstrative pronoun is likely to refer to the entire phrase *the court does not* (a), while the personal pronoun prefers the NP antecedent *the court* (b), resulting in a different reading. The reason is that non-nominal antecedents are less likely to be *in focus* than nominal antecedents. Note that in the case of German and Dutch personal and d-pronouns this latter differentiation would only hold for the neuter forms (German: *es, das*; Dutch: *het, dat*), but not for the masculine forms (German: *er, der*; Dutch: *hij, die*) which are used in the

experimental materials in this thesis.

(28)

a. We believe her, the court does not, and **that** resolves the matter
b. We believe her, the court does not, and **it** resolves the matter

Gundel (2003), sentence a) adopted from NY Times, 5/24/00

Although, demonstrative pronouns can refer to antecedents which are *in focus* (as *the kitchen* in (29)) since they by definition are also *activated*, from the interaction of the Givenness Hierarchy with general pragmatic principles (Q-Principle: "Make your contribution as informative as possible") usually follows that when more than one possible antecedent is available, the demonstrative pronoun prefers a *non-focal* entity (as *the hallway* in (30)b). The personal pronoun, however is marked for the *in focus* status (as *the kitchen* in (30)a) and can therefore only refer to the *focal* entity. Since the demonstrative pronoun is "unspecified" for this *in focus* status, it typically marks the *non-focal* entity for discourse continuation and therefore signals focus shift.

(29) John's kitchen is really cozy. **That**'s my favorite room in the house.

(30)

a. Anyway, going back from the kitchen$_i$ then is a little hallway$_j$ leading to a window. Across from the kitchen$_i$ is a big walk-through closet. And next to **it**$_j$...
b. Anyway, going back from the kitchen$_i$ then is a little hallway$_j$ leading to a window. Across from the kitchen$_i$ is a big walk-through closet. And next to **that**$_j$...

What does this mean for the distribution of personal and d-pronouns in German and Dutch? For unstressed personal pronouns in German and Dutch, the Givenness Hierarchy predicts their antecedents should be *in focus*. In the case of Dutch where the d-pronoun is also a demonstrative form, it predicts that the d-pronoun behaves like the English demonstrative *that*. Thus, the entity referred to by a d-pronoun can be *in focus*, but needs at least to be *activated*. In the case of the German d-pronoun, we assume that the Givenness Hierarchy also predicts that the entity referred to needs to be activated, as it can refer to *non-focal* entities as *the cook* in (31). However, it is unclear how to differentiate between the German d-pronoun *der* and the demonstrative pronoun *dieser*, as the hierarchy would predict an *activated* status for the two types of pronouns.

(31) **Der Arzt** wollte mit **dem Koch** Tennis spielen. Doch ***der*** war krank.
De arts wilde met **de kok** tennissen. Maar ***die*** was ziek.
The doctor *wanted to play tennis with* ***the cook****. But* ***he*** *[D] was sick.*

Furthermore, with regard to the *in focus* status entailing the *activated* status, it is predicted that the personal pronoun is marked for referring to an entity *in focus*, while the antecedent of the d-pronoun can be *in focus* but must be *activated*. Thus, when two potential antecedents are available for reference, the d-pronoun is more flexible in its assignment. Although, according to the pragmatic Q-principle being preferentially assigned to the non-focal entity, it could be resolved towards either entity, and could thus result in more ambiguity. While Gundel (2003) mentions that the syntactic subject occupies a prominent position in English and thus signals that the entity is *in focus*, the question arises as to whether this is also true for flexible word order languages. Thus, while the theory straightforwardly assigns a cognitive status to a certain type of referential expression, it is not clear which (cross-linguistic) factors bring the representation of a referent *in focus* or results in an *activated* status. For the first set of experiments in this thesis, I will examine pronoun resolution after canonical antecedent structures like that shown in (31), where the first-mentioned topical entity is assumed to be the *focal* entity, while the second-mentioned non-topical entity is *non-focal*. In the case of non-canonical word orders, the second-mentioned pragmatically focused entity is predicted to be in the current center of attention.

As is clear from the above review of theories of reference, there are quite distinct predictions for the resolution of personal and d-pronouns, on the one hand; and between German and Dutch on the other, even though the languages are typologically close and both have personal and d-pronouns available. At this point, it may be useful to take a closer look at the pronominal systems in German and Dutch and the differences between them.

2.3 German and Dutch Pronouns

As pointed out in the introduction, in German and Dutch contrary to English, an utterance as in (31) can be realized by using a personal pronoun like in English as well as by using a d-pronoun. But there are also differences with regard to the characteristics of the pronominal forms in both languages. One important difference is how both languages encode gender.

German has a three-fold gender system with masculine, feminine and neuter gender which is realized in the pronominal and article system. Dutch only realizes the three-fold gender

distinction on personal pronouns, where a distinction between masculine, feminine and neuter is made. However, in the case of d-pronouns a two-fold distinction is made with regard to neuter and common (masculine and feminine) forms. This is also found in the article system in Dutch which differentiates between *de* (common[20]) and *het* (neuter[21]) nouns (see Table 2.5).

	German	Dutch	English
masculine	der Löffel	de lepel	the spoon
feminine	die Gabel	de vork	the fork
neuter	das Messer	het mes	the knive

Table 2.5: Definite articles within noun phrases in German, Dutch and English

Personal as well as d-pronouns can be used to refer to animate as well as inanimate entities. But in the case of Dutch personal pronouns, the use of *zij* is mainly restricted to signal biological gender. According to Audring (2006) this is because of the loss of knowledge about masculine and feminine nouns in Dutch speakers (excluding the dialects of Limburg, Brabant and Flemish speakers), due to the bipartite article system. When referring to inanimates, gender is used on pronouns to differentiate between individual (countable) and mass objects. She proposes the following hierarchy for the distribution of pronominal genders in Dutch:

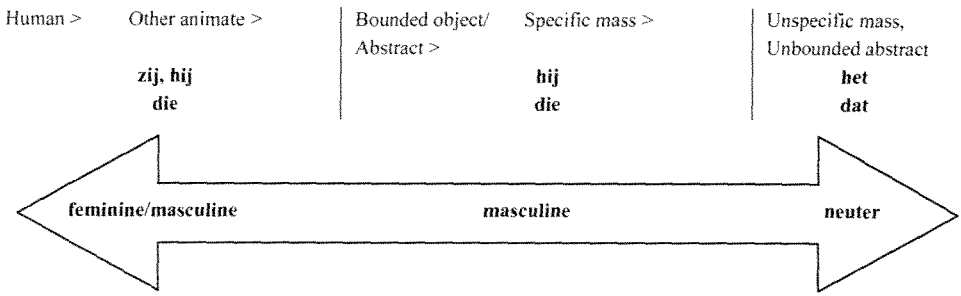

Figure 2.4: Dutch Individuation Hierarchy accounting for the distribution of pronominal genders in Dutch (Audring, 2006; p.103)

Animals have a special position on this hierarchy. Their biological gender is referred to if they have a special importance for the speaker as example (32) (from Audring, 2006; p.104-105) illustrates.

[20] in the following: common = [C]
[21] in the following: neuter = [N]

(32) A: Als je zo'n intelligente **kip** [C] hebt heb je eigenlijk geen hok voor nodig. 'k bedoel dan blijft **ie** [M] ook wel binnen de ruimtes die je '**m** [M] geeft.
B: **Ze** [F].
A: Gewoon zeggen van nou blijf daar zitten dan doet **ie** [M] dat wel.
B: **Ze** [F].
A: Ja, een kip is een "ze."

*A: When you have such an intelligent **chicken**, then you don't really need a cage. I mean **it (lit.: he)** will stay inside the space that you give it **(lit.: him)**.*
*B: **Her**.*
*A: Just say stay here and **it (lit.: he)** will do that.*
*B: **She**.*
A: Yes, a chicken is a "she."

Audring (2006) points out that d-pronouns in Dutch have the same gender pattern as the Dutch article system: common and neuter. Therefore, they provide an alternative to the mismatch problem of personal pronouns, and might be used as an avoidance strategy. However, in the Corpus of Spoken Dutch she also finds cases where the noun gender and the gender of the d-pronoun mismatch (see (33)).

(33) neuter > common
dan kreeg ik van Els namelijk nog een **berichtje** [N],
die [C] had ze blijkbaar vrijdag gestuurd.
*Then I got a **message** [N] from Els,*
*it seems that she sent **it (lit.: him/her)** [C] on Friday.*

With regard to the classification of d-pronouns, both German and Dutch grammar books usually introduce them as demonstratives. While in Dutch, *die* is a "true" demonstrative form, in the sense that it can be used to encode distance relationships, the German d-pronoun *der* is neutral in this regard. Therefore, a linguistic debate about the classification of d-pronouns has mainly taken place in the German literature, where some researchers have argued against the demonstrative denomination, in particular Ahrenholz (2007) who named them d-pronouns to indicate that they are to be distinguished from other pronominal forms, as well as Klein and Rieck (1982), Lambrecht (1994) and Weinrich (1993) who see them as a second system of personal pronouns.

Besides their similarity to demonstrative pronouns, German d-pronouns have one more

special feature which Dutch d-pronouns do not share, i.e., they resemble morphologically German definite articles (see Table 2.6 and Table 2.7). Despite the morphological proximity, the forms differ when case marked in the genitive singular and plural (*des Löffels* vs. *dessen*, *der Löffel* vs. *derer*) and dative plural (*den Löffeln* vs. *denen*) (unlike their demonstrative counter parts).

		definite NP *the spoon*	**d-pronoun**
Singular	Nominative	der Löffel	der
	Genitive	**des** Löffels	**dessen**
	Dative	dem Löffel	dem
	Accusative	den Löffel	den
Plural	Nominative	die Löffel	die
	Genitive	**der** Löffel	**derer**
	Dative	**den** Löffeln	**denen**
	Accusative	die Löffel	die

Table 2.6: Morphological similarity of the masculine definite article and d-pronoun in German

	definite NP *the spoon*	**d-pronoun**
Singular	**de** lepel	**die**
Plural	**de** lepels	**die**

Table 2.7: Common definite article and d-pronoun in Dutch (Dutch does not mark case)

Both, the German and the Dutch d-forms are morphologically congruent with German relative pronouns. Historically, both the definite article and the relative pronoun have developed from ("demonstrative") d-pronouns (Drosdowski et al., 1995). Their interpretations differ with regard to specificity (Lambrecht, 1994; p. 82). In example (34)a, the syntactic subject of the subordinate clause is a relative clause which has a specific and a non-specific interpretation. The book which is looked for may be any red book, thus at the time of the utterance it is unidentifiable for both speaker and addressee. Nevertheless, it could also be a specific red book which is identifiable to the speaker. For the d-pronoun in the second main clause of sentence (34)b, however, only a specific reading is possible.

(34)

a. Ich suche ein Buch, *das* rot ist.
I am looking for a book that's red.

b. Ich suche ein Buch, *das* ist rot.
I am looking for a book that's red.

The similarity that exists between the anaphoric use of personal and d-pronouns, and demonstrative and d-pronouns is illustrated in example (35) by Zifonun et al.(1997; p.558). Both the personal and the d-pronoun can be used to refer to the syntactic subject as well as the object antecedent. However, this is not true for the demonstrative pronoun *dieser*, for only sentence (35)b would license its use. Whenever the non-topical entity (*the Benz*) is not a semantically possible candidate for the demonstrative pronoun, the pronoun is interpreted as referring to a discourse-external referent. This is different for the d-pronoun which is preferentially resolved towards the non-topical entity, but may as well refer to the topical entity (*Peter*).

(35) Peter will einen Benz kaufen.
a. Er/Der/*Dieser hat wohl zuviel Geld. (Peter)
b. Er/Der/Dieser soll aber nicht so teuer sein. (Benz)

Peter wants to buy a (Mercedes) Benz.
*a. He [P]/He [D]/*He [DEM] might have too much money. (Peter)*
b. It [P]/It[D]/It [DEM] should however not be so expensive. (Mercedes Benz)

The Dutch pronouns *die* [C] and *dat* [N], which I call d-pronouns here, are actually classified as demonstrative pronouns which encode distal relationships (as opposed to the proximal demonstrative pronouns *deze* [C] and *dit* [N]). However, their frequent occurrence in spoken Dutch and their anaphoric use which constitutes an alternative to the use of personal pronouns (Audring, 2006; Comrie, 1997) make them comparable to the German d-pronouns (*der, die, das*) which are also frequently used in spoken German and (unlike demonstrative pronouns such as *dieser*) may appear in contexts of personal pronouns, as the comparative German-French example (36) by Schreiber illustrates (1999; p.219, the example is from Alfred Dölin's book "Berlin Alexanderplatz" and its French translation by Zoya Motchane). However, when two possible antecedents are available, as in example (37), then French also makes use of another type of pronoun, a demonstrative, to disambiguate the relationship.

(36) Der Arm ist ab, ***der*** wächst nicht mehr. (d-pronoun in German)
Le bras est amputé, ***il*** ne repoussera plus. (personal pronoun in French)
The arm is amputated, ***it*** *won't grow anymore.*

(37) Der Wirt lächelt, als sie alleine sind, **Franzen** an. ***Der*** streckt behaglich die Beine unter den Tisch. (d-pronoun in German)
Restés seuls, le patron sourit à **Franz**. ***Celui-ci*** très à son aise, allonge ses jambes sous la table. (demonstrative pronoun in French)
Being among themselves, the bar keeper smiles at ***Franz****.* ***He*** *[D] stretches comfortably his legs out under the table.*

Ahrenholz (2007; p.244) finds evidence for this type of disambiguating use in his corpus (see (38), transcription symbols omitted). The use of the d-pronoun clarifies the fact that reference is made to *the table* and not to *he*, representing the protagonist (most likely a person) in this example. Ahrenholz argues that if the personal pronoun had been used in this example, it would result in more ambiguity.

(38) I-R: aber als **er** sich auf-m stu/eh **tisch** abstützen wollte fiel **der** also zusammen
I-R: but when ***he*** *wanted to prop himself on the chai/eh* ***table it*** *[D]/****he*** *[D] then collapsed*

The disambiguation function of d-pronouns is of special interest to this thesis, as it implies that d-pronouns are marked for disambiguation while personal pronouns are neutral in this regard. This might then result in personal pronouns being more ambiguous and therefore more difficult to interpret than d-pronouns.

Ahrenholz (2007) analyzed two corpus data (P-Moll and KIH Corpus) of spoken German and found that *der, die* and *das* are among the most frequently occurring words in German. Out of 4945 occurrences (which also include definite articles, relative pronouns, and others), 1801 (36%) are occurrences of d-pronouns (see Figure 2.5).

Part of his data originates from taped student-professor conversations during consultation hours (KIH Corpus)[22]. The figures in Table 2.8 show that d-pronouns are used in more than half of all the occurrences of anaphoric pronominal reference. However, this analysis includes the neuter pronouns *es* (*it*, personal pronoun), *das* (d-pronoun), and *dis* (dialectical form of *dieses*, demonstrative pronoun). These forms are to be differentiated from masculine and feminine forms, in that they have additional specific functions such as referring to an entire clause.

[22] Ahrenholz does not find evidence that the use of d-pronouns implicates impertinence, moreover they are used to the same extend by students and professors.

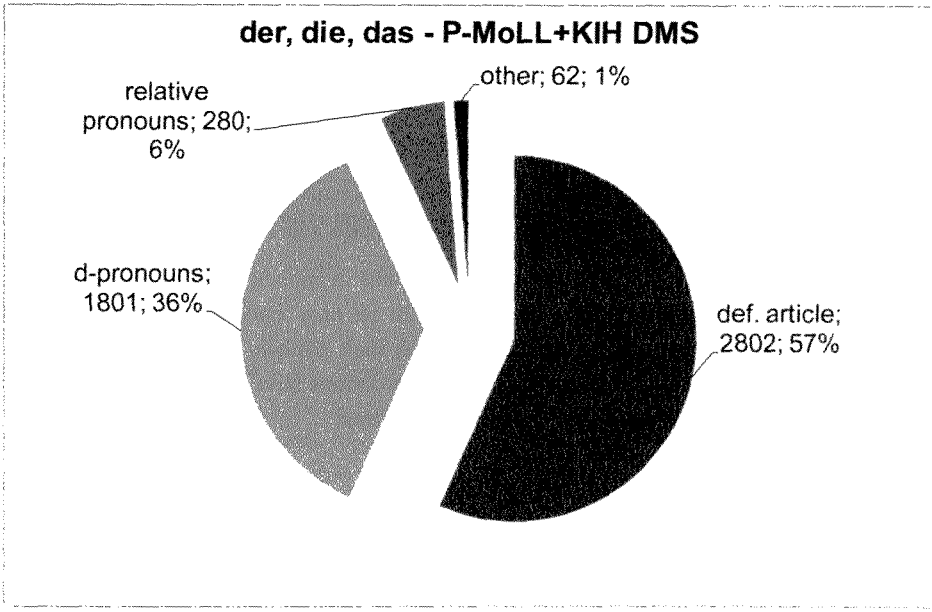

Figure 2.5: Ahrenholz (2007; p.212), Total occurrences of *der, die, das* in the corpus data (P-Moll and KIH Corpus) of spoken German (L1 data)

Form	Classification	Occurrences
er/sie/...	personal pronouns: he/she/...	437
der/die/...	d-pronouns	302
es	neuter personal pronoun: it	17
das	neuter d-pronoun	660
dis	dialectal form of the neuter demonstrative pronoun "dies"	0
dies-	demonstrative pronouns: this	8
Total	**Total**	**1424**

Table 2.8: Ahrenholz (2007; p.217), Corpus analysis of anaphoric occurrences of personal, d- and demonstrative pronouns in spoken German (P-Moll Corpus)

By subtracting the neuter forms from the analysis, Table 2.9 is obtained. As can be seen, the amount of occurrences of the personal pronouns *er* and *sie* (*he* and *she*) are more frequent (59%) than the occurrences of the d-pronouns *der* and *die* (40%). Nevertheless, the occurrence of d-pronouns in spoken German appears to be high, especially when compared to the occurrences of the demonstrative *dies-* (which reflects both the masculine and the feminine realization of *dieser* and *diese*) with 1%[23].

[23] The frequent use of German d-pronouns seems to be bound to spoken German. Bosch, Rozario and Zhao (2003)

	Amount	Percentage
personal pronouns	437	59%
d-pronouns	302	40%
demonstrative pronouns	8	1%
Total	**747**	**100%**

Table 2.9: adapted from Ahrenholz (2007; p.217), Corpus analysis of anaphoric occurrences of personal, d- and demonstrative pronouns in spoken German (P-Moll Corpus); neuter forms excluded

Helbig and Buscha (2001; p.229) claim that d-pronouns in spoken German usually receive stress. But Ahrenholz (2007; p.231) points out that he only finds a few occurrences of stressed d-pronouns in his corpus. He believes that their stressed occurrences have been overestimated by grammarians in that the use of d-pronouns have mainly been considered in emphatic and pejorative assertions and to signal contrast.

As we have seen in this section, although German and Dutch have similar pronominal systems, there are also some differences between the forms, such as the German three-way gender distinction for personal and d-pronouns. Critically for the current set of experiments, while in German the personal and d-pronoun carry the same amount of lexical information, in Dutch they appear to be differently informative in that the personal pronoun can disambiguate between masculine and feminine biological gender, while the d-pronoun cannot. With regard to the above outlined theories of reference (Ariel, 1990, 1994, 2001; Levinson, 1987, 1991) there should be complementary resolution behavior for the two types of pronouns, in that personal pronouns prefer more salient and d-pronouns less salient antecedents. A similar differentiation between the functions of personal and d-pronouns was assumed under Lambrecht's (1994) general rule which predicts that personal pronouns prefer topical antecedents while d-pronouns prefer antecedents which are not (yet) topical. The assumption that d-pronouns have a disambiguation function in contrast to the relatively neutral personal pronouns (Ahrenholz, 2007; Zifonun, et al., 1997) predicts that when more than one possible antecedent is available for pronoun resolution, disambiguation is the chief job of the d-pronoun in comparison to the personal pronoun, which signals neutral or default co-reference.

In the next section I will turn to psycholinguistic accounts of reference which have developed from empirical work.

---

analyzed a German Newspaper corpus (the NEGRA corpus), and found that 94.8% of all occurrences of *der* represented the definite article, 4.8% were occurrences of the relative pronoun, and only 0.4% were occurrences of the d-pronoun.

2.4 Resolving Pronominal Reference – Psycholinguistic Evidence

In this section, first, I outline two accounts of reference which have developed from psycholinguistic work on the resolution of personal and d-pronouns. Then, I summarize the empirical findings on pronoun resolution.

2.4.1. Bosch et al. and The Complementary Hypothesis

The Complementary Hypothesis assumes a complementary distribution of resolution preferences for personal and d-pronouns, usually based on one or multiple antecedent criteria, particularly with regard to the notion of salience. The hypothesis does not specify which antecedent criteria form salience, but when e.g. topicality is seen as the decisive factor, then according to the complementary hypothesis one would predict that if personal pronouns are resolved towards topical entities, d-pronouns will be resolved towards the non-topics. Indeed this is what Bosch et al. (2003) claimed when assuming that "[A]naphoric personal pronouns prefer referents that are established as discourse topics, while demonstratives [d-pronouns] prefer non-topical referents". However, the hypothesis could also be formulated in terms of grammatical role or other antecedent properties. Importantly, the resolution preference for one pronoun indicates the resolution preference for the other pronoun; therefore, the resolution preferences are in complementary distribution.

The Complementary Hypothesis for pronoun resolution was formulated on the basis of Kameyama's (1999) Complementary Preference Hypothesis which works out a differentiation between unstressed and stressed pronouns in English, and according to which "A focused pronoun [i.e. stressed pronoun] takes the complementary preference of the unstressed counterpart (p.315)." This idea is illustrated by Lakoff's (1971) example in (39), where in case of two unstressed personal pronouns a parallel reading is achieved, with the subject pronoun (*he*) referring to the subject antecedent (*John*) and the object pronoun (*him*) referring to the object antecedent (*Fred*). In the case of two stressed pronouns, this relationship is reversed and the subject pronoun (*HE*) now refers to the object antecedent (*Fred*) and vice versa.

(39)

a. John called Fred a Republican. Then he inSULted him.
b. John called Fred a Republican. Then HE insulted HIM.

The same idea also exists in other domains of anaphora resolution, such as the distinction between zero and personal pronouns in pro-drop languages like Italian and Spanish. As

Carminati (2002) pointed out, the distinction between the use of overt and null pronouns is often made with reference to the fact that only the overt is phonologically realized and can thus carry stress. According to Luján (1985, 1986), pro drop languages use a null pronoun where English would use an unstressed pronoun, and an overt pronoun in a pro drop language functions like an English stressed pronoun. However, Carminati (2002) only found partial evidence for this hypothesis. She reported that the use of null pronouns in Italian indeed resembled the use of unstressed pronouns in English, but she noted that some occurrences of overt pronouns in Italian resembled occurrences of reduced rather than stressed pronouns.

In the case of personal and d-pronouns, there are also two pronominal options where other languages such as English only have one (personal pronoun). However, in Dutch and German a one-to-one mapping to English stressed and unstressed pronouns is not reasonable, since in both languages one can also stress and unstress the pronominal forms (leading to a total of four options).

2.4.2. Kaiser and Trueswell and The Form-Specific Multiple-Constraints Approach

Previous studies carried out in flexible word order languages such as Finnish have tested the influence of *multiple factors*, particularly the influence on pronoun resolution of order of mention and grammatical role. Kaiser and Trueswell's (2008) study also investigated the difference in resolution preferences of the Finnish personal pronoun (*hän*) in comparison to the resolution of the Finnish demonstrative pronoun (*tämä*).

Kaiser and Trueswell's Form-Specific Multiple-Constraints Approach is based on three principles:

- Multiple factors rather than one single factor are involved in anaphor resolution.
- Different anaphors are sensitive to different characteristics of the antecedents, e.g. Finnish personal pronouns were found to be more sensitive to grammatical information while demonstrative pronouns were sensitive to both grammatical and topicality information.
- Anaphors may be sensitive towards the same antecedent characteristics but the degree of sensitivity towards these factors may differ. For example, if personal and d-pronouns were both sensitive towards discourse information, it might be the case that this is a very strong criterion for d-pronouns, while it has a relatively lower influence on the resolution of personal pronouns.

The account is interesting for this thesis, since it not only allows different anaphoric expressions to be differentially sensitive to different factors, but it also allows this

sensitivity to differ cross-linguistically.

As has become clear from the above-outlined accounts, there is common acceptance of the idea that multiple factors influence reference resolution. In the next section, we will review some empirical findings on the influence of some of these factors.

2.4.3. Empirical Findings on Pronoun Resolution

While there seems to be common acceptance of the idea that personal pronouns refer to more salient/accessible entities than fuller forms, the identification of the factors that lead to this elevated degree of accessibility, their relative contributions and interactions have occupied researchers over the past 30 years. These studies have found that different factors have a primary influence in pronoun resolution, e.g. order of mention according to Gernsbacher's First Mention Account (1989), grammatical role according to the Subject Preference Account (Crawley, et al., 1990; Frederiksen, 1981) or parallel functions in the Parallel Functions Account (Grober, Beardsley, & Caramazza, 1978; Smyth, 1994). Relatively recently researchers have started to elaborate these questions in flexible word-order languages which allow for an investigation of multiple factors and to test their influence on different pronominal forms (Bosch, et al., 2007a; Bosch, et al., 2003; Bosch & Umbach, 2007b; Bouma & Hopp, 2007; Järvikivi, Van Gompel, Hyönä, & Bertram, 2005; Kaiser & Trueswell, 2004, 2008; Schumacher, Roberts, Ellert, & Järvikivi, *in prog.*; Wilson, 2009). These studies have mainly used different antecedent word orders to disentangle the factors order of mention and grammatical role.

In the next sections, I summarize the English findings on the influence of the factors order of mention and grammatical role in language comprehension, then, I outline how later studies turning to flexible word-order languages and different pronominal forms have attempted to disentangle these two factors.

2.4.3.1. Order of Mention

The *first-mentioned account*[24] claims that while building a mental representation of a sentence, the first-mentioned entity may function as a basis on which subsequent information will be added to (Carreiras, Gernsbacher, & Villa, 1995; Gernsbacher, 1990; Gernsbacher & Hargreaves, 1988; Gernsbacher, Hargreaves, & Beeman, 1989). Thus, the importance of the first-mentioned entity is attributed to general cognitive mechanisms rather than to linguistic prominence.

[24] also called *the structure building framework* (Gernsbacher, 1990)

Using a probe recognition task[25], Gernsbacher and Hargreaves (1988) showed that readers were faster in verifying that the referent *Tina* was part of the sentence when it was mentioned first as in (40)a than when it was mentioned second as in (40)b. This first-mention effect was also shown for sentences in passive voice, which presented *Tina* as the semantic patient of the utterance as in (40)c. The first-mention effect (for *Tina*) even persisted when *Tina and Lisa* were both the syntactic subject of the sentence as in (41)a, as opposed to only *Tina* being the subject as in (41)b. Gernsbacher and Hargreaves (1988) concluded that the cognitive factor *order of mention* was more important than linguistic factors, such as agentivity or subjecthood, for comprehenders to build a coherent mental representation of the sentence.

(40)

a. **Tina** beat **Lisa** in the state tennis match.
b. **Lisa** beat **Tina** in the state tennis match.
c. **Tina** was beaten by **Lisa** in the state tennis match.

(41)

a. **Tina** and **Lisa** argued during the meeting.
b. **Tina** argued with **Lisa** during the meeting.

In English however, the first-mentioned position is also linguistically a very prominent position. Syntactic subjects, agents or topical elements usually appear in this position and this ranks its informational value very high. Therefore, order of mention may be a more important cue for English comprehenders than for comprehenders of more flexible word order languages, such as Spanish, who may encounter syntactic objects or verbs in initial position. To further examine this possibility, Carreiras, Gernsbacher and Villa (1995) replicated Gernsbacher and Hargreaves (1988) study in Spanish. Instead of presenting sentences in active and passive voice, they presented the sentences in SVO and OVS order as in (42). The first-mention effect was present in both Spanish word orders; this was taken as cross-linguistic evidence that the first-mention advantage is a general cognitive mechanism, as the *structure building framework* suggests.

(42)

a. **María** invitó à **Diana** a cenar en casa.
Mary$_{SUBJ}$ *invited Diane*$_{OBJ}$ *for dinner at home.*

[25] In this task, a written sentence is presented word by word on a screen. At a certain point of the unfolding sentence, a probe word appears on the screen and the participant is asked to indicate as fast as possible if the word has appeared in the sentence or not. Reaction times are measured.

b. A **Diana** la invitó **María** a cenar en casa.
Diane$_{OBJ}$, Mary$_{SUBJ}$ invited (her) for dinner at home.

Although, the above studies were able to show the importance of the factor order of mention in language comprehension, they both used probe recognition tasks. These tasks have been criticized in psycholinguistic research (Gordon, Hendrick, & Foster, 2000). It has been shown that they measure task-dependent strategies: participants try to keep track of the words which qualify as probes and engage in memory specific strategies rather than establishing a mental representation of the sentence itself. Therefore, it has been argued that probe recognition tasks are not suitable for capturing co-reference interpretations and language comprehension processes in general (Gordon, et al., 2000)[26].

2.4.3.2. Grammatical Role

The importance of the grammatical function of the antecedent has been postulated by two different accounts in the psycholinguistic literature: the *subject preference account* and the *parallel functions account.*

The *subject reference account* was based on Frederiksen's (1981) work which showed that a subject antecedent was more likely to be attached to a subsequent pronoun as compared to an antecedent in object position. Furthermore, Frederikson found shorter reading times for sentences beginning with a pronoun when the pronoun was referring to a subject antecedent than when it was referring to an object antecedent. He did not find a reading time difference for subject and object pronouns referring to the subject antecedent, suggesting that pronouns, irrespective of their own grammatical sentence function, have a strong preference to co-refer with subject antecedents. But as Frederikson pointed out, his sentences were not completely free of semantic bias and could have therefore biased the resolution preferences.

The *Parallel Functions Account* was first proposed by Sheldon (1974). Her experimental evidence comes from the area of relative clause acquisition by English-speaking children, but she extended her findings to the area of pronominalization. Under this account it is assumed that a pronoun co-refers with the antecedent with the same grammatical function. That is, an unstressed subject pronoun is universally attached to a subject antecedent (43)a, while an object pronoun is resolved towards an object antecedent as in (43)b.

---

[26] Furthermore in a later study, Gernsbacher (1989) used full NPs and pronouns as probe words and found a facilitatory effect for repeated names, formulated as the explicitness hypothesis. These findings violated well established evidence showing that on the contrary, there exists a repeated name penalty when full NPs are used to maintain reference in comparison to pronouns (Gordon, Grosz, & Gilliom, 1993). Gordon et al. (2000) mentioned that this finding might have occurred, because of the task-dependent strategies.

(43)

a. John hit Bill and then he kicked Sarah. (he = John)
b. John hit Bill and then Sarah kicked him. (him = Bill)

Grober, Beardsley and Caramazza (1978) tested the parallel functions strategy in pronoun resolution with a written sentence completion task. They presented test booklets with sentence beginnings such as in (44), which contained an ambiguous subject pronoun. Grober et al. found that in over 70% the subject pronouns were resolved towards subject antecedents and took this as evidence for the parallel function account. However, since their materials did not provide any object pronouns, this result could also be taken as evidence of support for the subject preference account.

(44) *John must scold Bill because he...*

Grober et al. (1978) also used two different types of connectives preceding the pronoun: *but* and *because*. Interestingly, the completion pattern differed according to the type of connective. When following *but* the pronoun was most frequently interpreted as referring to the subject antecedent (90%). In contrast, after the connective *because* there was no significant subject preference for the pronoun (52%)[27]. The 'denial of expectation' character of the connective *but* seemed therefore to initiate more subject interpretations (this finding was also robust across different types of verbs), while the interpretation of the pronoun after *because* highly depended on the verb and the direction of its implicit causality (e.g. NP1: *apologize, accuse*; NP2: *criticize, forgive*). This finding indicates that there are also semantic factors involved in pronoun resolution. Grober et al. (1978) state that when the implicit causality information of the verb marks the direction of the object antecedent, then "the pronoun is [...] assigned coreferential with it, in violation of the proposed strategy [parallel functions strategy] (p.129)".

Crawley, Stevenson and Kleinman (1990) used a reading task and an assignment task to systematically test the resolution preferences for object pronouns in sentences which were controlled for semantic bias (see (45)). They found a preference to interpret the object pronouns with the subject antecedents, and thus proposed a re-interpretation of Grober et al.'s (1978) results in favor of the Subject Preference Account.

---

[27] Grober et al. (1978; p.125) used a Chi Square test and found a significant association of type of connective and resolution preference. However, they did not test whether the resolution pattern for *because* (NP1=279; NP2=255) showed a significant first-mentioned preference or no preference. Therefore, based on their results, I calculated a Chi Square test comparing the frequency with which the first- and the second-mentioned antecedent was chosen compared to a chance level frequency (NP1=267; NP2=267), which revealed no antecedent preference for *because*, $\chi^2(1) = .540$, $p > .05$.

(45) Brenda and Harriet were starring in the local musical. Bill was in it too and none of them were very sure of their lines or the dance steps. Brenda copied Harriet and Bill watched **her**. (her = Brenda)

Smyth (1994) examined Crawley et al.'s results and criticized their materials. Specifically, it was noted that some items consistently showed a subject preference, while others had an object preference. Smyth argued, therefore, that the sentences were not fully parallel (e.g. the first main clause had an accusative object, while the second presented a dative object), therefore Crawley et al. did not find a parallel function assignment. Smyth argued that parallel function assignment occurs when the antecedent and the anaphor share the same grammatical and thematic role, when both clauses have the same attachment site and constituent structure. For the parallel function assignment to occur, special requirements need to be met, which is also in line with Grober et al.'s (1978) findings. Thus, it does not seem to constitute a default strategy, as has been argued to be the case for the subject preference account (e.g. Centering Theory).

2.4.3.3. On First-mentioned Objects and Second-mentioned Subjects

In this section, I review the findings of two psycholinguistic studies which systematically investigated how order of mention and grammatical role information are processed when resolving pronouns in the flexible word order language, Finnish. By manipulating the word order of the antecedent structures, order of mention effects were disentangled from grammatical role effects. Pronoun resolution was measured after SVO and OVS structures, which enabled the researchers to directly compare e.g. first-mentioned objects to first-mentioned subjects and to second-mentioned objects, in order to draw conclusions about the relative influence of the two salience factors. Like the experiments in this thesis, both studies used a visual-world eye-tracking task to measure the resolution preferences in real time.

Järvikivi et al. (2005) investigated the resolution preferences for the Finnish personal pronoun *hän* after different antecedent word orders (46). They found both an early subject advantage (480 – 1110 ms) and a later first-mention effect (690 – 1110 ms) for subject and object antecedents[28]. For the personal pronoun following SVO antecedent structures, this led to an overall preference to look at the first-mentioned subject (*Tony Blair*). For the personal pronoun after OVS antecedent structures, there was no first-mention preference;

[28] Instead of analyzing word order as a factor, Järvikivi et al. (2005) were interested in effects of grammatical role and order of mention. They found significant main effects for both order of mention and grammatical role, but no interaction. Also, their data suggests that there was no main effect of word order; however, whether word order would have interacted with grammatical role cannot be determined, even though the data suggests it might.

"if anything, there were more looks to the subject [*Tony Blair*] than to the object [*George Bush*] character (p.264)". Järvikivi et al. (2005) took this as evidence against one-factor models such as the *first mention account* and the *subject preference account*, and suggested that pronoun resolution was rather guided by an interplay of both factors.

(46)

a. SVO
Tony $\text{Blair}_{\text{SUB}}$ kätteli George $\text{Bushia}_{\text{OBJ}}$ valkoisessa talossa. Hän halusi keskustella Irakin tilanteesta.
Tony $\textit{Blair}_{\text{SUB}}$ *shook hands with George* $\textit{Bush}_{\text{OBJ}}$ *in the White House. He [P] wanted to discuss the situation in Iraq.*

b. OVS
George $\text{Bushia}_{\text{OBJ}}$ kätteli Tony $\text{Blair}_{\text{SUB}}$ valkoisessa talossa. Hän halusi keskustella Irakin tilanteesta.
George $\textit{Bush}_{\text{OBJ}}$ *shook hands with Tony* $\textit{Blair}_{\text{SUB}}$ *in the White House. He [P] wanted to discuss the situation in Iraq.*

Kaiser and Trueswell (2008) base their Form-Specific Multiple-Constraints Account (introduced in section 2.4.2 above), on data from experiments comparing the resolution preferences for the Finnish personal pronoun *hän* and the demonstrative *tämä* following SVO and OVS antecedent structures similar to those in (46). In contrast to Järvikivi et al. (2005), for the personal pronoun *hän* they found an overall subject preference, but no first-mention preference. For the demonstrative *tämä* they found different influences. After SVO sentences, there was an initial subject preference which, over time, switched into a second-mention preference (700 ms). After OVS sentences, there was initially no preference for either antecedent, but later, a second-mention preference was observed. The different resolution patterns for the two pronouns, which were not simply an asymmetric distribution, lead the authors to the assumption that personal and d-pronouns are not necessarily equally sensitive towards the same antecedent properties; rather their resolution might be guided by different factors (or the interplay between them).

Although both sets of Finnish results support the Form-Specific Multiple-Constraints Account, the two studies above find differences in the interpretation for the personal pronoun *hän*: Järvikivi et al. (2005) find an influence of grammatical role and order of mention but Kaiser and Trueswell (2008) only find a grammatical role effect. Kaiser and Trueswell (2008) attribute this difference to the fact that they embedded their antecedent and pronominal structures in a discourse context, whereas Järvikivi et al. (2005) presented

their structures in isolation, which might have affected the resolution after the non-canonical, OVS sentences. This then raises the question of what sort of influence the discourse context might have had, and in particular, what differences occur after canonical and non-canonical antecedent structures. Kaiser and Trueswell (2008) claim that different pronominal forms are sensitive to information from different levels of representation, namely syntactic and discourse information. They found that after OVS sentences the postverbal entity was chosen for the personal as well as the d-pronoun, and point out that OVS sentences are mainly used to signal that the object is discourse-old (or topical) and the subject is discourse-new information (non-topical). Alternatively, the fact that they found the same interpretation pattern across pronouns in these information-structurally marked environments could also point to the fact that only the second-mentioned entity was in fact considered for reference, and that the first-mentioned entity had been discarded from discourse continuation in general. This could be due to the fact that OVS structures are used to indicate topic shift. The subject is presented in the focus position and it may be the case that in these contexts, focus is a better predictor of salience than topichood. In other words, maybe there was not so much ambiguity after all, since the antecedent structure itself had delivered a cue for disambiguation (before the pronoun had even been encountered). This possibility is discussed in more detail in chapter 3.2.

2.4.4. Pronoun Resolution in this Thesis

This chapter reviewed some critical findings with regard to pronoun resolution and the factors which have been proposed to influence it. While earlier studies have made use of probe recognition tasks, which have been argued to measure task-dependent strategies, later studies have used the visual-world paradigm which allows one to investigate how pronouns are resolved in spoken language in real time, and without interrupting the speech signal.

Using flexible word order languages such as Finnish can help to disentangle order of mention effects from grammatical role effects. However, the results from earlier studies are not consistent. Furthermore, it is unclear whether pronoun resolution after OVS sentence structures is influenced by the same constraints as pronoun resolution after SVO sentence structures. Therefore, the experimental part of this thesis will investigate resolution preferences of personal and d-pronouns after comparative sentences. Special attention will be paid to the role of order of mention of the antecedent candidates, and the influence of canonical and non-canonical antecedent structures on resolution preferences. While previous studies have investigated order of mention by controlling for grammatical function in different word orders, in this thesis order of mention will be examined in antecedent structures which do not present a grammatical subject-object distinction, by embedding the antecedents in double nominative comparative structures respectively. Furthermore, the

experiments will systematically examine resolution preferences in the same materials across the two languages German and Dutch, in order to find out whether there is evidence for cross-linguistic resolution patterns, as claimed by theories of reference.

Resolving Ambiguous Pronouns in L1 German and Dutch

Chapter 3

3.1. Pronoun Resolution after Canonical Antecedent Structures[29]

3.1.1. Introduction

This part of the chapter explores the influence of the order of mention of the antecedent candidates on the resolution preferences in German and Dutch after canonical antecedent structures. The results will be used as a cross-linguistic basis to which the L2 learner results from chapter 4.1 of this thesis will be compared.

As shown in the previous chapter, both German and Dutch have two pronominal forms to refer to a discourse entity, namely personal (G: *er*, D: *hij*) and d-pronouns (G: *der*, D: *die*), as can be seen in (47), where in English a personal pronoun would be used.

(47) German: Peter will einkaufen gehen. **Er/Der** hat wohl zuviel Geld.
Dutch: Peter wil gaan boodschappen. **Hij/Die** moet wel te veel geld hebben.
English: *Peter wants to go shopping. **He** might have too much money.*

Furthermore, while the use of the personal pronoun is unambiguous in English in (48) (example from Zifonun, et al., 1997, p.558), in German and Dutch the personal and the d-pronouns need to be disambiguated (here by the semantic context), since they both could refer to *Peter* as well as to *the Benz*.

(48) German: Peter will einen Benz kaufen. **Er/Der** soll aber nicht so teuer sein.
Dutch: Peter wil een mercedes gaan kopen. **Hij/Die** moet echter niet te duur zijn.
English: *Peter wants to buy a (Mercedes) Benz. **It** should however not be too expensive.*

Personal pronouns and d-pronouns have been argued to have different functions, as indicated by the fact that most linguists and psycholinguists call d-pronouns

[29] Parts of the research presented in this chapter will appear in Ellert, M., Järvikivi, J. & Roberts, L. (to appear). Information structure affects the resolution of the subject pronouns er and der in spoken German discourse. In Sarda, L., Thomas, S.C. & Fagard, B. [eds.] *Linguistic and Psycholinguistic Approaches to Text Structuring.* Amsterdam: John Benjamins.

"demonstratives". But there are several properties which distinguish d-pronouns from proper demonstratives, and some linguists argue that they are better understood as a second set of personal pronouns (Klein & Rieck, 1982; Lambrecht, 1994; Weinrich, 1993). We adopt here Ahrenholz's (2007) notion of *d-pronouns* to distinguish them from other pronominal forms.

The question of what precisely the function of *d-pronouns* is has not yet been fully answered. A common view is that they refer to entities which are not (yet) topics (Ahrenholz, 2007; Bosch & Umbach, 2007b; Lambrecht, 1994; Zifonun, et al., 1997). But whether this is due to an inherent feature of the d-pronoun or not is an open issue. It may be that it is marked to refer to a non-topical entity in the previous discourse (Ahrenholz, 2007; Bosch & Umbach, 2007b; Zifonun, et al., 1997), or it may just prefer non-topical antecedents in the way that personal pronouns prefer topical antecedents (Bosch, et al., 2003; Comrie, 1994; Diessel, 1999; Lambrecht, 1994).

According to Lambrecht, personal as well as d-pronouns refer to identifiable and active referents (the highest of Chafe's activation states), and their use differs with the topic-non-topic distinction. As a general rule, personal pronouns refer to topic antecedents, while d-pronouns refer to entities which are not yet topical. This is similar to Bosch et al. (2007b) who claim that personal pronouns refer to the topical and d-pronouns to the non-topical entities on the basis of the Complementary Hypothesis according to which the distribution of resolution preferences for both forms will be asymmetric.

This complementary distribution of interpretation preferences would also be predicted by Levinson's Pragmatic Neo-Gricean Theory of Anaphora (1987, 1991) according to which reference maintenance interpretations are given priority over reference shift interpretations, thus the topic antecedent is a good candidate for future reference and the use of a personal pronoun favors a neutral/default co-referential interpretation, towards the topical entity, while the use of a d-pronoun favors a disjoint/non-coreferential interpretation, towards the non-topical entity. Co-referentiality is achieved by using the most reduced yet recognizable form, which in this case is the personal pronoun. The d-pronoun, being a fuller form in terms of semantic content and length, favors a disjoint interpretation to the reference maintaining entity. But the information criterion "semantic content" points in the opposite direction for Dutch, since the Dutch personal pronoun distinguishes between three types of gender, while the Dutch d-pronoun only differentiates between two types of gender. Interestingly, this would predict differences in pronoun resolution between German and Dutch.

Gundel et al.'s (2003; 1993) Givenness Hierarchy principally makes the same predictions for the reference distributions of personal and d-pronouns, but there is one crucial difference to other theories of reference. Following an antecedent sentence which introduces two potential antecedents, such as (49), the personal pronoun *must* be resolved towards the first-mentioned topical entity (*the package*). This is because this entity is in the current center of attention, as opposed to the second-mentioned non-topical entity (*the table*) which is not in the center of attention, even though it is currently active in working memory. The demonstrative pronoun on the other hand can refer to both antecedents, as its reference only requires the antecedents to be in working memory (which they both are); the second-mentioned topical entity may be most frequently chosen according to preferences, but not due to cognitive requirements. Thus, the theory predicts that personal pronouns are marked for topical co-reference relations. In the case of a d-pronoun, more ambiguity may be involved in reference resolution because of the referential "unspecification" of this pronominal form.

(49) The package$_i$ was on the table$_j$.
  a. **It**$_i$ looked new.
  b. **That**$_{i/j}$ looked new.

Concerning Ariel's (1990, 2001) classification criteria for lower and higher accessibility markers on the Accessibility Hierarchy, d-pronouns are lower accessibility markers than personal pronouns, indicating to the addressee that the mental representation of the co-referred entity is of intermediate activation as opposed to high activation in the case of personal pronouns. Thus, personal pronouns should be resolved towards a more salient/accessible entity than d-pronouns. Note that while in Ariel's theory d-pronouns refer to entities whose representation is of an intermediate level of activation, according to Lambrecht they both refer to activated and not semi-activated referents, and so they only differ with regard to the topicality of their antecedents. Moreover, while d-pronouns in German fit the criteria to be classified as lower accessibility markers relatively well, Dutch d-pronouns do not. Since most theories have been developed on data from languages which do not have d-pronouns, they do not directly address these subtle differences between personal and d-pronouns. The same is true for psycholinguistic accounts, such as the First Mention Account (Gernsbacher, 1989) or the Subject Preference Account (Crawley, et al., 1990; Frederiksen, 1981), which were based on data from English personal pronouns. It is not clear, therefore, how to extend these findings in order to make predictions for German and Dutch personal and d-pronouns.

Psycholinguistic studies on the resolution of personal and d-pronouns have found different

results. Following their Complementary Hypothesis, Bosch et al. (2007a; 2007b) predicted that personal pronouns would be resolved towards the grammatical subject and/or topical entity and d-pronouns towards the object and/or non-topic, but they did not find support for their hypothesis. In items with canonical antecedent structures (50), they found no preference for the personal pronoun, and a second-mention preference for the d-pronoun (*the emergency patient*). Therefore, they suggested that while d-pronouns are marked to refer to the non-topical entity, personal pronouns may be neutral in this regard (see also Ahrenholz, 2007; Zifonun, et al., 1997).

(50) Im Krankenhaus.
At the hospital
Der Oberarzt untersucht den Notfallpatienten.
The senior doctor is examining the emergency patient.
Er/Der ist gerade erst gekommen.
He [P/D] has only just arrived.

Final-interpretation-preferences task
Der _____ ist gerade erst gekommen.
The _____ *has only just arrived.*

In contrast, Bouma and Hopp (2007), testing the resolution of *er* in an off-line interpretation-preferences task, found that it was highly sensitive to syntactic information. Following different word orders of main and subordinate antecedent structures, the personal pronoun was resolved towards the syntactic subject, which is not in line with what Bosch et al. (Bosch, et al., 2007a; Bosch & Umbach, 2007b) have found for the personal pronoun.

In general, off-line tasks might not be suitable to study whether the markedness of a pronominal form affects pronoun resolution preferences, since the relative degree of ambiguity (which is assumed to be higher for unmarked forms) can still lead to quite firm final interpretation patterns. That is, while there might be no preference during real-time resolution, ultimately (and especially in tasks which force participants to make a choice) a decision will be made.

Kaiser and Trueswell (2004) addressed the question of real-time reference resolution in spoken Dutch by using the visual-world eye-tracking paradigm. They presented items such as (51), all consisting of an SVO antecedent sentence which was followed by a subsequent personal or d-pronoun.

(51) Het begon uit de hand te lopen in het klaslokaal.
Things were beginning to get out of hand in the classroom.
De leerling stak de leraar speels met een scherp potlood.
The student poked the teacher jokingly with a sharp pencil.
Hij/Die was gekleed in een groene trui, omdat het buiten koud was.
He [P/D] was wearing a green sweater, because it was cold outside.
Het lijkt erop dat ze naar de rector moeten.
It looks like they will have to go see the principal.

The participants listened to the discourses, while their eye-movements to a corresponding visual scene containing the potential antecedents along with other distractors were measured. The looks towards the antecedent pictures were analyzed from pronoun onset onwards and it was presupposed that the more looks there were to one of the antecedents, the more likely it was to be a good co-reference candidate for the momentary pronoun interpretation. The advantage of this task in comparison to off-line tasks is that it measures moment-by-moment inferences about the unfolding discourse material without disrupting the input. Kaiser and Trueswell found that the personal pronoun preferred the first-mentioned antecedent, while the d-pronoun was resolved towards the second-mentioned entity. They concluded that personal pronouns favored syntactic subjects and d-pronouns objects. Since syntactic information, order of mention and topicality information all marked the first-mentioned antecedent as more salient, and the second-mentioned antecedent as less salient, the results cannot disentangle the influence of these factors. Nevertheless, these psycholinguistic findings are evidence for the fact that in Dutch as well as in German personal and d-pronouns have two different functions, even if it is not yet clear what they are.

In a visual-world study on the resolution of personal and d-pronouns in German, Wilson (2009) observed that the two pronominal forms were sensitive to different factors. The personal pronoun following SVO antecedent sentences showed an initial second-mention preference which ultimately switched to a first-mention preference. After OVS structures, no preference was observed for the personal pronoun. Wilson concluded that the personal pronoun was sensitive to both syntactic and discourse information, in the sense that it preferred topical subjects (SVO - S); but that due to a trade-off between these two constraining preferences, no preference was observed in the OVS condition (OVS – S/O). For the d-pronoun she found an overall second-mentioned non-topical preference (SVO – O; OVS - S). Therefore, she concluded that the d-pronoun was mainly guided by discourse information because it preferred non-topical antecedents. The direction of the results for the personal pronoun needs some more reflection. Particularly what is surprising is the above-

mentioned switch in preferences observed after SVO structures (from O→ S). Wilson used mixed models to analyze the eye-tracking data. She analyzed a time window of 2000 ms after pronoun onset, dividing it up into eight 250 ms time slices. She calculated a mixed model for each of the four conditions (SVO-*er*, SVO-*der*, OVS-*er*, OVS-*der*). *Time* and *order of mention*[30] were entered into the model as predictors and the interaction term was only kept when it was a significantly better predictor of the outcome. Looks to the first-mentioned entity were taken as the base level. For the SVO-*er* condition, the interaction ($b_{mention2nd\ x\ time} = -0.10$, $p < .01$), as well as both main effects ($b_{mention2nd} = 0.47$, $p < .05$; $b_{time} = 0.18$, $p < .001$), were significant. The beta coefficient of the main effect of order of mention indicated that there were overall more looks to the second-mentioned entity than to the first-mentioned entity (thus O > S). The beta coefficient of the interaction term indicated that as time progressed, fewer looks were initiated to the second-mentioned object. However, this does not necessarily entail a significant first-mention preference. When looking at the plotted log odd second-mentioned advantage scores, it in fact appears that there was no preference after the initial second-mention preference.

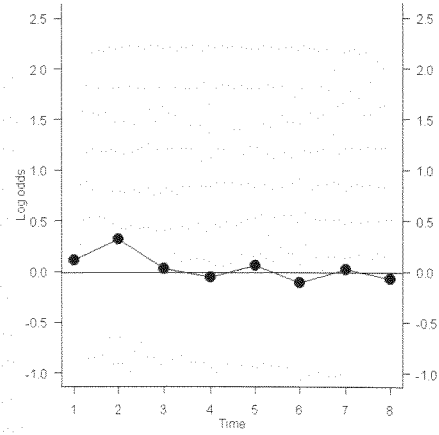

Figure 3.1: Log odds of looks to the second-mentioned and first-mentioned entity in the SVO-personal pronoun condition in Wilson (2009). A value greater than zero indicates a second-mention preference, and a value of less than zero indicates a first-mention preference.

Furthermore, in a second visual-world experiment on the resolution of *er* and *der* (active vs. passive antecedent structures), Wilson found no preference at all in the same SVO-*er* condition, which arguably strengthens the conclusion that there was no preference for *er* overall. This is also indicated by Wilson's findings of an additional acceptability judgment

[30] Wilson (2009) actually calls this factor *object* with the levels *NP1* and *NP2* which correspond to *first-* and *second-mentioned*.

task, where there was no preference for the SVO-*er* condition (p.143). This finding is in line with Bosch et al.'s (2007a; 2007b) results who also found no preference for the personal pronoun and a second-mention preference for the d-pronoun after SVO sentences, as reported above.

Note that the pattern of results from this set of studies is surprising because all of the above-mentioned accounts (Subject Preference Account, First-Mention Account, Theories of Reference) predict a first-mention effect for the personal pronoun. Wilson's conclusion does not fit this pattern of results very well. If the personal pronoun relied on syntactic and discourse information, then the question is raised as to why was it not resolved towards the first-mentioned entity after a canonical SVO antecedent structure. It would appear that Bosch et al.'s (2007b) explanation is more likely. The personal pronoun, being unmarked, did not elicit a clear preference because of its relatively high degree of flexibility, whereas the d-pronoun shows a clear second-mention preference, because it is marked to refer to the non-topical antecedent.

In a visual-world study by Schumacher, Roberts, Ellert and Järvikivi (in prog.), the German personal pronoun was shown to be sensitive mainly to syntactic information, in that it preferred subject over object antecedents in SVO and OVS contexts. This is in line with Bouma and Hopp's (2007) findings. However, the effect for the personal pronoun was not as robust after OVS sentences, which indicates that there was also an influence of word order. Taken together with the finding that the d-pronoun preferred object antecedents regardless of the type of antecedent structure, the results of this study overall fit with those of Bosch et al. (Bosch, et al., 2007a; 2003), according to which personal pronouns were mostly used to refer to antecedents in nominative case, such as syntactic subjects, versus d-pronouns which preferred antecedents in non-nominative case, such as syntactic objects. However, since the personal pronoun was sensitive to both grammatical role information and word order variations, Schumacher et al. (in prog.) concluded that the d-pronoun is the more marked form and therefore showed a robust effect of grammatical role. In contrast, the personal pronoun despite its preference for grammatical subjects was more affected by word order variation, due to its unmarkedness. This explanation is in line with previous assumptions on the disambiguating function of d-pronouns in ambiguous contexts (Ahrenholz, 2007; Bosch & Umbach, 2007b; Zifonun, et al., 1997).

Given the range of results on the functions of personal and d-pronouns, our understanding of the resolution constraints and preferences of the two pronominal forms is limited. There are two more problems. First, because of the different experimental set-ups and the small number of studies in general, it is unclear whether the resolution differences which have

been observed between the Finnish (see chapter 2.4.3.3) and the German experiments can be attributed to cross-linguistic differences rather than more trivial methodological differences. Therefore, a cross-linguistic perspective was chosen for the studies in this dissertation.

Second, all of the previous studies tested the resolution of personal and d-pronouns following SVO and OVS sentences in order to disentangle the effects of order of mention and grammatical role. Thus, the resolution to, for instance, first-mentioned objects was directly compared to the resolution to first-mentioned subjects and second-mentioned objects. However, these studies could not control for potential information structure influences, because the SVO word order is unmarked in contrast to OVS. It may be that pronoun resolution differs following OVS versus SVO structures, because in the former case, they are marked and their antecedents have different information statuses than antecedents in SVO sentences. Therefore, in the studies presented in this section, the experimental design comprised two pronominal conditions, and one type of antecedent structure, which was a canonical comparative sentence. In other words, the role of order of mention is investigated independently from grammatical role cues by presenting the two potential antecedents in a comparative double nominative construction such as: *NP1-verb-comparative-NP2* (52).

3.1.2. The present study

How does the factor order of mention of the antecedent candidates affect the resolution of personal and d-pronouns? Is there a difference in the influence on the two types of pronouns? Theoretical accounts of reference and psycholinguistic results as discussed in chapter 2, suggest that if there is a positional effect, it may guide the resolution of personal pronouns towards the first-mentioned antecedent because it is the topic, while d-pronouns prefer the second-mentioned non-topical antecedent. Furthermore, if the d-pronoun is marked for co-reference to non-topical entities and the personal pronoun is neutral in this regard as has been suggested previously, we might expect a higher degree in ambiguity for the personal pronoun than for the d-pronoun. On the other hand, if according to the Givenness Hierarchy (Gundel, 2003; Gundel, et al., 1993) we consider d-pronouns to be more neutral in their co-reference relations and personal pronouns to be more constrained (due to the necessary *in focus* criterion), then we expect to find the reverse pattern, with more ambiguity in the resolution of the d-pronoun. As mentioned above, off-line tasks might not be sensitive enough to capture these differences in markedness, as final interpretation preferences might be subject not only to inherent features of the pronouns but to discourse constraints in general. Thus the research questions are addressed both with a questionnaire task and an on-line visual-world eye-tracking experiment. In the latter task,

the prediction is that if a pronoun is ambiguous, it should elicit equal numbers of looks to the two potential antecedents for a longer period of time in comparison to a less ambiguous pronoun.

In this experiment we ask how German and Dutch native listeners use positional cues to resolve personal and d-pronouns, when there is no subject-object distinction in the antecedent clause (52).

(52) German: **Der Schrank** ist schwerer als **der Tisch.** ***Er /Der*** stammt aus einem Möbelgeschäft in Belgien.
Dutch: **De kast** is zwaarder dan **de tafel.** ***Hij/Die*** is afkomstig uit een meubelwinkel in België.
English: ***The cupboard*** *is heavier than* ***the table.*** ***It*** *[P/D] originates from a furniture store in Belgium.*

A second issue to be addressed here is whether the resolution preferences for personal and d-pronouns are found cross-linguistically: Are personal and d-pronouns resolved in the same way in German and in Dutch or is there, despite the typological proximity, a difference? Thus, the two research questions are:

1. Does the order of mention of the antecedent candidates influence the resolution of personal and d-pronouns?
2. Is there cross-linguistic evidence between German and Dutch for general resolution preferences for personal and d-pronouns?

3.1.3. Pronoun resolution in German

The first experiment investigates the resolution of personal and d-pronouns in German.

3.1.3.1. Methods

Participants

Twenty-eight German native speakers (22 female, 6 male) participated in the study. The participants were students at the Radboud University Nijmegen or employees at the Max-Planck Institute for Psycholinguistics in Nijmegen. They were aged between 20 and 31 years (mean = 23.25; SD = 2.68). All participants were tested individually and were paid for their participation. All participants had normal or corrected-to-normal vision.

Experimental Tasks

All participants undertook two experimental tasks which both consisted of the same experimental items. This allowed a better comparison of the results across the tasks. The experiment started with a visual-world eye-tracking task. Participants' eye movements were recorded while they viewed three pictures on a screen and listened to short discourses related to the pictures. Each picture displayed one referent which was mentioned in the discourse. The task involved answering content questions which appeared randomly after some of the stimuli to assure that the participants were paying attention to the meaning of the discourses.

After the eye-tracking task, the participants filled in a forced-choice questionnaire asking them explicitly to choose an antecedent for the pronoun. In that sense, the off-line-task would provide information about the participants' final decisions, and the on-line-task would provide information about if and when they disambiguated the critical pronoun. Since the off-line task explicitly asked for their preferred interpretation, it was administered after the eye-tracking task, to ensure as far as possible that during the eye-tracking task the participants would remain unaware of the constructions under investigation. After the experiment, the participants completed a background questionnaire which provided information about their age and made sure that German was their native language. The whole session took about 45 min (eye-tracking 25 min and questionnaires 20 min).

Materials and design

Both experimental tasks contained the same items. Twenty-four experimental items were constructed (see Table 0.1 in the Appendix), each beginning with a comparative antecedent sentence of the type *NP1-verb-comparative-NP2* that introduced both referents with a singular masculine definite NP, one in preverbal and one in postverbal position (see (53)). Both NPs appeared in nominative case. An SVO main clause followed, which constituted the target clause and started with a subject pronoun. The subject pronoun was either a personal pronoun (53)a or a d-pronoun (53)b, yielding two experimental conditions. The discourses ended with a third wrap-up context sentence. The sentence segments following the pronoun were constructed to be free of semantic bias to make the discourses fully ambiguous throughout the duration of the whole trial.

In both conditions, the subject pronoun could co-refer to either the first-mentioned NP *the cupboard* or the second-mentioned NP *the table*. Thus, for the two types of pronouns a first-mentioned as well as a second-mentioned interpretation was available which made the resolution ambiguous.

(53) Conditions: Sample Item

a. Personal Pronoun Condition
Der Schrank ist schwerer als der Tisch. **Er** stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.
The cupboard is heavier than the table. ***It*** *[P] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

b. D-pronoun Condition
Der Schrank ist schwerer als der Tisch. **Der** stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.
The cupboard is heavier than the table. ***It*** *[D] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

According to the above outlined linguistic and psycholinguistic accounts, the first-mentioned topical antecedent was predicted to be more accessible for the personal pronoun, thus favoring a first-mentioned interpretation, while the second-mentioned non-topical antecedent was predicted to be more accessible for the d-pronoun, favoring a second-mentioned interpretation.

For the on-line task, twenty-four experimental items were constructed. Two versions of each of the 24 experimental items were then created, either containing a personal or a d-pronoun, counterbalanced in a latin square design. The experimental items were interspersed among 48 filler items, half of which started with a comparative structure, and the other half containing only non-comparative clause structures. The comparative filler items were constructed to better mask the experimental items. They presented two NPs of the same gender without being followed by a subsequent subject pronoun (see (54)).

(54) Das Telefon ist lauter als das Radio. Die Zuschauer fühlten sich sehr gestört, als das Telefon im Theater während der Vorstellung klingelte. Das war eine peinliche Situation.
The phone is louder than the radio. The audience felt very annoyed, when the phone was ringing during the theater performance. That was an embarrassing situation.

The total number of 72 items was split into two experimental blocks, and the order of the blocks was counterbalanced between-participants. The order of the stimuli within each block was pseudorandomized. Five non-comparative practice items were added to the material, to familiarize the participants with the eye-tracking task.

Each item was digitally recorded to computer. A male native German speaker read each item out loud. The experimental items were recorded separately for each condition to avoid splicing effects. Intonational stress on the pronouns was avoided. The items were cut into two separate sound files with the PRAAT software (Boersma & Weenink, 2009) as illustrated in (55), so that the experimental items were separated into a first sound file playing the antecedent sentence and a second sound file starting with the critical pronoun. Thus, the second sound file was the target sound file.

(55) Sound file 1: Der Schrank ist schwerer als der Tisch.
The cupboard is heavier than the table.
Sound file 2: Er stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.
It originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.

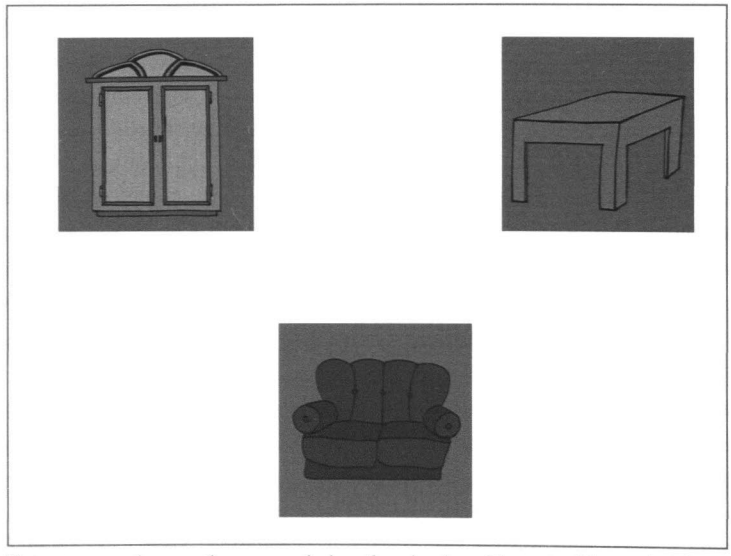

Figure 3.2: Pictures appearing on the screen during the visual-world eye-tracking task. The upper pictures (cupboard, table) show the two target referents and the picture below (sofa) shows a discourse-related non-target referent (appearing in the wrap-up sentence)

Each experimental screen showed three pictures for each trial. The pictures originated from the MPI picture database. They were each presented in a 288x288 pixel frame. They appeared on three positions on the screen (1024x768 pixel) in a triangular mode: top left

corner (at 171,167 pixel), top right corner (at 855,167 pixel) and down central position (at 512,599 pixel). Each experimental trial contained two target pictures (e.g. *the cupboard, the table*) which either appeared in top left or top right position, and a discourse-related non-target picture (e.g. *the sofa*) which appeared in down central position. The position of the target pictures was counterbalanced within conditions. Each of the target pictures only appeared once during the experiment.

The experiment was programmed with the Experiment Builder Software of SR Research. Participants' eye movements were recorded with an SR Research EYELINK II eye tracker. The eye tracker is an infrared video-based system with a head-mounted camera. Only the dominant eye was recorded. A sampling rate of 500-Hz was used which monitored gaze locations every 2 ms. The calibration of the camera which links the position of the eyes with a certain location on the screen, ensured that spatial accuracy was at least 0.5°.

The off-line task was a forced-choice questionnaire which presented the discourses in a text format (see Figure 3.3). The twenty-four experimental items were interspersed among 36 filler items of different types which included a pronoun of any sort (demonstratives, feminine pronouns or plural pronouns). The two conditions resulted in two versions of the off-line questionnaire which had a total of 60 items (see Table 0.2 in the Appendix). Each participant saw the same number of experimental items in each condition, but no participant saw the same item more than once.

1	Die Biene ist fleißig und arbeitet jeden Tag an den Honigwaben. Abends kehrt **sie** erschöpft zum Bienennest zurück.
2	Die Zauberin hat ihren Zauberstab immer bei sich. Und bei Vollmond füllt **sie** einen großen Kessel mit Zaubertrank.
3	Der Schrank ist schwerer als der Tisch. **Der** stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.

Figure 3.3: Three example items from the off-line task. Item 3 is an experimental item.

Procedure

The participants first undertook the eye-tracking task. Each trial began with a drift correction to control for minor head movements between the trials. After this, the experimental display containing the three pictures was presented. The display was shown for 1000 ms before the onset of the first sound file allowing for a preview to inspect the scene. After the first sound file, the pictures on the screen disappeared and a fixation cross was presented in the middle of the screen at equal distance from each of the three pictures previously shown. This method was chosen to avoid any spillover effects from the end of the first sentence into the start of the second sentence containing the critical region. Participants fixated the cross in the middle of the screen for 1500 ms. The experimental

display reappeared and the second sound file, which started with the critical pronoun, was presented simultaneously. Participants were presented with three practice items prior to the experimental blocks. Additionally, one practice item was placed at the beginning of each experimental block. Between the two experimental blocks, the participants paused (ca. 5 min) and the camera was turned off. At the beginning of each block, the camera was recalibrated and validated.

Participants were told that they would hear several mini stories and that after some of them, they should answer a content question (yes/no-question). The questions were presented after some of the filler trials, but never after the experimental items. The participants were instructed to answer by clicking the left mouse button for a "yes"-answer and the right mouse button for a "no"-answer. They received immediate visual feedback on the correctness of their answer. The feedback was given to ensure that they would listen to the content of the discourses. The accuracy of the responses was very high with 95% correct answers (24 questions; mean correct answers = 22.79, SD = 1.01).

After the eye-tracking task, the participants received the forced-choice questionnaire which presented the discourses in a text format. The pronouns were written in bold letters and the participants were asked to circle the part of the text which the word in bold letters was referring to. This task was created in order to elicit the participants' final interpretations of the pronouns. With their choice, the participants indicated whether they resolved the pronoun towards the first- or the second-mentioned antecedent. This was coded respectively as 1 or 2 and subsequent calculations were performed on the frequency values.

3.1.3.2. Results

This section is divided into two parts. First, I will present the results of the questionnaire and then report the results of the eye-tracking task, although both tasks were presented to the participants in the other order.

Forced Choice Questionnaire

Figure 3.4 and Table 3.1 show that overall the personal pronoun was resolved towards the first-mentioned topical entity (97%), suggesting a strong first-mention preference for *er*. However, the resolution of the d-pronoun looks much less clear. With a weak majority of 56%, *der* was resolved towards the second-mentioned non-topical entity.

A Chi-Square test was conducted on the data. There was a significant association between the type of pronoun and whether the first- or the second-mentioned antecedent was chosen as the co-referential entity, χ^2 (1) = 231.99, $p<.001$. The odds of choosing the first-

mentioned entity were 46.71 times higher for personal than for d-pronouns.

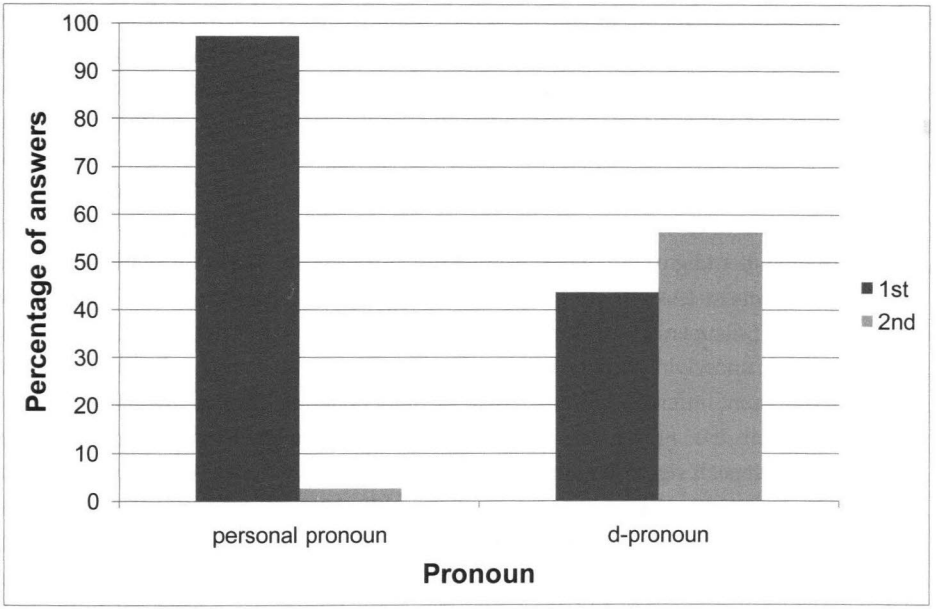

Figure 3.4: L1 German Results on the Forced Choice Questionnaire; pronoun being coreferential either with the first- or the second-mentioned entity (1st, 2nd) in the two conditions (personal pronoun, d-pronoun)

	1st	2nd
personal pronoun	97.32% (327)	2.68% (9)
d-pronoun	43.75% (147)	56.25% (189)

Table 3.1: L1 German Results on Forced Choice Questionnaire; pp = 28; items = 24

In order to investigate further whether the distribution, especially for *der*, was different from a 50%-chance level (indicating no preference), two more Chi Square tests were conducted in which the distribution of responses for one pronoun (e.g. *der,* 1st = 147, 2nd = 189) was compared to a 50%-chance level distribution (1st = 168, 2nd = 168; preference vs. no preference). These tests revealed that the association of *er* and the antecedent choice was significantly different from a no-preference-level, χ^2 (1) = 193.90, p<.001. The first-mentioned entity was 36.33 times more likely to be chosen for *er* than a 50%-chance level would have predicted. For *der* the distribution of responses was only marginally different from no preference, χ^2 (1) = 2.64, p = .061[31] with a tendency for the second-mentioned

[31] p>.05

entity.

Visual-World Eye-Tracking

Since we were interested in the resolution strategies for the personal and the d-pronoun, we required a task that would be able to measure resolution processes during real-time interpretation. Additionally, with regard to the German d-pronouns mainly appearing in spoken language, we needed a task that measured spoken language comprehension; thus, we chose the visual-world eye-tracking method (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). More specifically, we employed an eye-movement-during-listening paradigm in which participants listened to stories matching pictures presented on a screen (Altmann & Kamide, 1999; Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Cooper, 1974; Järvikivi, et al., 2005; Kaiser & Trueswell, 2004, 2008). The assumption underlying this task is that as the speech signal unfolds, listeners make moment-by-moment inferences about what they are hearing. These inferences can be captured by their eye movements, i.e. when confronted with visual scenes depicting elements of the speech signal (Figure 3.2), they are more likely to look at a certain picture upon hearing information about it (Altmann & Kamide, 1999; Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Cooper, 1974; Järvikivi, et al., 2005; Kaiser & Trueswell, 2004, 2008), or to even look at the position of a blank screen where the picture of a referent of the current speech input had been presented previously (Altmann, 2004; Richardson & Spivey, 2000). Language-mediated eye movements have been demonstrated for a large variety of language comprehension phenomena (Henderson & Ferreira, 2004), as well as for pronoun resolution in particular (Arnold, et al., 2000; Järvikivi, et al., 2005; Kaiser & Trueswell, 2004, 2008). Therefore, we decided to adopt this technique to investigate the resolution of personal and d-pronouns.

Since the duration of a spoken pronoun takes about 100 - 250 milliseconds, it is highly advantageous that the temporal resolution accounts for fixation durations in the millisecond area (here: every 2 to 4 ms). In addition, the technique allows for the study of the comprehension of spoken language without interrupting the speech signal. In the case of ambiguity resolution, this enabled us to study participants' moment-by-moment hypotheses about the linguistic input without explicitly asking them for their preferred co-reference interpretations.

Data Analysis

The proportions of fixations over time to both antecedent pictures (1st & 2nd) from the total fixations to the screen (looks to pictures and blanks) were analyzed for the experimental items in both conditions (personal and d-pronoun). The areas of interest,

which define the location and size of the regions to be analyzed, had the same size (288x288 pixels) as the pictures. The reference point for the analysis was the onset of the second sound file, since it began with the critical pronoun (reference point of timeline at 0 ms, see Figure 3.5). The interest period for analysis included the time window 200 ms prior (-200 ms) to the onset of the pronoun to control for looks to the target regions before pronoun onset. Note that at this time (and until the onset of the pronoun), participants viewed a fixation cross in the middle of the screen, i.e. no pictures were shown to them. However, if they had already looked at a target region although it was blank, then these looks were excluded from the analysis[32]. As Altmann (2004) has shown, looks to "blank" target regions may appear subsequently to picture presentation, when the current speech input refers to these previously depicted entities. In our study, participants had viewed the depicted elements on the screen during the presentation of the first sound file and prior to the screen showing a fixation cross in the middle. Although they were instructed to always look at the fixation cross in the middle, a few looks were made to the "blank" target regions. To make sure these looks were not due to postprocessing of the first sound file or memory specific effects, we excluded these prior to pronoun onset target looks from the analysis. In other words, we wanted to make sure that the target looks entering the analysis would inform us about the pronoun resolution preferences. Therefore, 42 looks (< 1%) to either target position that started before pronoun onset (-200 till 0 ms) were excluded from the analysis (for this type of cleaning, see Järvikivi, et al., 2005; Pyykkönen, Matthews, & Järvikivi, 2010). At the pronoun onset, the experimental display reappeared. While the speech signal was unfolding, participants' looks were recorded. The fixations to the target pictures were generated for every 4 ms time point between 200 ms before and 2000 ms after pronoun onset, resulting in 551 time points per participant and item. Figure 3.5 shows the proportion of target looks that were included in the analysis (4577 fixations in total).

The programming of a saccade until its launch takes approximately 200 ms (Matin, Shao, & Boff, 1993). In addition, the duration for a pronoun to be uttered takes about 100 to 250 ms. Thus, effects due to pronoun resolution should not be visible before 300 ms.

[32] Due to technical problems with the hardware set-up (namely the presentation of the stimuli was realized with a DirectX soundcard instead of an ASIO soundcard), there were random processing delays between the reappearance of the experimental screen and the start of the second sound file, so that an accidental second preview time was given to some of the participants for some of the trials. However, this problem was neglectigable, since the delays occurred randomly and because of our data cleaning procedure which prevented any looks to the target picture prior to the onset of the second sound file to enter the analysis.

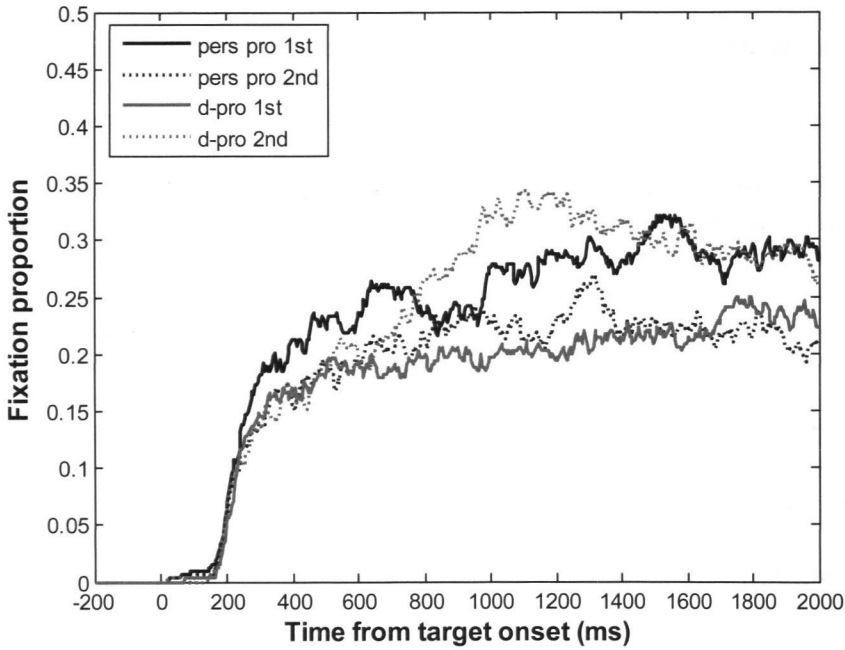

Figure 3.5: Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Overview of the eye movements

As can be seen in Figure 3.5, the looks to both target pictures for both pronouns *er* and *der* start rising at around 200 ms until approximately 600 ms.

Personal pronoun. The figure shows that in the *er* condition, 200 ms after the pronoun onset, the looks to both target pictures rise steadily and slowly for about 1300 ms, after which the looks to the second-mentioned target decrease while the looks to the first-mentioned target further increase until about 1600 ms after which they also decrease. This pattern of results reflects a high degree of ambiguity for the personal pronoun.

D-pronoun. Effects of the d-pronoun *der* start to become visible at about 600 ms after pronoun onset, where there is a sharp rise in second-mentioned target looks for about 500 ms, after which the looks decrease again. This pattern of results reflects that the d-pronoun is disambiguated relatively early and clearly towards the non-topical antecedent.

Taken together, the order of mention has an inverse influence on the two types of pronouns, with an earlier second mention preference for the d-pronoun *der* and a later first mention preference for the personal pronoun *er* which might indicate a higher amount of ambiguity for the personal pronoun compared to the d-pronoun.

Statistical Analysis of the eye movement patterns

We analyzed the on-line data using linear mixed-effect models (Baayen, 2008; Baayen, Davidson, & Bates, 2008) with participants and items as a crossed-random factor[33]. As in previous studies (e.g. Arnold, et al., 2000; Järvikivi, et al., 2005; Kaiser & Trueswell, 2008), the analyses were conducted on different time windows. The individual samples were averaged over time windows of 200 ms, resulting in ten time windows in addition to the prior time window (-200 – 0 ms) which was used for cleaning the data (see Table 3.2).

window	1	2	3	4	5	6	7	8	9	10
time (ms)	0-200	200-400	400-600	600-800	800-1000	1000-1200	1200-1400	1400-1600	1600-1800	1800-2000

Table 3.2: Time windows used for analysis; pronoun onset at 0 ms

The dependent variable, the frequency with which a picture was fixated, was transformed into empirical logits for each of the 200 ms bins. Empirical logits transformations were chosen rather than natural log odds, to avoid problems when the probability was approaching zero or one (Barr, 2008; Jaeger, 2008). The performance of data analysis on logits rather than fixation proportions was motivated by the fact that parametric assumptions regarding the homogeneity of variances (and therefore the normal and independent distribution of the mean and the standard error) are not met for proportional data since the values are bounded to a range between 0 and 1; a problem that can be dealt with by using the log odds scale which ranges from – infinity to + infinity and is symmetrical around zero (50%-chance) for positive and negative odds. By running the analysis on logits rather than proportions, multiplicative effects can thus be transformed into additive effects (a prerequisite for linear models which assume "constant effect size over the entire scale", Barr, 2008; p.5).[34]

For each time window, two linear mixed models were calculated. In a first model, pronoun type (called *condition* in the analysis with 2 levels: personal vs. d-pronoun) and order of mention (called *mention* in the analysis with 2 levels: 1st vs. 2nd) were entered as the

[33] A pdf-file containing a full overview of all statistical models calculated on the data of this thesis can be downloaded on the following web pages:

- Pdf-file with annotation: http://corpus1.mpi.nl/ds/imdi_browser?openpath=MPI1175800%23
- Direct access to pdf-file: http://hdl.handle.net/1839/00-0000-0000-0011-F0F9-4

[34] Empirical logits formula: $\eta' = \ln((y+.5)/(n-y+.5))$ (Barr, 2008; p.13)

predictors of interest (*simple model*), and in a second model, their interaction was added (*interaction model*). The *d-pronoun* condition and the *1st*-mentioned level were mapped onto the intercept. Thus, a positive beta coefficient for condition indicated more looks to the personal pronoun. A positive beta coefficient for mention indicated more looks to the second-mentioned entity.

(56) Formulas in R for both models[35]:
Simple model = lmer(looks ~ condition + mention + (1 | pp) + (1 | item))
Interaction model = lmer(looks~ condition*mention+ (1 | pp) + (1 | item))

This type of analysis was chosen because linear mixed-effect models can deal with participants and items as a crossed-random factor (Baayen, et al., 2008), and it takes into account the physical dependence of eye-movement data[36].

A stepwise approach to model building was used, applying a forward elimination method, which ensures that only the minimum number of predictors accounting for the variance in the data enter the model. The *simple model* containing both main effects of order of mention and type of pronoun was chosen as the most basic model, since we were interested in the estimates for order of mention and the two types of pronouns. The second model, the *interaction model*, investigated whether the looks to both types of target pictures were affected differently by the two pronouns. Thus, for the analyses, first both the *simple model* and the *interaction model* were calculated for a specific time window, and subsequently a loglikelihood test was calculated on the fit of the models. The fit of a model was also indicated by a smaller value of the model selection criteria such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The *interaction model* was only chosen when it significantly better explained the variance in the data. This procedure was repeated for the other time windows.

The *interaction model* explained the data better for all time windows starting 600 ms after pronoun onset (see Table 0.4 in the Appendix). This patterns with the above-mentioned observation that the earliest effect of type of pronoun becomes visible at around 600 ms. It suggests that an influence of type of pronoun on resolution patterns emerged at this time.

Turning to the analysis of fixed effects (see Table 3.3), the main effects of pronominal

[35] We used the lme4 software application in R (Bates & Sarkar, 2007).
[36] It is physically impossible to look at different objects on the screen every 4 ms, rather does a data point depend on the location duration of its preceding data point, knowing that it takes 200 ms to trigger a saccade (Matin, et al., 1993).

condition and order of mention were either marginally significant or significant across the same time windows (600 – 2000 ms), except for time window 4 where there was only a marginally significant main effect of pronominal condition (600 – 800 ms). The positive sign of the coefficient (beta) indicated that overall there were more looks for the personal than for the d-pronoun, while at the same time there were overall more looks to the second-mentioned target picture than to the first-mentioned target picture.

Time window	in ms	Fixed predictors		
		pronoun condition	order of mention	interaction
1	0-200	0.06 (0.901)	-0.01 (-0.167)	
2	200-400	0.2 (1.223)	-0.17 (-1.019)	
3	400-600	0.18 (0.92)	-0.23 (-1.191)	
4	600-800	0.5 (1.79)†	0.4 (1.422)	-0.81 (-2.033)*
5	800-1000	0.53 (1.823)†	0.93 (3.168)**	-1.16 (-2.81)**
6	1000-1200	0.91 (3.026)**	1.35 (4.527)***	-1.84 (-4.354)***
7	1200-1400	0.65 (2.139)*	1.05 (3.487)***	-1.47 (-3.428)**
8	1400-1600	0.86 (2.829)**	0.96 (3.15)**	-1.76 (-4.096)***
9	1600-1800	0.53 (1.769)†	0.7 (2.314)*	-1.3 (-3.063)**
10	1800-2000	0.56 (1.868)†	0.55 (1.826)†	-1.29 (-3.015)**

Table 3.3: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*er* vs. *der*) and order of mention (1st vs. 2nd). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

To break down the interaction, we calculated individual analyses for each type of pronoun entering order of mention as a predictor (see Table 3.4). These analyses showed that for the personal pronoun there were more looks to the first-mentioned entity than to the second-mentioned entity. This effect was significant in time windows 8 to 10 (1400 – 2000 ms). In time window 3, there was a marginally significant effect of mention; however, as the interaction was not significant in this time window, there were not more looks overall to the first-mentioned entity for the personal pronoun than for the d-pronoun.

For the d-pronoun we found significantly more looks to the second-mentioned than to the first-mentioned entity across time windows 5 to 9 (800 – 1800 ms). The main effect stayed marginally significant in time window 10 (1800 – 2000 ms). Thus, the interaction was due to there being significantly more looks to the second-mentioned antecedent for the d-pronoun than for the personal pronoun, and later more looks to the first-mentioned entity for the personal pronoun than for the d-pronoun, indicating an effect of order of mention.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	-0.07 (-0.725)	0.05 (0.615)
2	200-400	-0.28 (-1.162)	-0.06 (-0.245)
3	400-600	-0.49 (-1.756)†	0.03 (0.125)
4	600-800	-0.41 (-1.402)	0.4 (1.463)
5	800-1000	-0.24 (-0.807)	0.93 (3.143)**
6	1000-1200	-0.49 (-1.647)	1.35 (4.495)***
7	1200-1400	-0.41 (-1.342)	1.05 (3.547)***
8	1400-1600	-0.81 (-2.656)**	0.96 (3.142)**
9	1600-1800	-0.61 (-2.018)*	0.7 (2.314)*
10	1800-2000	-0.74 (-2.457)*	0.55 (1.805)†

Table 3.4: Results of the individual time course analyses for each type of pronoun (*er* and *der*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

Taken together, we observed a relatively late first-mention preference for the personal pronoun and an earlier second-mention effect for the d-pronoun.

3.1.3.3. Discussion

The eye-movement measures showed a clear effect of order of mention. After hearing the personal pronoun *er*, participants fixated more often pictures of the first-mentioned than the second-mentioned character of the antecedent sentence, while in the d-pronoun condition the participants fixated pictures of the second-mentioned more often than the first-mentioned image. The off-line task revealed the same direction of results: a first-mention preference for *er* and a marginally significant second-mention preference for *der*. The time course analyses revealed a first effect of the interaction of order of mention and pronoun type in the 600-800 ms segment. However, there was no main effect of pronoun type or order of mention. The order of mention effect emerged later, in the 800-1000 ms segment, in which there was also an interaction of order of mention and pronoun type. This indicates that the interaction might have been driven by the materials.[37]

Our findings are in line with the assumption that personal and d-pronouns have different co-reference functions, in that personal pronouns prefer first-mentioned topical antecedents,

[37] Note that half of the items consisted of animate antecedents and the other half of inanimate antecedents. Both materials entered this overall analysis. In this early segment, the effects differed for animate and inanimate materials giving rise to this interaction (see chapter 3.3.).

while d-pronouns prefer second-mentioned non-topical antecedents (Ariel, 1990, 2001; Bosch, et al., 2007a; Bosch, et al., 2003; Kaiser & Trueswell, 2004, 2008; Lambrecht, 1994; Levinson, 1987, 1991; Schumacher, et al., in prog.; Wilson, 2009). Interestingly, the first-mention preference for the personal pronoun was at ceiling in the off-line task, reflecting a strong first-mention preference in final interpretations. During the eye-tracking task this preference emerged quite late compared to the earlier second-mentioned non-topical preference for the d-pronoun, indicating that there was a higher degree of ambiguity for the personal pronoun. This is in line with the assumption that the d-pronoun is marked for non-topical co-reference relations while the personal pronoun is unmarked and therefore more ambiguous (Ahrenholz, 2007; Bosch & Umbach, 2007b; Schumacher, et al., in prog.; Zifonun, et al., 1997). At the same time the on-line results reject the prediction made based on the Givenness Hierarchy (Gundel, 2003; Gundel, et al., 1993) which according to observations for English personal and demonstrative pronouns, suggests that personal pronouns are marked for topicality (due to the *in focus* requirement of its antecedent), while d-pronouns are more flexible in their co-reference relations. The direction of the off-line results could be explained with Gundel's hypothesis. However, if so, it is unclear why the same pattern did not emerge in the on-line data[38].

The pattern of results obtained could be explained in line with Kaiser and Trueswell's Form-Specific Multiple-Constraints Approach (2008) according to which personal and d-pronouns are sensitive towards different factors. For German, Wilson (2009) has claimed that personal pronouns are sensitive to syntactic and discourse factors, while d-pronouns are only sensitive to discourse factors. Our results also fit this explanation. The fact that the antecedents in the current experiment were not differentially case-marked might have caused a higher level of ambiguity for personal pronouns. This suggests that the resolution of personal pronouns might depend on both syntactic and discourse factors, while the d-pronouns were resolved on the basis of the order of mention/discourse information. However, this explanation fits the off-line results less well, since a highly significant first-mention preference was observed there. Furthermore, the fact that Wilson (2009) did not find a preference for *er* after SVO antecedent structures cannot be accounted for by this explanation, which would predict topical first-mentioned subjects of SVO sentences to be very prominent antecedent candidates for the resolution of the personal pronoun.

[38] Moreover, as will be shown in chapter 3.4, the off-line pattern for *der* was due to variation in individual preferences.

Taken together, German personal and d-pronouns are found

- to have different co-reference functions after canonical antecedent structures, in that personal pronouns prefer first-mentioned topical antecedents, and d-pronouns second-mentioned non-topical antecedents
- to be differently marked for these co-reference preferences, in that the d-pronoun is marked for non-topical co-reference and the personal pronoun is an unmarked form

The question I address in the next section is whether the resolution of personal and d-pronouns in Dutch is comparable to the resolution in German. Do personal and d-pronouns also have these two different co-reference functions in Dutch, and are they also differently marked for these preferences? By investigating these issues, we will get a clearer picture of the possible cross-linguistic differences and similarities with regard to pronoun resolution preferences.

3.1.4. Pronoun resolution in Dutch

This experiment replicates the German experiment.

3.1.4.1. Methods

Participants

Twenty-eight Dutch native speakers (26 female, 2 male) participated in the study. The participants were students at the Radboud University Nijmegen. They were aged between 18 and 25 years (mean = 19.79; SD = 1.86). All participants were tested individually and were paid for their participation. All participants had normal or corrected-to-normal vision.

Experimental Tasks

The experimental tasks and the order of presentation were identical to those of the German experiment.

Materials and design

The items of the German experiment were translated into Dutch. This was easily possible except for the Dutch NP *potlood* (pencil) which has neuter gender. Therefore, we changed it to *viltstift* (felt tip) which has common gender. The picture was changed accordingly. In (57)a Dutch example of an experimental item in both conditions is given.

(57) Conditions: Sample Item

a. Personal Pronoun Condition
De kast is zwaarder dan de tafel. **Hij** is afkomstig uit een meubelwinkel in België. De sofa zal volgende week geleverd worden.
The cupboard is heavier than the table. ***It*** *[P] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

b. D-pronoun Condition
De kast is zwaarder dan de tafel. **Die** is afkomstig uit een meubelwinkel in België. De sofa zal volgende week geleverd worden.
The cupboard is heavier than the table. ***It*** *[D] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

Procedure

The procedure was the same as for the German experiment. Small adaptations were made which are reported in the following. Each item was digitally recorded to computer. The Dutch items were read out loud by a female native Dutch speaker. Intonational stress on the referents and the pronouns was avoided.

We used exactly the same pictures as for the German experiment, except that the picture of a *pencil* was replaced with a picture of a *felt tip*. The presentation of the pictures was changed with regard to the German experiment in that we rotated the positions on the screen for all three pictures, i.e. the first target picture, the second target picture and the discourse-related non-target picture. This was changed in order to maximally counterbalance the position of the pictures (not only between target pictures, but including the discourse-related non-target picture). The target pictures could thus also appear in the lower center position, and the discourse-related non-target could appear in the upper positions as well.

Participants' eye movements were recorded on a portable eye-tracker, namely SR Research EYELINK II. The dominant eye was recorded. A sampling rate of at least 250-Hz was used which monitored gaze locations every 4 ms. The calibration of the camera which links the position of the eyes with a certain location on the screen, ensured that spatial accuracy was at least 0.5°. The accuracy of the responses to the content questions was comparably high to the German experiment with 97% correct answers (24 questions; mean correct answers = 23.29, SD = 0.84).

3.1.4.2. Results

Forced Choice Questionnaire

Figure 3.6 and Table 3.5 show the results for the off-line task in Dutch. The figure reveals a similar pattern of results as for the German native speakers. The first-mentioned antecedent was chosen as co-referential with the personal pronoun *hij* in 95% of the cases. The d-pronoun *die* showed a tendency to be more often resolved towards the second-mentioned antecedent (57%).

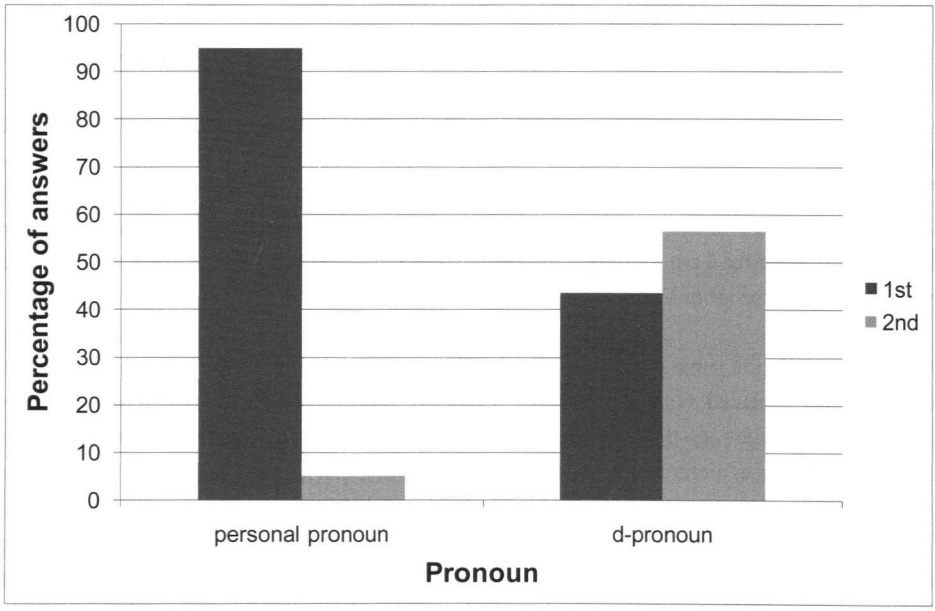

Figure 3.6: L1 Dutch Results on the Forced Choice Questionnaire; pronoun being coreferential either with the first- or the second-mentioned entity (1st, 2nd) in the two conditions (personal pronoun, d-pronoun)

	1st	2nd
personal pronoun (*hij*)	94.94% (319)	5.06% (17)
d-pronoun (*die*)	43.45% (146)	56.55% (190)

Table 3.5: L1 Dutch Results on Forced Choice Questionnaire; pp = 28; items = 24

There was a significant association between the type of pronoun and whether the first- or second-mentioned antecedent was chosen as the coreferential entity, χ^2 (1) = 209.00 $p<.001$. The odds for choosing the first-mentioned entity were 24.42 times higher for personal than for d-pronouns.

In order to further investigate whether the distribution was different from a 50%-chance level which would indicate no preference, two more Chi Square tests were conducted in which the distribution of responses for one was compared to the no-preference distribution. These tests revealed that the association of *hij* and the antecedent choice was significantly different from not showing a preference, χ^2 (1) = 170.07, p<.001. The first-mentioned entity was 18.76 times more likely to be chosen for *hij* than the 50%-chance level would predict. For *die* the distribution of responses was marginally significantly different from the no-preference level, χ^2 (1) = 2.893, p = .052[39].

Taken together, the German and Dutch off-line results are very similar. They both reveal a strong first-mention preference for the personal pronoun, and a marginal second-mention preference for the d-pronoun.

Visual-World Eye-Tracking

Data Analysis

The data analysis was conducted identically to the German native experiment. Although we used an ASIO sound card in this experiment, we kept the data cleaning method (as explained in the previous section) to make sure that the eye-tracking data maximally informed us about the resolution behavior due to the pronoun. Therefore, we excluded 21 target looks (< 1%) that had started before pronoun onset (-200 till 0 ms), resulting in 4605 analyzable samples.

Overview of the eye movement patterns

Figure 3.7 shows the proportion of fixations over time to both antecedent pictures (1st & 2nd) in both conditions *hij* and *die*. Overall the looks start rising at around 200 ms to both target pictures across conditions.

[39] p>.05

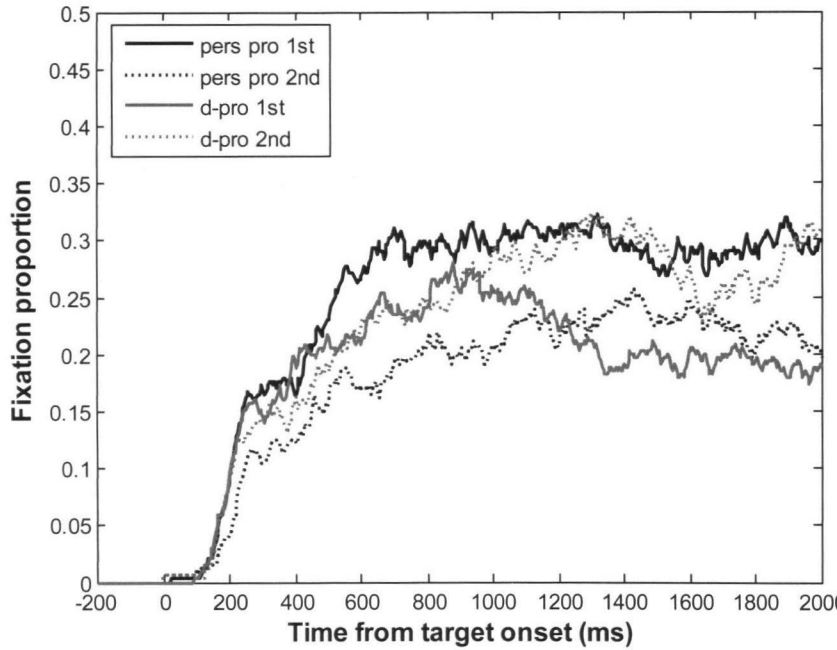

Figure 3.7: Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Personal pronoun. For the personal pronoun *hij*, around 500 ms after pronoun onset, there is a sharp rise in looks to the first mentioned entity with a peak at around 700 ms which lasts until about 1400 ms. In the time segment between 1400 and 1600 ms the looks to the first-mentioned entity for *hij* decrease, but increase again towards the end (1600 – 2000 ms), while the looks to the second-mentioned target drop. This pattern of results suggests an early and clear first-mention preference for the personal pronoun *hij*.

D-pronoun. For the d-pronoun *die*, looks to both targets rise until about 1000 ms after the pronoun onset, at which point there is an increase in looks to the second-mentioned target and a drop of looks to the first-mentioned target. At around 1400 ms the target looks to the second-mentioned entity start to drop until about 1600 ms after which they increase again. This pattern of results reflects that ambiguity persists relatively long for the d-pronoun *die* compared to the personal pronoun *hij*, which is then resolved towards the second-mentioned entity.

Taken together, order of mention had an inverse influence on the two types of pronouns, with an earlier first mentioned effect for the personal pronoun *hij* and a later second mention effect for the d-pronoun *die* which indicates a relatively low degree of ambiguity for the personal pronoun *hij*.

Statistical Analysis of the eye movement patterns

The statistical analysis was conducted as with the German data. Two linear mixed models, the *simple model* and the *interaction model*, were calculated for each of the ten 200 ms time slices from pronoun onset till 2 seconds after the pronoun onset. The *simple model* contained the terms for the main effects of order of mention (*mention*) and type of pronoun (*condition*), while the interaction term between order of mention and type of pronoun was added to the *interaction model*. A likelihood test was calculated to determine which model better explained the variance in the data for each time window (Baayen, 2008) and the interaction was only added when it significantly better predicted the outcome of the data.

The *interaction model* significantly better predicted the outcome starting at 600 ms after pronoun onset and lasting till 2000 ms (see Table 0.5 in the Appendix). This is in line with the observation that the first-mention preference for *hij* emerges around that time.

Time window	in ms	Fixed predictors		
		pronoun condition	order of mention	interaction
1	0-200	-0.07 (-1.135)	-0.08 (-1.187)	
2	200-400	-0.17 (-0.997)	-0.38 (-2.252)*	
3	400-600	-0.09 (-0.473)	-0.4 (-2.099)*	
4	600-800	0.5 (1.716)†	-0.08 (-0.286)	-0.98 (-2.388)*
5	800-1000	0.49 (1.672)†	0.05 (0.161)	-1.08 (-2.588)**
6	1000-1200	0.58 (1.927)†	0.52 (1.707)†	-1.34 (-3.126)**
7	1200-1400	0.99 (3.276)**	1.11 (3.678)***	-1.8 (-4.23)***
8	1400-1600	0.82 (2.74)**	0.98 (3.278)**	-1.45 (-3.435)***
9	1600-1800	0.84 (2.849)**	0.53 (1.797)†	-1.14 (-2.747)**
10	1800-2000	1.09 (3.689)***	1.01 (3.422)***	-1.82 (-4.345)***

Table 3.6: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*hij* vs. *die*) and order of mention (1st vs. 2nd). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

Concerning the model estimates for the fixed effects and their significance levels (see Table 3.6), we observed a significant main effect of order of mention in time windows 2 and 3

(200 – 600 ms). The negativity of the beta coefficient indicates that there were overall more looks to the first-mentioned entity than to the second. In time windows 4 and 5 (600 – 1000 ms), we found a significant interaction between order of mention and pronoun type, but there was only a marginally significant main effect of type of pronoun. The interaction stayed highly significant in the remaining time windows (1000 – 2000 ms) and the main effects of pronominal condition and order of mention were both either marginally significant or significant across these same time windows. The positivity of the beta coefficient indicated that overall there were more looks for the personal than for the d-pronoun; at the same time there were overall more looks to the second-mentioned target picture than to the first-mentioned target picture.

To break down the interaction effect, individual comparisons were conducted for the two types of pronouns with order of mention as a predictor (see Table 3.7). For the *hij* condition the analyses revealed a significant main effect of order of mention across time windows 2 to 10 (200 – 2000 ms), except for time window 8 (1400 – 1600 ms). The beta coefficient for mention was negative indicating a first-mention preference.

For the d-pronoun *die*, the main effect of order of mention was either marginally significant or significant across time windows 6 to 10 (1000 – 2000 ms). The positivity of the beta coefficient for mention reflected a second-mention preference.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	-0.14 (-1.507)	-0.02 (-0.215)
2	200-400	-0.54 (-2.295)*	-0.21 (-0.886)
3	400-600	-0.72 (-2.608)**	-0.09 (-0.323)
4	600-800	-1.06 (-3.648)***	-0.08 (-0.284)
5	800-1000	-1.03 (-3.527)***	0.05 (0.158)
6	1000-1200	-0.82 (-2.693)**	0.52 (1.693)†
7	1200-1400	-0.69 (-2.257)*	1.11 (3.697)***
8	1400-1600	-0.47 (-1.56)	0.98 (3.285)***
9	1600-1800	-0.62 (-2.04)*	0.53 (1.823)†
10	1800-2000	-0.81 (-2.66)**	1.01 (3.453)***

Table 3.7: Results of the individual time course analyses for each type of pronoun (*hij* and *die*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

Taken together, the individual analyses for the two Dutch pronouns revealed a strong first-mention preference for the personal pronoun *hij* and a very weak second-mention preference for the d-pronoun *die*.

3.1.4.3. Discussion

As a whole, the results of the eye-tracking study showed that participants' interpretation of *hij* and *die* was influenced by the order of mention of the antecedent candidates. The off-line results showed the same preferences: a first-mention preference for *hij* and a marginally significant second-mention preference for *die*. The first-mention effect for *hij* combined with the second-mention effect for *die* is compatible with what was observed for the German personal and d-pronoun. Thus, we found cross-linguistic evidence for the assumption that personal and d-pronouns have different co-reference functions (Bosch, et al., 2003; Comrie, 1994; Diessel, 1999; Lambrecht, 1994). While personal pronouns preferred first-mentioned topical antecedents, d-pronouns preferred second-mentioned non-topical antecedents. However, the Dutch data is not consistent with the observation in the German data that the d-pronoun is marked for non-topicality while the personal pronoun is more neutral. This seems to be an inherent property of the German d-pronoun.

According to Gundel's Hypothesis, the personal pronoun was resolved early and clearly towards the first-mentioned entity in the on- and off-line task, because it may be marked for topical reference in comparison to the relatively unspecified d-pronoun which may be resolved towards both entities. But the fact that we found a first-mention effect for *hij* emerging already in time window 2 (200 – 400 ms) rather points to the possibility that more general top-down discourse effects influenced its resolution. In pronoun comprehension, top-down processes have been claimed to represent processes initiated by the discourse context, while bottom-up processes represent processes triggered by the pronoun, such as search processes for an adequate antecedent (e.g. Stevenson, 1996). Thus, the early first-mention preference for the personal pronoun might have arisen because the first-mentioned antecedent occupied a privileged position in the discourse. The subsequent mention of the personal pronoun triggered bottom-up processes which enhanced this interpretation, while the d-pronoun did not. The d-pronoun took some time to be resolved, maybe because the bottom-up processes it triggered modified the top-down interpretation.

In general, our results are difficult to interpret with regard to the predictions made by theories of reference[40]. Both Ariel (1990, 2001) and Levinson (1987, 1991) classify the

[40] The Expectancy Hypothesis by Arnold (1998, 1999, 2001, 2008) is not discussed here, because we are not able to falsify it by our experiments. It predicts that co-reference relations between anaphoric expressions and their antecedents are determined by a likelihood estimation of discourse expectation. The frequency with which these

form fullness of anaphoric expressions by the richness of semantic content they carry and their phonological size. This would predict the personal pronoun to be a fuller form than the d-pronoun in Dutch, since the Dutch personal pronoun differentiates between three types of gender, whereas the d-pronoun differentiates only two. We therefore think that the classification criteria for form fullness formulated by the theories of reference are not well suited criteria to classify Dutch d-pronouns. In chapter 2.2.2, we discussed Broeder's (1991; p.128) observation that personal pronouns are frequently used in production to signal topic shift, whereas d-pronouns are not as likely to occur in this function. This distribution is not found in our processing data. Rather we find that the personal pronoun is resolved earlier than the d-pronoun[41].

Our results are compatible with Kaiser and Trueswell (2004) who also found a first-mention preference for *hij* and a second-mention preference for *die* when presenting them after SVO-antecedent structures in a sentence completion task and in a visual-world eye-tracking task. They attribute these findings to saliency operationalized as grammatical role: "*hij* prefers subjects and *die* prefers objects" (p.146). But since our materials did not include a grammatical subject-object distinction, grammatical role information cannot be the sole source. In the light of the linguistic accounts, we might assume that it is rather topicality that underlies the effect, although we cannot exclude the importance of first-mention (these two often coincide – non-accidently). In this sense, *hij* is more likely to be resolved towards the first-mentioned topical entity, while *die* prefers the second-mentioned non-topical entity.

In general, the order of mention of the antecedent candidates influenced the resolution of personal and d-pronouns. This shows that even when syntactic subject-object distinctions are not available in the input, order of mention information is used to resolve pronouns. Although it is very likely that the first-mention preference was due to topicality, and the preference for the second-mentioned entity occurred because it was non-topical, the relative influence of the two factors order of mention and topicality cannot be torn apart, since in the materials of the experiments reported here, the first-mentioned entity was always topical and the second-mentioned entity non-topical. Therefore, in a second set of German-Dutch experiments, we investigated whether the resolution preferences of personal and d-

---

relations have been encountered in the language input determines the preferred co-reference relations. Although this hypothesis provides an explanation for cross-linguistic as well as L1-L2 resolution differences since we do not have any reference data on the frequency in the input, any empirical results could be argued to appear due to differences in cue validity.

[41] Since Broeder (1991) explains his observations on the basis of the disambiguating function of pronominals due to their more precise encoding of biological gender, we might not find this here, because the animate antecedents in our materials were all males; thus the disambiguating function of the personal pronoun did not show up.

pronouns varied when following non-canonical antecedent structures. We asked whether the information status of the antecedent candidates had an influence on the order of mention preferences for personal and d-pronouns. These experiments are reported in the next section.

3.2. Pronoun Resolution after Non-Canonical Antecedent Structures[42]

3.2.1. Introduction

In chapter 3.1, we saw that when a canonical antecedent structure contains more than one possible antecedent, personal pronouns are preferentially resolved towards the first-mentioned topical entity, while d-pronouns are resolved towards the second-mentioned non-topical entity. These preferences were found after double nominative comparative antecedent structures, which lack a formal subject-object distinction. Thus, compared to earlier resolution studies on personal and d-pronouns (Bosch, et al., 2007a; Bosch, et al., 2003; Bosch & Umbach, 2007b; Kaiser & Trueswell, 2004, 2008; Schumacher, et al., in prog.; Wilson, 2009), whose antecedent structures consisted of a grammatical subject antecedent and an object antecedent, and where the results can at least be partially attributed to influences of grammatical role, our study found that the position of the antecedent candidates was crucial for disambiguation. However, as we take position (as well as grammatical role) to be an indicator of the information structure of a sentence and the pragmatic functions of its constituents, in that topics generally appear in initial positions and non-topical entities in later positions, this finding corroborates all accounts that postulate that personal pronouns should be resolved towards the topical and d-pronouns towards the non-topical entity (Ariel, 1990, 2001; Bosch & Umbach, 2007b; Comrie, 1994; Diessel, 1999; Lambrecht, 1994; Levinson, 1987, 1991). The accounts differ with regard to the reasons they assume account for such differences. Perhaps topichood entails a higher degree of accessibility (Ariel, 1990, 2001). Or perhaps, motivated by Neo-Gricean pragmatic reasons, the personal pronoun is a minimal yet recognizable form and as such allows for co-reference to the preferred reference maintaining entity, which is the topical entity, whereas the d-pronoun initiates a non-coreferential/disjoint interpretation, and is therefore resolved towards the non-topical entity (Levinson, 1987, 1991). Or as others have pointed out (Bosch, et al., 2003; Comrie, 1994; Diessel, 1999; Lambrecht, 1994), it could simply be a general rule that personal pronouns are resolved towards the topical entity, while d-pronouns are resolved towards the non-topical/not yet topical entity, for the fact that they are only to be differentiated on this information structural dimension (all other things, like mental activation states and identifiability being equal).

It is only until recently that psycholinguists have started to investigate the influence of multiple factors and their interaction on the resolution preferences of different kinds of

[42] Parts of the research presented in this chapter will appear in Ellert, M., Järvikivi, J. & Roberts, L. (to appear). Information structure affects the resolution of the subject pronouns er and der in spoken German discourse. In Sarda, L., Thomas, S.C. & Fagard, B. [eds.] *Linguistic and Psycholinguistic Approaches to Text Structuring*. Amsterdam: John Benjamins.

pronouns. While all earlier studies were confined to English personal pronouns and to the influence of single factors, such as first-mention (Gernsbacher, 1989) or subject preference (Crawley, et al., 1990; Frederiksen, 1981), more recent studies have turned to flexible word-order languages such as Finnish and German and have applied real time processing methods to disentangle the influences of grammatical role and position on pronoun interpretation (Bouma & Hopp, 2007; Järvikivi, et al., 2005; Kaiser & Trueswell, 2008; Schumacher, et al., in prog.; Wilson, 2009). These studies used sentences like (58) (taken from Wilson, 2009), where the grammatical antecedent roles were switched by case marking, from SVO (58)a to OVS (58)b. As stated earlier, since changing the word order of a sentence also affects its information structure, the resolution preferences observed after OVS structures in these studies might have been caused by information structure differences rather than to features inherent to the pronominal forms.

(58)

a. Der Kellner erkennt den Detektiv, als das Bier umgekippt wird. ***Er/Der*** ist offensichtlich sehr fleißig.
The waiter$_{SVO}$ recognizes the detective$_{ACC}$ as the beer is tipped over. ***He*** *[P/D] is clearly very hard working.*

b. Den Kellner erkennt der Detektiv, als das Bier umgekippt wird. ***Er/Der*** ist offensichtlich sehr fleißig.
The waiter$_{ACC}$, the detective$_{SVO}$ recognizes as the beer is tipped over. ***He*** *[P/D] is clearly very hard working.*

In this section, I review the results obtained with a focus on the non-canonical OVS structures (see chapter 3.1 for a review of results in the SVO conditions). Table 3.8 shows that the results obtained across studies do not present a consistent picture. While some report a primary influence of grammatical role, others report an influence of topicality on pronoun resolution. Considering only the resolution preference found after OVS structures, with the exception of the Schumacher et al. (in prog.) study, there are no first-mention preferences. This is perhaps because the first-mentioned topical antecedent in OVS word order is not as accessible as in the SVO topic-comment structure, because of the focused, second-mentioned syntactic subject in OVS. It may be that it is the function of the non-topic that is influential (i.e., comment vs. focus) in pronoun resolution, rather than the distinction between topic and non-topic.

			Antecedent Structures	
		Pronouns	**SVO**	**OVS**
Schumacher et al. (in progress)	German	**personal pronoun**	1st	2nd
		d-pronoun	2nd	**1st**
Wilson (2009)	German	**personal pronoun**	no preference	no preference
		d-pronoun	2nd	**2nd**
Bosch et al. (2007a)	German	**personal pronoun**	no preference	no preference[43]
		d-pronoun	2nd	**no preference**[44]
Bouma & Hopp (2007)	German	**personal pronoun**	1st	2nd
Kaiser & Trueswell (2008)	Finnish	**personal pronoun**	1st	2nd
		demonstrative pronoun	2nd	**2nd**
Järvikivi et al. (2005)	Finnish	**personal pronoun**	1st	2nd → no preference

Table 3.8: Overview of results from previous visual-world studies on the resolution of personal and d-pronouns/demonstrative pronouns in German and Finnish (resolution preference for the d-pronoun in **bold**)

Focus information, operationalized by the use of clefted antecedent structures, has been shown to influence pronoun resolution in English (Arnold, 1998, 1999, 2001; Cowles, 2003; Cowles, Walenski, & Kluender, 2007). The clefted element in such structures is in contrastive focus which is different from informational focus. The latter type of focus represents information which is unpredictable and predicates the topic (Lambrecht, 1994). Contrastive focus on the other hand is an element out of a set of alternatives which are presupposed to exist. Cowles et al. (2007) used a cross-modal priming task to investigate the influence of topic and focus information on resolution preferences of personal pronouns. They used a 3x2 design with three different types of antecedent sentence types (discourse topic, sentence topic, clefted focus) and two positions for the topical/focused entity to appear (first vs. second), as in (59). They observed that the pronoun was resolved towards the more prominent entity (discourse topic, sentence topic, clefted focus) in all three conditions. Thus, topic information as well as focus information increased the

[43] only two items in this condition
[44] only two items in this condition

cognitive prominence of the antecedents. Furthermore, the focused constituent was only significantly preferred when it carried contrastive intonation. The results of this study underline the importance of different pragmatic functions (and not only topicality) on the resolution of pronouns.

(59)

1. *Discourse topic*
   Anne wanted to see the new movie with Sarah.
   a. So, **Anne** called Sarah. (first-mentioned)
   b. When Sarah came home, **Anne** called. (second-mentioned)
2. *Sentence topic*
   A new movie opened in town.
   a. So, **Anne** called Sarah. (first-mentioned)
   b. When Sarah came home, **Anne** called. (second-mentioned)
3. *Clefted focus*
   A new movie opened in town.
   a. It was **Anne** who called Sarah. (first-mentioned)
   b. The one who called Sarah was **Anne.** (second-mentioned)

   Target Sentence
   But later that night, **she** couldn't go to the movie after all.

Example item in Cowles et al. (2007: p.7)

In the present study, we addressed the question of whether pragmatically focusing one of the antecedents had an effect on the resolution of personal and d-pronouns in German and Dutch. More specifically, we wanted to find out what the functions of personal and d-pronouns in such marked contexts are. As in chapter 3.1, we used comparative antecedent structures to test this assumption trying to minimize grammatical role effects. While the first-mentioned entity in such structures (as in (60)) belongs to the topic, the second-mentioned entity is focused. Since the sentence initial comparative constituent implies that something must be *heavier than the table*, the second-mentioned entity carries contrastive focus in that it is activated from a set of possibilities.

(60) German: Schwerer als **der Tisch** ist **der Schrank.** ***Er/Der*** stammt aus einem Möbelgeschäft in Belgien.
Dutch: Zwaarder dan **de tafel** is **de kast.** ***Hij/Die*** is afkomstig uit een meubelwinkel in België.

English: *Heavier than* ***the table*** *is* ***the cupboard.*** ***It*** *[P/D] originates from a furniture store in Belgium.*

As before, the experiment was undertaken in both German and Dutch.

3.2.2. The present study

Two issues are addressed in the present study. First, are the preferences after non-canonical structures different from the preferences after canonical structures? Second, which role does focus, as an antecedent property, play? The results of previous pronoun resolution studies (Bosch, et al., 2007a; Bouma & Hopp, 2007; Järvikivi, et al., 2005; Kaiser & Trueswell, 2008; Wilson, 2009) with the exception of one (Schumacher, et al., in prog.), have not obtained first-mentioned resolution preferences after non-canonical OVS antecedent structures irrespective of type of pronoun. This hints at a possible influence of the focus function of the second-mentioned antecedent on pronoun resolution, which we explicitly test in the current experiments. The same materials as those in the experiments reported in 3.1 were used, but the antecedent sentences were inverted from a canonical comparative structure to a non-canonical comparative structure: *Comparative-NP2-verb-NP1* (as in (60)).

Theoretical accounts of reference (Ariel, 1990, 2001; Bosch, et al., 2003; Levinson, 1987, 1991) predict an asymmetric resolution pattern for the two types of pronouns, irrespective whether salience is achieved via pragmatic topic or focus encodings. A non-canonical word order influence, on the other hand, as has been assumed to be visible in the previous OVS findings, would make the focused entity particularly prominent for future reference, and would therefore predict the same resolution preferences for the two types of pronouns.

The research questions are thus as follows:

1. How does a non-canonical antecedent structure affect the resolution preferences of personal and d-pronouns in comparison to canonical antecedent structures?
2. How does focus influence the resolution preferences for the two types of pronouns in this context?
3. Is there cross-linguistic evidence for general resolution preferences for personal and d-pronouns after the inverted comparative antecedent structures between German and Dutch?

3.2.3. Pronoun resolution in German

The first experiment investigates the resolution of personal and d-pronouns in German in order to find out if and how the inversion of the antecedent structure influences the preferences.

3.2.3.1. Methods

Participants

Thirty-two German native speakers (18 female, 14 male) participated in the study. They were aged between 17 and 25 years (mean = 19.47; SD = 2.22). All participants were tested individually and were paid for their participation. All participants had normal or corrected-to-normal vision. Seven participants were students at the Educational University Karlsruhe, and twenty-five participants were pupils at the high school "Goethe Gymnasium Karlsruhe" (grades 12 and 13[45]).

Experimental Tasks

The experimental tasks were identical to those used of the experiments in chapter 3.1. The entire experimental session took approximately 45 min.

Materials and design

The stimuli were the same as used in the German experiment reported in chapter 3.1 (see Table 0.1 in the Appendix), only the comparative sentences were changed from *NP1-verb-comparative-NP2* to *Comparative-NP2-verb-NP1* (as in (61)). This included all 24 experimental items, as well as 24 comparative fillers.

(61) Conditions: Sample Item

a. Personal Pronoun Condition
Schwerer als der Tisch ist der Schrank. **Er** stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.
Heavier than the table is the cupboard. ***It*** *[P] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

[45] Note, that German pupils who pass high school with a diploma enabling them to study at a University, need to attend school for 13 years, usually starting at age 6 or 7 and ending at age 19 or 20. Therefore, German high school students of grades 12 and 13 are as old as undergraduate students in the Netherlands.

b. D-pronoun Condition
Schwerer als der Tisch ist der Schrank. **Der** stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.
*Heavier than the table is the cupboard. **It** [D] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

Procedure
The procedure was essentially the same as for the German experiment in chapter 3.1. The items were read out loud by the same male German speaker as in chapter 3.1. The pictures were the same as in the German experiment in chapter 3.1, only that the position of all pictures was rotated among all three available positions.

The participants' eye movements were recorded on a portable eye-tracker, namely SR Research EYELINK II. The dominant eye was recorded. A sampling rate of at least 250-Hz was used which monitored gaze locations every 4 ms. The calibration of the camera which links the position of the eyes with a certain location on the screen, ensured that spatial accuracy was at least 0.5°.

The accuracy of the responses to the content questions was very high with 94% correct answers (24 questions; mean correct answers = 22.61, SD = 0.97).

3.2.3.2. Results

Forced Choice Questionnaire
As can be seen in Figure 3.8 and Table 3.9, the German participants chose the second-mentioned antecedent as being coreferential with the pronoun in 99% of all cases irrespective of whether it was a personal pronoun or a d-pronoun.

Since, two cells (personal pronoun – 1st, d-pronoun – 1st) had an expected count less than 5 (namely 2), Fisher's exact test was conducted on the data, and as expected showed no significant association between the type of pronoun and the type of antecedent chosen, P(Fisher, 2-sided) = .688[46]. The second-mentioned entity was chosen with an equal frequency for the two types of pronouns. This means that there was a very strong overall preference for the entity in focus.

[46] P(Fisher, 2-sided)>.05

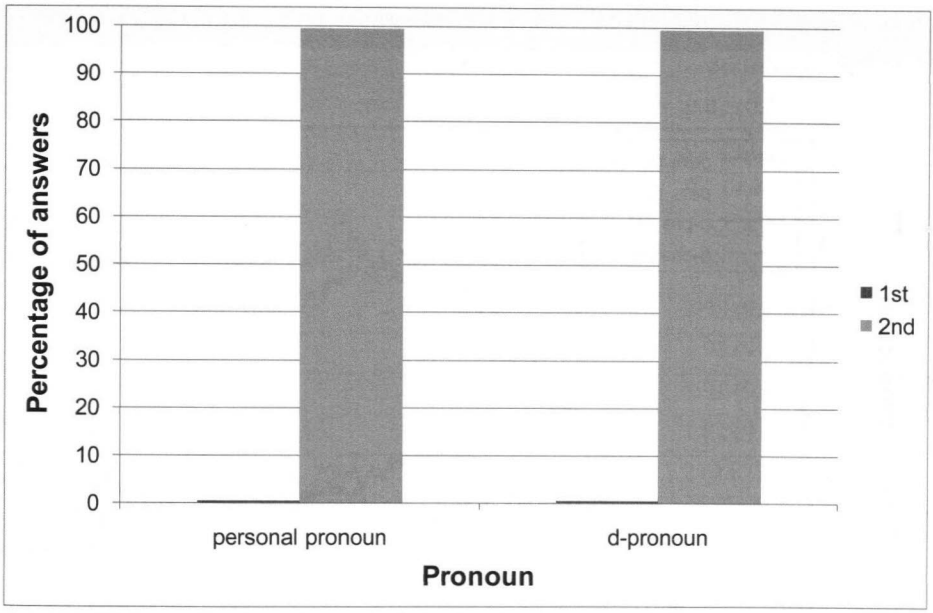

Figure 3.8: L1 German Results on the Forced Choice Questionnaire; pronoun being coreferential either with the first- or the second-mentioned entity (1st, 2nd) in the two conditions (personal pronoun, d-pronoun)

	1st	2nd
personal pronoun	0.52% (2)	99.48% (382)
d-pronoun	0.52% (2)	99.48% (382)

Table 3.9: L1 German Results on Forced Choice Questionnaire; pp = 32; items = 24

Visual-World Eye-Tracking

Data Analysis

Due to data cleaning procedure as explained for the experiments in chapter 3.1, we excluded 70 target looks (1%) which had started before pronoun onset (-200 till 0 ms), resulting in 4914 analyzable samples in total.

Overview of the eye movement patterns

As can be seen in Figure 3.9, in both conditions the looks to the two targets start to rise sharply at about 200 ms, which reflects the time which is needed to launch a saccade (Matin, et al., 1993). But the target looks to the second-mentioned entity continue to increase until they reach a 40%-level, while the proportions of fixations towards the first-mentioned entity stay below a 20%-level. This difference persists until 2 seconds after the

onset of the pronoun and reflects a strong second-mention preference for the two types of pronouns.

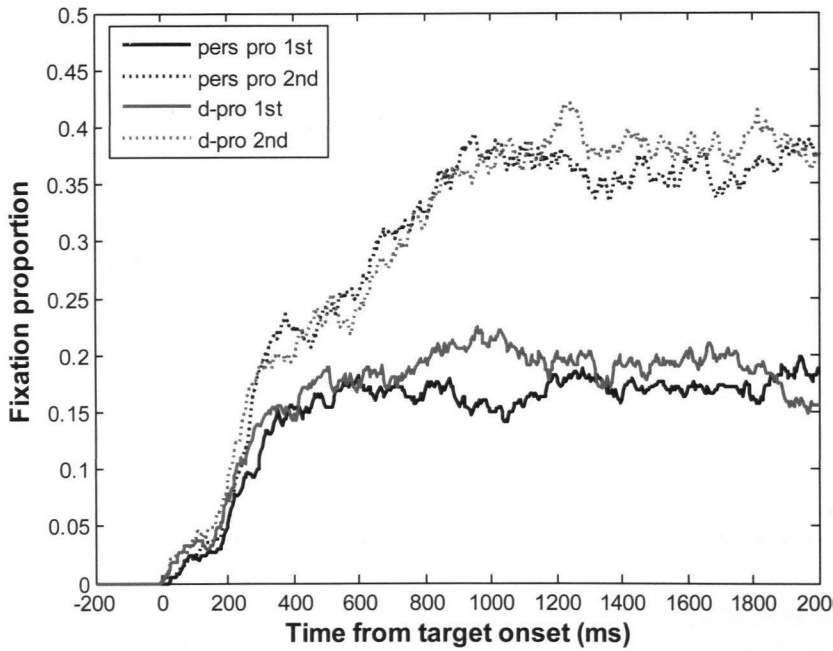

Figure 3.9: Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Statistical Analysis of the eye movement patterns

The *simple model* explains the data better than the *interaction model* for all time windows. This means that the model predicts the outcome better when the interaction term is not included. Thus, contrary to the results obtained for personal pronouns and d-pronouns after canonical structures (chapter 3.1), after non-canonical structures the resolution of the two pronouns was not affected differently by order of mention.

The analysis of the fixed effects (see Table 3.10) showed a highly significant main effect of order of mention and at the same time no effect of type of pronoun across time windows 2 to 10 (200 – 2000 ms). The positive sign of the beta coefficient indicated an overall second-mention preference.

Time window	in ms	Fixed predictors	
		pronoun condition	order of mention
1	0-200	-0.05 (-0.552)	0.02 (0.259)
2	200-400	0 (0.009)	0.52 (3.255)***
3	400-600	0.02 (0.112)	0.7 (3.727)***
4	600-800	-0.06 (-0.317)	1.22 (6.198)***
5	800-1000	-0.23 (-1.137)	1.75 (8.587)***
6	1000-1200	-0.24 (-1.206)	1.98 (9.824)***
7	1200-1400	-0.27 (-1.324)	1.93 (9.484)***
8	1400-1600	-0.17 (-0.859)	1.91 (9.431)***
9	1600-1800	-0.25 (-1.241)	1.92 (9.413)***
10	1800-2000	0 (-0.014)	2 (9.879)***

Table 3.10: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*er* vs. *der*) and order of mention (1st vs. 2nd). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

3.2.3.3. Discussion

The eye-movement data showed a very strong effect of the information status of the antecedents on pronoun resolution. Shortly after pronoun onset (200 ms), participants fixated more often pictures of the focused second-mentioned than the non-focused first-mentioned entity of the antecedent sentence, as is indicated by the overall persisting main effect of mention. We also found an overall second-mention preference in the off-line task which was at ceiling, indicating that the final interpretation preferences were equally affected by the focus information. This is in line with previous research on the pronoun resolution in English where the personal pronoun preferred focused entities compared to non-focused entities (Cowles, et al., 2007). The current results differ from those of Cowles et al., however. They found that pronoun resolution was affected similarly by focus information as compared to topic information, whereas the results reported above show that the focus–non-focus distinction was a stronger cue than the topic–non-topic distinction (chapter 3.1).

In fact, the effects after marked antecedent structures emerge as early as possible (200 ms), and remain very strong throughout the whole analysis window (until 2000 ms), indicating that there is no ambiguity after marked antecedent structures. The early emergence of this effect was probably due to top-down processes. In pronoun comprehension, such processes are typically triggered by the context and the expectations it creates for the subsequent

discourse (e.g. Stevenson, 1996)[47]. In our materials, the antecedent sentence created a certain discourse expectation (discourse representation) which was to continue with the entity in focus. Bottom-up processes are initiated by the pronominal form, and may maintain or modify these top-down processes. Search processes for an adequate referent of a pronoun are therefore defined as bottom-up processes[48]. In these non-canonical materials, the bottom-up processes triggered by the personal pronoun or the d-pronoun did not modify the top-down processes.

This underlines the importance of information structural cues in pronoun resolution. The fact that the second-mentioned antecedent was focused overrode all other factors, and made only this antecedent available for subsequent resolution. The non-focused first-mentioned antecedent was thus backgrounded, and not available for resolution. This leads to the conclusion that personal and d-pronouns may not only have overlapping functions when only one potential antecedent is available, but also when more than one potential antecedent is available. That is, the discrimination of the functions of the two pronouns (as observed in chapter 3.1) may only take place when other discourse cues for disambiguation are unavailable.

This finding is particularly important with regard to previous research on the resolution of personal and d-pronouns which mainly found second-mentioned or no preferences for the two types of pronouns following marked OVS sentence structures (Bosch, et al., 2007a; Kaiser & Trueswell, 2008; Wilson, 2009). It shows that direct comparisons of the grammatical role and order of mention criteria of antecedents are not licensed across different sentence structures, as information structural cues have a strong influence on pronoun resolution. Rather we should re-evaluate these findings in the light of information structure and start to investigate its influence on pronoun resolution systematically.

In the next section, I report the resolution preferences found for Dutch personal and d-pronouns after marked antecedent structures.

---

[47] Stevenson (1996) defines top-down and bottom-up processes in pronoun comprehension with regard to mental models (Johnson-Laird, 1983). Mental models are non-linguistic representations of the discourse. Top-down processes are defined as the way that the structure of a certain mental model influences the interpretation of the subsequent discourse, whereas bottom-up processes are triggered by the semantic content of a linguistic form, such as a pronoun.

[48] Note that from this point of view, Gernsbacher's First-Mention Account is defined as a top-down process (it creates a discourse representation which is build on the basis of the first-mentioned entity), whereas the Parallel Functions Account is a bottom-up process (which is initiated upon encountering the pronoun).

3.2.4. Pronoun resolution in Dutch

This experiment investigates if and how the markedness of the antecedent structure influences the preferences in Dutch in comparison to the German preferences.

3.2.4.1. Methods

Participants

Twenty-four Dutch native speakers (20 female, 4 male) participated in the study. They were aged between 18 and 24 years (mean = 20.67; SD = 1.65). All participants were tested individually; they were paid for their participation. They spent approximately 45 min participating in the experiment. They were all students at the Radboud University Nijmegen.

Experimental Tasks

The experimental tasks were identical to those used in the previous experiment.

Materials and design

The stimuli of the above German experiment were translated into Dutch (see example (62)). As in the Dutch experiment reported in chapter 3.1, one experimental item needed to be changed, because the Dutch NP *potlood* (pencil) had neuter gender. Therefore, we changed it to felt tip, *viltstift* in Dutch which has common gender. The picture was changed accordingly. In the following, a Dutch example of an experimental item in both conditions is given.

(62) Conditions: Sample Item

a. Personal Pronoun Condition
Zwaarder dan de tafel is de kast. **Hij** is afkomstig uit een meubelwinkel in België. De sofa zal volgende week geleverd worden
Heavier than the table is the cupboard. ***It*** *[P] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

b. D-pronoun Condition
Zwaarder dan de tafel is de kast. **Die** is afkomstig uit een meubelwinkel in België. De sofa zal volgende week geleverd worden
Heavier than the table is the cupboard. ***It*** *[P] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

Procedure

The procedure was the same as for the German experiment. The Dutch items were read out loud by the same female native Dutch speaker as for the experiment reported in chapter 3.1. We used exactly the same pictures as for the German experiment except that we exchanged the picture of a *pencil* with a picture of a *felt tip*.

Participants' eye movements were recorded on an SR Research EYELINK II eye tracker. The dominant eye was recorded. A sampling rate of 500-Hz was used which monitored gaze locations every 2 ms. The calibration of the camera which links the position of the eyes with a certain location on the screen, ensured that spatial accuracy was at least 0.5°.

The accuracy of the responses to the content questions was comparably high to the German experiment with 95% correct answers (24 questions; mean correct answers = 22.71, SD = 1.10).

3.2.4.2. Results

Forced Choice Questionnaire

Since, two cells (personal pronoun – 1st, d-pronoun – 1st) had an expected count less than 5 (namely 3), Fisher's exact test was conducted on the data, and as expected showed no significant association between the type of pronoun and the type of antecedent chosen, P(Fisher, 2-sided) = .657[49]. The second-mentioned entity was chosen with an equal frequency for the two types of pronouns.

Thus, the Dutch results are almost identical to the German results; there is an overall preference for the focused second-mentioned antecedent, irrespective of pronoun type.

[49] P(Fisher, 2-sided)>.05

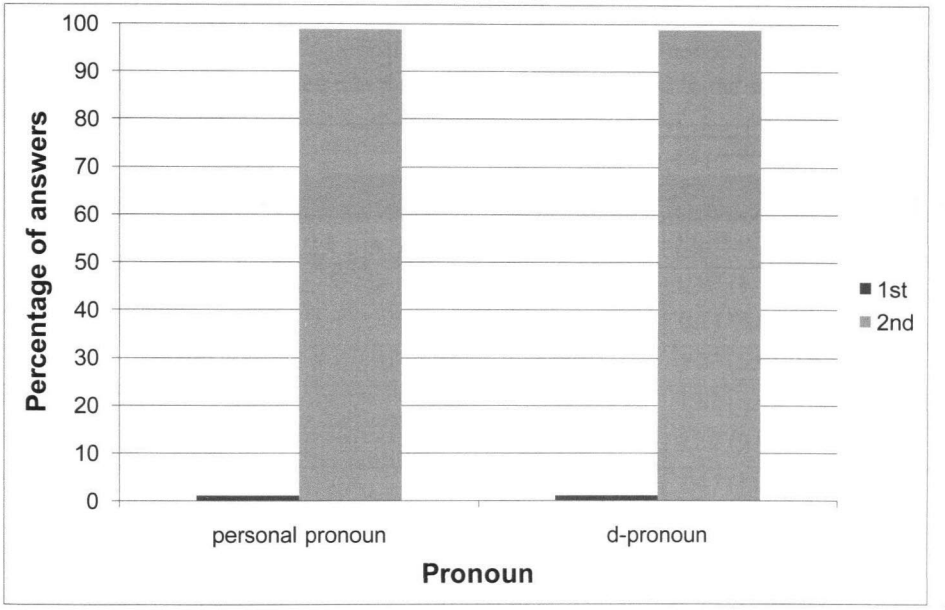

Figure 3.10: L1 Dutch Results on the Forced Choice Questionnaire; pronoun being coreferential either with the first- or the second-mentioned entity (1st, 2nd) in the two conditions (personal pronoun, d-pronoun)

	1st	2nd
personal pronoun	1.04% (3)	98.96% (285)
d-pronoun	1.04% (3)	98.96% (285)

Table 3.11: L1 Dutch Results on Forced Choice Questionnaire; pp = 24; items = 24

Visual-World Eye-Tracking

Data Analysis

The data analysis was the same as in the German experiment. 15 looks (< 1%) to either target picture that started before pronoun onset (-200 till 0 ms) were excluded from the analysis, resulting in a total of 3859 samples which entered the analysis.

Overview of the eye movement patterns

The striking similarity of the Dutch results in Figure 3.11 to the German results in Figure 3.9 already reveals that the two types of pronouns were also resolved likewise during the eye-tracking task. The target looks to the second-mentioned entity rise immediately (at 200 ms) until they reach a 40%-level, while the proportions of fixations towards the first-mentioned entity stay at a 15%-level. This difference persists until 2 seconds after the onset

of the pronoun and reflects a strong second-mention preference for the two types of pronouns.

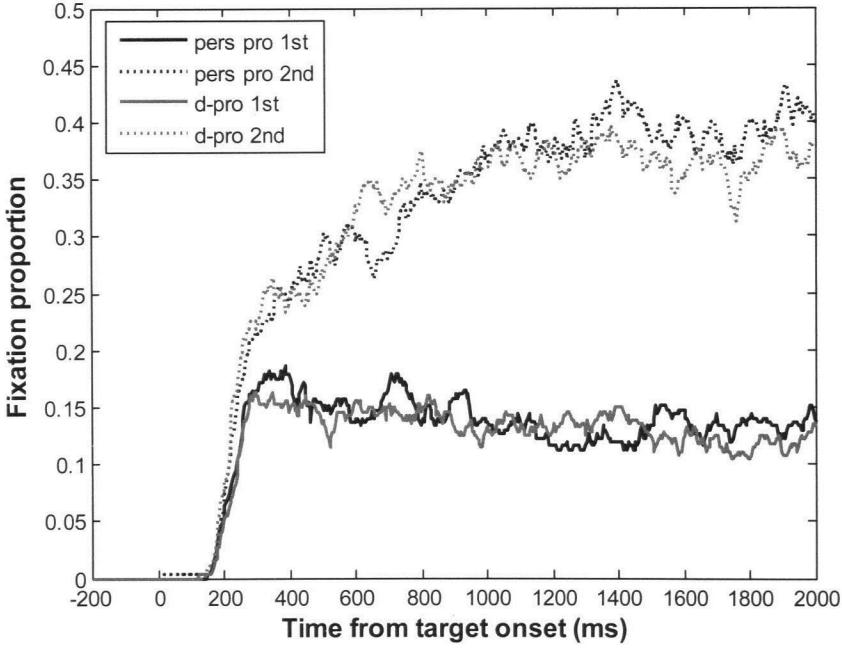

Figure 3.11: Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Statistical Analysis of the eye movement patterns

Like the German data, the Dutch data were better explained by the *simple model* than the *interaction model* in all time windows. The model predicted the outcome better when the interaction term was not included, suggesting that the resolution of the personal and the d-pronoun was not affected differently by order of mention.

The analysis of the fixed effects (see Table 3.12) showed a significant main effect of pronoun condition in the first time window (0 – 200 ms) with more looks for the d-pronoun than for the personal pronoun. Considering the low amount of overall looks in this time window and the fact that they stayed below a 5%-level, this effect was not to be attributed to pronoun resolution preferences. Rather it reflected a faster rise in looks for the d-pronoun due to its shorter duration in the speech input (mean = 229 ms; SD = 31 ms) compared to

the personal pronoun (mean = 277 ms; SD = 35 ms)[50]. Across time windows 2 to 10 (200 – 2000 ms), there was a highly significant main effect of order of mention. The positive sign of the beta coefficient indicated that there were more target looks to the second-mentioned target picture than to the first.

Time window	in ms	Fixed predictors	
		pronoun condition	order of mention
1	0-200	-0.12 (-2.151)*	0.07 (1.346)
2	200-400	-0.01 (-0.027)	0.71 (3.696)***
3	400-600	0.11 (0.502)	1.23 (5.839)***
4	600-800	-0.07 (-0.34)	1.82 (8.452)***
5	800-1000	-0.01 (-0.067)	2.12 (9.697)***
6	1000-1200	0.09 (0.394)	2.47 (11.323)***
7	1200-1400	0.04 (0.176)	2.62 (11.989)***
8	1400-1600	0.09 (0.421)	2.54 (11.661)***
9	1600-1800	0.2 (0.917)	2.54 (11.76)***
10	1800-2000	0.12 (0.532)	2.71 (12.385)***

Table 3.12: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*hij* vs. *die*) and order of mention (1st vs. 2nd). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

3.2.4.3. Discussion

Strikingly, the Dutch results replicate the German results almost one-to-one. The results of the resolution of personal and d-pronouns after marked antecedent structures in Dutch showed a strong effect of the information status of the antecedents, that is, whether in focus or not. Contrary to the resolution of the two pronouns after canonical topic-comment antecedent structures (chapter 3.1), after non-canonical topic-focus structures the two pronouns were resolved towards the same entity. The type of antecedent structure thus affected pronoun resolution. The off-line results showed the same preference in that the focused second-mentioned entity was favored almost 100% of the time across pronouns. Moreover, the timing of the on-line effect suggested that the focus–non-focus distinction was such a strong discourse cue that it created the expectation that the discourse would be continued with the focused entity before the pronoun was processed (at 200 ms). Upon encountering the pronoun, the focused entity was highly available for resolution. This is in

[50] An independent-samples *t*-test showed that the difference in duration between the two pronouns was significant ($t(46) = 5.12$, $p<.001$).

line with the view that the comprehension of the pronouns was influenced by top-down processes created by the context; the early emergence of the effect supports this idea.

The high comparability between the Dutch and German results underlines the importance of investigating the factor *focus* in pronoun resolution cross-linguistically. Since pronoun resolution after non-canonical antecedent structures has never been studied in Dutch, it was particularly useful to use inverted double nominative comparative structures which enabled us to compare the results to the German results since these structures present the antecedents in the nominative in both languages[51].

As mentioned before, the results are particularly interesting with regard to the earlier studies which after OVS structures mainly found a second-mentioned preference or no preference across pronouns (Bosch, et al., 2007a; Kaiser & Trueswell, 2008; Wilson, 2009) which should not be attributed solely to influences of grammatical role and order of mention, but rather to influences of the information structure of the antecedent sentence.

Another issue, which has been addressed in earlier studies, is the influence of animacy on pronoun resolution. Previous studies have controlled for the animacy of their antecedents, most likely to avoid a confound with potential grammatical role and order of mention effects. But all conclusions about the resolution of personal and d-pronouns are based on animate antecedents and the question arises as to whether this generalizes over resolution preferences for inanimate antecedents. Since personal pronouns and d-pronouns in German and Dutch can be used to refer to animate as well as inanimate antecedents, half of the items constructed contained two animate antecedents and the other half comprised two inanimate antecedents. Therefore, the question of whether there was an influence of the semantic factor *animacy* on the resolution of different pronominal forms was addressed, and the results are presented in the next section.

[51] Unlike German, Dutch does not mark accusative and dative case on definite lexical NPs.

3.3. The Influence of Animacy on Pronoun Resolution

3.3.1. Background

In the studies presented in chapters 3.1 and 3.2, half of the items consisted of two animate antecedents, and the other half of two inanimate antecedents. We were interested in finding out whether this antecedent property had an influence on pronoun resolution. Since in every item both antecedents had the same level of animacy, it was hypothesized that they had no disambiguating function. But the pronominal forms may not have been equally sensitive towards being used with animate or inanimate items. Therefore, the data were recoded and additional analyses were conducted to investigate whether including this factor in the statistical model improved the fit of the model.

To the best of our knowledge this factor has not been systematically investigated in on-line pronoun resolution, with most utilizing only on animate antecedents. This is partly due to the fact that some of these studies have been conducted on the Finnish pronouns *hän* and *tämä*, and only *tämä* can refer to both animate and inanimate antecedents. Earlier personal-d-pronoun-resolution studies on German controlled for the animacy of their antecedent candidates, because it was necessary to make both animate, since they presented them in SVO and OVS structures which encoded agent-patient relationships.

Animacy effects on pronoun resolution have been investigated using off-line methodology, however. For instance, Bittner (2007) investigated the influence of animacy and grammatical role on the interpretation preferences of German personal and d-pronouns. This study was part of a larger study investigating the production and comprehension of zero, personal and d-pronouns in children and adults. All antecedent sentences were SVO structures which contained either two animate (AA) or two inanimate antecedents (II), or two antecedents of mixed animacy (AI or IA), and were followed by a target sentence containing either a zero, personal or d-pronoun. Subsequently, participants were asked to *whom* the sentence referred to (e.g. *The ball is touching the bear. Ø/It [P]/It [D] was white. – Who was white?*). Bittner found that the adults preferred to interpret the personal pronoun as co-referential with the first-mentioned entity when presented after two animate, or after two inanimate antecedents. However, there was no preference for either antecedent in mixed animacy items (AI, IA). For the d-pronoun, a second-mention preference was found only when the second-mentioned antecedent was animate (AA, IA). Bittner (2007) concluded that the personal pronoun showed a preference for both animate and inanimate first-mentioned subjects, while d-pronouns only showed a preference for second-mentioned animate objects, because animate entities are more likely to become the topic of the next utterance. But for the personal pronoun it is unclear why there was no preference in the

mixed conditions, especially in the animate-subject-first condition (AI), where the first-mentioned entity is highly prominent in terms of grammatical role/topicality and animacy[52]. Unfortunately, no statistical analyses are provided; only percentages are given. This makes it difficult to draw any firm conclusions. Interestingly, however, it indicates that animacy might have an effect on the resolution of personal and d-pronouns.

In the animacy literature, animacy has been differentiated from other salience factors such as grammatical role and order of mention in that it encodes *inherent accessibility* by its intrinsic semantic properties as opposed to *derived accessibility* (such as grammatical role or order of mention) which is provided temporarily by the linguistic or non-linguistic context (Prat-Sala & Branigan, 2000). Psycholinguistic studies have mainly investigated animacy effects on word order, and have found an animate-first effect, i.e. animates are preferred to occur in agent-first utterances (and patient-first utterances such as in passivizations occur more frequently with inanimates) (Dietrich & Van Nice, in press; Igoa, 1996; Prat-Sala & Branigan, 2000; Van Nice & Dietrich, 2003). For example, comprehension studies on the resolution of local structural ambiguities such as relative clause processing (Clifton et al., 2003; Mak, Vonk, & Schriefers, 2002, 2006; Traxler, Morris, & Seely, 2002; Trueswell, Tanenhaus, & Garnsey, 1994; Weckerly & Kutas, 1999) have shown that animacy guides the parser to a certain interpretation and thereby helps to disambiguate sentences. Trueswell et al. (1994) found that sentences like *the defendant/the evidence examined by the lawyer* are preferred with inanimates, since animates prefer a main clause reading and therefore result in processing difficulties. But there is still disagreement as to how animacy interacts with other factors, such as grammatical role and discourse information in comprehension (see De Swart, Lamers, & Lestrade, 2008, for an overview). The animate-first effect might be of importance in pronoun resolution. Animates in word order studies have been shown to be preferentially put in first sentence position. It may be that personal pronouns prefer to be resolved towards an animate-first entity in comparison to an inanimate-first entity, because it is more likely to be the topic of the following utterance. The same could be derived for d-pronouns. D-pronouns prefer a second-mentioned non-topical entity more when it is animate versus inanimate, because it marks topic shift and topic shift is more likely to occur with an animate versus an inanimate entity.

Fukumura and van Gompel (accepted) studied how animacy influenced the choice of referring expressions in English when presented in fully animate, inanimate or mixed

[52] Note that Bittner (2007) presented the entities not only in the linguistic material, but parallely in the non-linguistic context to make the task more suitable to be used with children. Inanimate objects such as trains and balls were represented by moving puppets. This may have let to an animation of these inanimate objects.

conditions (AA, II, AI, IA). In a set of written sentence completion experiments, they presented an antecedent sentence (such as *The boats overtook the swimmers in the end. Clearly, ...*) in four different animacy conditions: AA (*rowers-swimmers*), AI (*rowers-yachts*), IA (*boats-swimmers*), and II (*boats-yachts*). Participants were asked to continue the sentence by referring to the syntactic subject (Experiment 1)[53]. If animacy increased salience of the antecedent candidates, it was assumed that participants would more often choose a pronoun to refer to the animate entity compared to the inanimate entity. They found an increased use of pronouns as opposed to definite NPs or names, when the antecedent was animate, but no effect of competitor animacy, supporting an absolute saliency account (AI, AA > IA, II). Overall, fewer pronouns were used in the identical animacy conditions (AA, II) than in the mixed animacy conditions (AI, IA). This suggests that pronoun use was felicitous in mixed animacy conditions, because animacy could disambiguate the relationship between the two antecedents. This finding is interesting to our study, since we exclusively presented our items in the AA or in the II condition. Thus, the factor animacy does not disambiguate between the antecedents. However, since the study found support for the absolute saliency account, this hints at a possible influence of presenting two animate versus two inanimate antecedents on pronoun resolution. That is, the personal pronoun may be preferentially resolved towards a first-mentioned animate antecedent even when the second-mentioned antecedent is also animate. But in the case of the d-pronoun, which we expect to favor second-mentioned non-topical antecedents, there are no clear predictions on how animacy affects its resolution. It could be that the d-pronoun prefers second-mentioned non-topical entities. In this case its resolution should not be affected by animacy. It could also be that the d-pronoun introduces a topic shift and therefore favors second-mentioned non-topical entities. Then, we expect an influence of animacy in that animate antecedents are favored over inanimate antecedents, because they qualify better as future topics.

This leads us to the following two research questions:

1. Does the *inherent accessibility factor* animacy influence pronoun resolution? And if yes, how does it interact with *derived factors* such as order of mention of the antecedents in referential processing?
2. Is there cross-linguistic evidence for the influence of animacy information on pronoun resolution between German and Dutch?

[53] In Experiment 2, they repeated the set-up, but asked the participants to continue with the syntactic object. The results also supported the absolute animacy account.

3.3.2. Results: Native Germans

To analyze these questions, we tested whether adding *animacy* as a fixed factor to the linear mixed model (see chapter 3.1.3) would explain the data better than without this factor. Thus, the goodness-of-fit of the following three models was calculated for each of the ten 200 ms time slices from pronoun onset till 2 seconds after the pronoun onset:

1. *simple model:* containing the terms for the main effects of order of mention (*mention*) and type of pronoun (*condition*)
2. *interaction model:* the interaction term between order of mention and type of pronoun was added
3. *animacy model:* the three-way interaction term between order of mention, type of pronoun and animacy was added (intercept = condition: *d-pronoun*, mention: *1st*, animacy: *animate*)

Log-likelihood tests were calculated to determine which model better explained the variance in the data for each time window (see Table 0.8 in the Appendix). As shown in chapter 3.1.3, the *interaction model* explained the data better than the *simple model* in the time windows between 600 ms and 2000 ms. However, the data was even better explained when the three-way interaction term *mention x condition x animacy* was included in time window 8 (1400 – 1600 ms), and the *animacy model* predicted the data marginally significantly better in time windows 4 (600 – 800 ms) and 7 (1200 – 1400 ms). This means that *animacy* had an effect on pronoun resolution.

Turning to the fixed effects (see Table 3.13), time window 4 (600 – 800 ms) showed a marginally significant interaction between order of mention and animacy. In time windows 7 and 8 (1200 - 1600 ms), the three-way interaction *condition x mention x animacy* became significant, and the two-way interactions of *mention x animacy* and *condition x animacy* were either marginally significant or significant. In addition, in window 8 there was a significant main effect of animacy. The beta coefficient suggested that there were more looks for animate than for inanimate items. In general, the significant interactions with animacy show that animacy had an influence on the resolution effects.

To further evaluate the effects found for the *animate model*, individual analyses are provided for animate and inanimate items.

Time window	in ms	Fixed predictors						
		main effects			2-way interactions			3-way interaction
		pronoun condition	order of mention	animacy	condition x mention	condition x animacy	mention x animacy	condition x mention x animacy
1	0-200	0.06 (0.901)	-0.01 (-0.167)					
2	200-400	0.2 (1.223)	-0.17 (-1.019)					
3	400-600	0.18 (0.92)	-0.23 (-1.191)					
4	600-800	0.28 (0.704)	-0.07 (-0.166)	-0.43 (-1.092)	-0.8 (-1.428)	0.45 (0.798)	0.93 (1.659)†	-0.01 (-0.017)
5	800-1000	0.53 (1.823)†	0.93 (3.168)**		-1.16 (-2.81)**			
6	1000-1200	0.91 (3.026)**	1.35 (4.527)***		-1.84 (-4.354)***			
7	1200-1400	-0.19 (-0.438)	0.52 (1.22)	-0.65 (-1.524)	-0.55 (-0.906)	1.67 (2.763)**	1.06 (1.768)†	-1.84 (-2.153)*
8	1400-1600	-0.14 (-0.323)	0.29 (0.668)	-0.93 (-2.16)*	-0.54 (-0.883)	2 (3.294)**	1.34 (2.215)*	-2.45 (-2.859)**
9	1600-1800	0.53 (1.769)†	0.7 (2.314)*		-1.3 (-3.063)**			
10	1800-2000	0.56 (1.868)†	0.55 (1.826)†		-1.29 (-3.015)**			

Table 3.13: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*er* vs. *der*), order of mention (1st vs. 2nd) and animacy (animate vs. inanimate). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

3.3.2.1. Animate Items

Figure 3.12 shows the resolution of personal and d-pronouns for the animate items (2325 looks; 51%) in the German participants. Interestingly, for animate items the resolution pattern differs from the overall personal pronoun resolution pattern reported earlier, where both item types are collapsed (see chapter 3.1.3), where there was a relatively late first-mention preference for *er* (1400 ms). When we look at animates alone, this preference seems to emerge much earlier (around 600 ms). To test for these preferences, a statistical analysis was conducted for the animate items only. This analysis tested the influence of the factor order of mention for each type of pronoun.

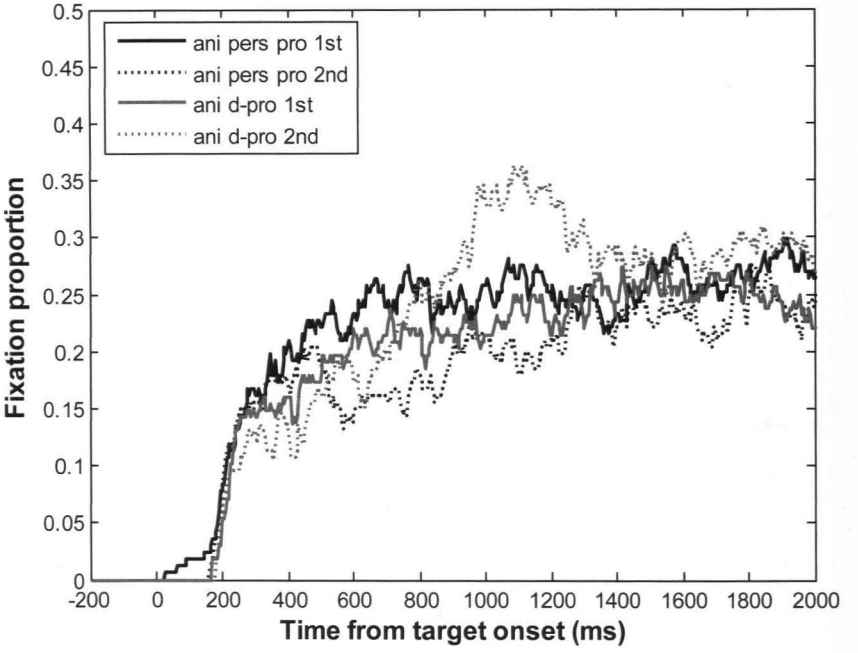

Figure 3.12: Animate items. Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

For the personal pronoun *er* (see Table 3.14), there was a significant main effect of order of mention relatively early between 600 and 800 ms. The negativity of the beta coefficient for mention reflected a first-mention preference.

For the d-pronoun *der*, there was a significant main effect of order of mention relatively late between 1000 and 1200 ms. The positivity of the beta coefficient for mention indicated

more looks to the second-mentioned target picture than to the first.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	-0.12 (-0.938)	0.04 (0.318)
2	200-400	-0.06 (-0.156)	-0.22 (-0.69)
3	400-600	-0.55 (-1.371)	-0.24 (-0.635)
4	600-800	-0.87 (-2.193)*	-0.07 (-0.165)
5	800-1000	-0.61 (-1.492)	0.65 (1.535)
6	1000-1200	-0.55 (-1.339)	1.16 (2.622)**
7	1200-1400	-0.03 (-0.063)	0.52 (1.213)
8	1400-1600	-0.25 (-0.59)	0.29 (0.647)
9	1600-1800	-0.21 (-0.489)	0.38 (0.888)
10	1800-2000	-0.3 (-0.697)	0.56 (1.294)

Table 3.14: Results of the individual time course analyses for each type of pronoun (*er* and *der*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

Taken together the individual analyses suggest that for the animate items, there was a relatively early first-mention effect for the personal pronoun, and a relatively late and short second-mention effect for the d-pronoun. The effect of the personal pronoun emerged in time window 4 (600 – 800 ms). We had seen before that when the factor animacy was included in a mixed model, the model could better explain the overall variance over all items. That means that although no significant three-way interaction was found, the marginally significant interaction between animacy and mention may be due to this early first-mention preference for the personal pronoun for animate items. The d-pronoun *der* showed a second-mention effect in time window 6 (1000 – 1200 ms) which was not affected by animacy. Therefore, this effect is not different from the overall effect. In time windows 7 and 8, we do not observe any significant effects for either type of pronoun for animate items.

3.3.2.2. Inanimate Items

The same type of analysis was conducted on the inanimate items. Figure 3.13 shows the resolution of personal and d-pronouns separately for the inanimate items (2252 looks; 49%). As can be seen in Figure 3.13, the direction of the preferences is the same as compared to the animate items. Strikingly, in contrast to the early effect for the personal pronoun for the animate items, there was a relatively late first-mention preference for

inanimate items. This means that the personal pronoun is strongly affected by the animacy of the antecedents. The second-mention preference for the d-pronoun seems to appear earlier for inanimate compared to animate items.

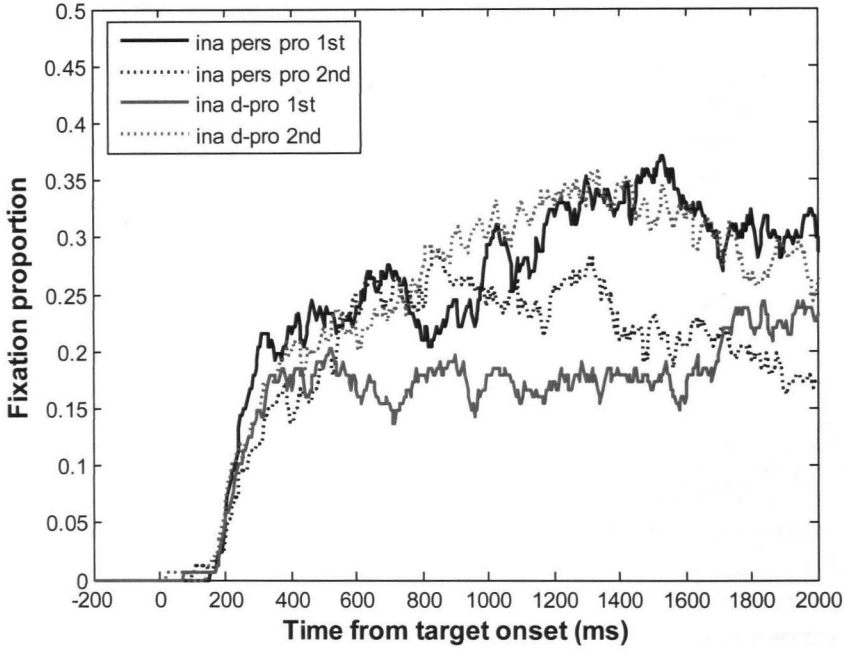

Figure 3.13: Inanimate items. Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Individual comparisons were conducted for the two types of pronouns with order of mention as a predictor (see Table 3.15). For the *er* condition the analyses revealed a marginally significant main effect of order of mention between 1200 and 1400 ms, after which the effect became significant and persisted until the last time windows (1400 – 2000 ms). The negativity of the beta coefficient for mention indicated a first-mention preference.

For the d-pronoun *der,* there was a significant main effect of order of mention between 600 and 1800 ms. The positivity of the beta coefficient for mention reflected more looks to the second-mentioned target picture than the second-mention preference for the d-pronoun *der.*

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	-0.02 (-0.131)	0.06 (0.579)
2	200-400	-0.51 (-1.497)	0.11 (0.332)
3	400-600	-0.44 (-1.103)	0.3 (0.781)
4	600-800	0.05 (0.115)	0.87 (2.288)*
5	800-1000	0.13 (0.317)	1.2 (2.929)**
6	1000-1200	-0.43 (-0.996)	1.55 (3.79)***
7	1200-1400	-0.8 (-1.803)†	1.58 (3.869)***
8	1400-1600	-1.36 (-3.156)**	1.63 (3.904)***
9	1600-1800	-1.01 (-2.35)*	1.01 (2.412)*
10	1800-2000	-1.17 (-2.819)**	0.54 (1.267)

Table 3.15: Results of the individual time course analyses for each type of pronoun (*er* and *der*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

Taken together the individual analyses suggest that for the inanimate items, there was a relatively strong and long lasting second-mention effect for *der* (600 – 1800 ms) and a quite late first-mention effect for *er* which only started to be significant at 1400 ms. Concerning time windows 4, 7, and 8, which had previously been shown to be affected by the factor animacy, the analyses for the personal pronoun revealed a significant first-mention effect during windows 7 and 8. This pattern is the opposite of what was found for animate items, where there was a significant effect in window 4 only. As for the d-pronoun, significant effects emerged in all of the three time windows for inanimate antecedents, as opposed to the animate items where no effect was found in these time windows.

3.3.3. Discussion

Animacy was found to have a strong effect on the resolution of personal and d-pronouns in German. This was shown by the fact that the mixed model explained the data better for time windows 4, 7 and 8 when animacy was included as a predictor. Animacy did not have an effect on the direction of the preferences for the pronouns, but strikingly, the personal pronoun was resolved earlier when following animate antecedents (600 ms) than compared to the overall analysis in chapter 3.1.3 (1400 ms).

	er	*der*
animate	1st at 600 ms	2nd at 1000 ms
inanimate	1st at 1400 ms	2nd at 600 ms

Table 3.16: Direction and emergence of effects of the personal and the d-pronoun for animate and inanimate antecedents

This is in line with our earlier assumption that the personal pronoun prefers animate first-mentioned topical antecedents over inanimate antecedents. In sum, the direction of the first-mention effect for the personal pronoun suggests that the order of mention/topicality information was used to resolve the pronoun. But the difference in timing of the effect between animate and inanimate items indicates that the personal pronoun was also sensitive to animacy information. Disambiguation was facilitated with animate items suggesting that the personal pronoun prefers animate over inanimate antecedents. To further investigate this issue mixed items of the type AI and IA would be needed in a future study.

The d-pronoun was comparatively unaffected by animacy. With regard to the timing, after inanimate antecedents the second-mention effect emerged in time window 4 (600 ms), whereas it only emerged at 1000 ms for the animate items. But compared to the overall results of chapter 3.1.3, the effect did not appear considerably earlier or later than the overall effect (800 ms). Thus with regard to the earlier predictions, it appears that the preference of d-pronouns to be resolved towards non-topical antecedents is robust across animate and inanimate items. As to the duration of the effects, we observe a longer lasting second-mention effect after inanimate antecedents (600 – 1800 ms) compared to the relatively short effect after animate antecedents (1000 – 1200 ms), which hints at a sensitivity of this form towards animacy information. This sensitivity cannot be interpreted here since we do not know what longer lasting effects might indicate. Taking the results together, the assumption that the d-pronoun prefers animate antecedents regardless of the animacy of the competitor antecedent, because they qualify better as subsequent topics (Bittner, 2007) can be rejected.

The animacy findings also explain the significant *condition x mention* interaction which was reported in chapter 3.1.3 for the 600-800 ms segment. There were no significant main effects which suggests a possible effect of the experimental materials. The interaction became significant during this fourth time window, because at this time the first-mention effect for the personal pronoun emerged for animate items, while for inanimate items a second-mention effect for the d-pronoun was in evidence.

In general, animacy affects German pronoun resolution, irrespective of whether the

antecedents are both animate or inanimate. The effects found in chapter 3.1.3 (*er*-1st at 1400 ms, *der*-2nd at 800 ms) for both animacy types of antecedent mask the early preference for *er* (at 600 ms) for animate items, and the late effect at 1400 ms seemed to be driven by the effect for inanimate items. For the d-pronoun, we found a second-mention effect in both cases around the same time (at 600 ms). Thus, the d-pronoun was not as strongly affected by animacy as the personal pronoun. However, the effect was longer lasting and clearer for inanimate items. This further underlines the earlier assumption that in German, d-pronouns are marked for non-topical co-reference relations and therefore show robustness across animate and inanimate antecedents, while personal pronouns are unmarked, and are therefore affected by animacy information.

The findings of these additional analyses are particularly interesting with regard to earlier visual-world studies on the resolution of personal and d-pronouns which based their findings on animate materials only. When interpreting such findings, attention should be paid to the fact that "inanimacy" might have an influence on the timing of these effects. The question of whether this sensitivity towards animacy information in pronoun resolution extends to Dutch is addressed in the next section.

3.3.4. Results: Native Dutch

As for the native German data, log-likelihood tests were calculated, to determine whether the full model containing the *condition x mention x animacy* interaction explained the data better than either the *simple model* (for time windows 1 to 3: 0 – 600 ms) or the interaction model (for time windows 4 to 10; 600 – 2000 ms) not containing animacy (see Table 0.9 in the Appendix). It was found that the *animacy model* predicted the data significantly better than the *simple model* in time window 3 (400 – 600 ms). In time window 10 (1800 – 2000 ms), the fit of the *animacy model* was marginally significantly better than the *interaction model*. This means that *animacy* had an effect on pronoun resolution.

Turning to the fixed effects (see Table 3.17), a marginally significant two-way interaction between order of mention and animacy was observed in time window 3 (400 – 600 ms) together with a significant main effect of animacy. The positivity of the beta coefficient indicated that there were overall more looks for the inanimate versus the animate items. In time window 10 (1800 – 2000 ms), the three-way interaction was highly significant. The *condition x mention* interaction was also highly significant, as well as both main effects whose beta coefficients indicated more overall looks for the personal than for the d-pronoun and at the same time more overall looks to the second-mentioned target picture than to the first. The *condition x animacy* interaction was marginally significant. The significant interactions containing animacy indicated an effect of this factor on pronoun resolution.

Time window	in ms	Fixed predictors						
		main effects			2-way interactions			3-way interaction
		pronoun condition	order of mention	animacy	condition x mention	condition x animacy	mention x animacy	condition x mention x animacy
1	0-200	-0.07 (-1.135)	-0.08 (-1.187)					
2	200-400	-0.17 (-0.997)	-0.38 (-2.252)*					
3	400-600	0.22 (0.575)	0.37 (0.965)	0.87 (2.274)*	-0.73 (-1.349)	0.01 (0.025)	-0.91 (-1.685)†	0.19 (0.251)
4	600-800	0.5 (1.716)†	-0.08 (-0.286)		-0.98 (-2.388)*			
5	800-1000	0.49 (1.672)†	0.05 (0.161)		-1.08 (-2.588)**			
6	1000-1200	0.58 (1.927)†	0.52 (1.707)†		-1.34 (-3.126)**			
7	1200-1400	0.99 (3.276)**	1.11 (3.678)***		-1.8 (-4.23)***			
8	1400-1600	0.82 (2.74)**	0.98 (3.278)**		-1.45 (-3.435)***			
9	1600-1800	0.84 (2.849)**	0.53 (1.797)†		-1.14 (-2.747)**			
10	1800-2000	1.61 (3.852)***	1.4 (3.365)***	0.06 (0.128)	-2.88 (-4.882)***	-1.03 (-1.753)†	-0.79 (-1.332)	2.13 (2.553)**

Table 3.17: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*hij* vs. *die*), order of mention (1st vs. 2nd) and animacy (animate vs. inanimate). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

To further evaluate the effects found for the *animate model*, individual analyses are provided for animate and inanimate items.

3.3.4.1. Animate Items

Figure 3.14 shows the resolution of personal and d-pronouns separately for the animate items (2354 looks; 51%). Overall the figure reveals similar preferences for the two pronouns as observed in chapter 3.1.4.

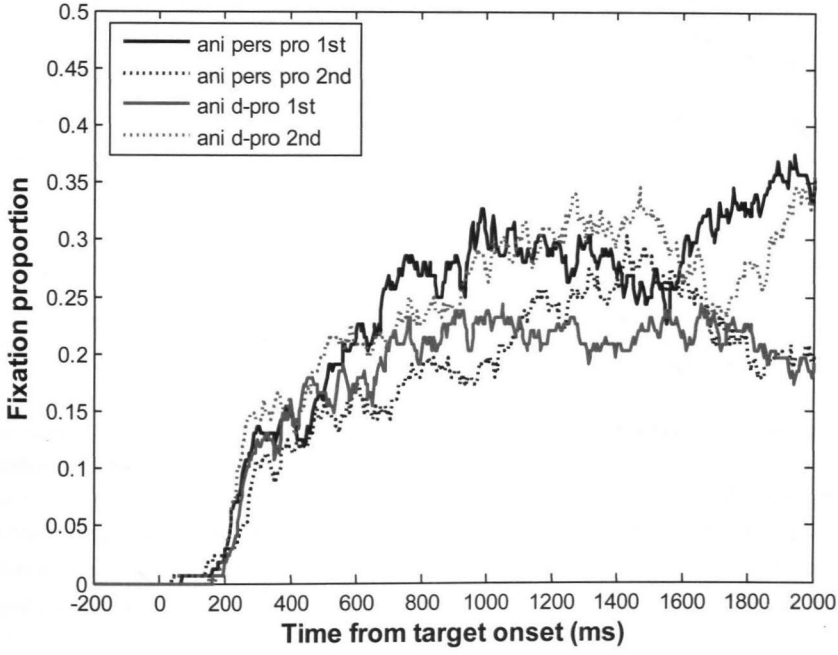

Figure 3.14: Animate Items. Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Separate comparisons were conducted for the two types of pronouns with order of mention as a predictor. For the *hij* condition (see Table 3.18), the analyses revealed a significant main effect of order of mention between 600 and 1000 ms, and during the last two time windows (1600 – 2000 ms). The negativity of the beta coefficients for mention indicated a first-mention preference.

For the d-pronoun, a significant main effect of order of mention was observed between 1000 and 1400 ms, and which was still marginally significant between 1400 and 1600 ms

and reappeared to be significant in the last time window (1800 – 2000 ms). This effect was due to a second-mention preference.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	-0.14 (-1.071)	0.14 (1.048)
2	200-400	-0.26 (-0.847)	0.18 (0.538)
3	400-600	-0.36 (-0.973)	0.37 (0.965)
4	600-800	-0.97 (-2.369)*	0.18 (0.437)
5	800-1000	-1.13 (-2.73)**	0.54 (1.286)
6	1000-1200	-0.54 (-1.235)	0.95 (2.233)*
7	1200-1400	-0.16 (-0.378)	1.13 (2.623)**
8	1400-1600	-0.07 (-0.163)	0.82 (1.907)†
9	1600-1800	-1.1 (-2.512)*	0.32 (0.763)
10	1800-2000	-1.48 (-3.435)***	1.4 (3.26)**

Table 3.18: Results of the individual time course analyses for each type of pronoun (*hij* vs. *die*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

Taken together, the individual analyses suggest that for the animate items, there was a first-mention effect for the personal pronoun and a second-mention effect for the d-pronoun. There was no effect during the time window 3 (400 – 600 ms). But in chapter 3.1 we found a significant first-mention effect for the personal pronoun at this early time window, suggesting that it might be driven by the inanimate items. During the last time window (1800 – 2000 ms) effects for both pronouns were found (which were also found in chapter 3.1).

3.3.4.2. Inanimate Items

Figure 3.15 shows the resolution of personal and d-pronouns for the inanimate items (2251 looks; 49%) in the Dutch native speakers. The direction of the preferences is the same as compared to that of the animate items. However, the first-mention effect for the personal pronoun was longer-lasting for the inanimate items, and indeed appeared earlier. The figure also indicates that the second-mention preference for the d-pronoun is slightly later for inanimate items than for animate items.

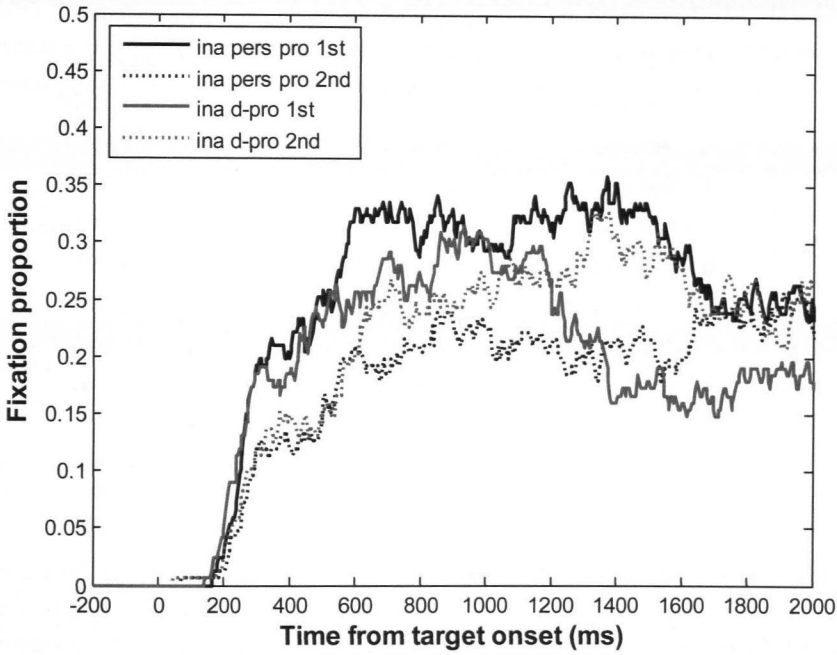

Figure 3.15: Inanimate items. Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

As before, individual analyses were conducted for the two types of pronouns with order of mention as a predictor (see Table 3.19). For the personal pronoun the analyses revealed a significant main effect of order of mention between 200 and 1600 ms. The beta coefficient for mention was negative indicating a first-mention preference.

For the d-pronoun the analyses revealed a marginally significant main effect of order of mention between 200 and 400 ms. Interestingly the beta coefficient for mention was negative which reflected a first-mention preference. Considering the direction of this effect as well as its early emergence, it was probably not due to pronoun resolution preferences, but rather to general discourse effects. There was a significant second-mention effect in time windows 7 and 8 (1200 – 1600 ms), which was marginal in window 9 (1600 – 1800 ms).

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	-0.13 (-1.069)	-0.18 (-1.35)
2	200-400	-0.83 (-2.286)*	-0.6 (-1.752)†
3	400-600	-1.08 (-2.625)**	-0.54 (-1.411)
4	600-800	-1.14 (-2.738)**	-0.34 (-0.795)
5	800-1000	-0.94 (-2.198)*	-0.45 (-1.033)
6	1000-1200	-1.11 (-2.592)*	0.08 (0.189)
7	1200-1400	-1.23 (-2.832)**	1.08 (2.58)**
8	1400-1600	-0.87 (-2.078)*	1.13 (2.745)**
9	1600-1800	-0.13 (-0.305)	0.73 (1.83)†
10	1800-2000	-0.13 (-0.31)	0.62 (1.521)

Table 3.19: Results of the individual time course analyses for each type of pronoun (*hij* vs. *die*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

Taken together, an early first-mention preference (200 – 400 ms) was found for the two types of pronouns. This preference disappears for the d-pronoun at 400 ms and becomes stronger for the personal pronoun. For the d-pronouns, a relatively late (1200 ms) second-mention effect is observed. Time windows 3 and 10 had been shown to better predict the data when animacy was included in the mixed model. The only significant effect for inanimate items here was a first-mention effect for the personal pronoun in time window 3. This contrasts with the observation for animate items where significant effects for the personal and the d-pronoun were found in the last time window, and no effects were observed in time window 3.

3.3.5. Discussion

The analysis of the factor animacy revealed a clear influence on the resolution of personal and d-pronouns in Dutch. But there were differences observed compared to the German results. In Dutch, inanimacy triggered a first-mention preference for the two types of pronouns between 200 and 400 ms. This effect emerged so early that it is unlikely that it reflected the resolution of the pronoun. It was perhaps driven by a discourse expectation which triggered top-down processes, namely a general preference to continue the discourse with the first-mentioned topical entity. In the case of the personal pronoun, this discourse expectation was met, and thus facilitated the resolution of the personal pronoun towards this first-mentioned entity; the personal pronoun thus triggered bottom-up processes which maintained the top-down interpretation. However, upon encountering a d-pronoun in this

context, the discourse expectation was not met and the top-down interpretation needed to be modified. The d-pronoun was not resolved until 1200 ms. In other words, for the inanimate-d-pronoun condition we observed a switch in preferences from the first-mentioned topical antecedent to the second-mentioned non-topical antecedent.

	hij	*die*
animate	1st at 600 ms	2nd at 1000 ms
inanimate	1st at 200 ms 1st at 400 ms	1st at 200 ms 2nd at 1200 ms

	er	*der*
animate	1st at 600 ms	2nd at 1000 ms
inanimate	1st at 1400 ms	2nd at 600 ms

Table 3.20: Direction and emergence of effects of the personal and the d-pronoun for animate and inanimate antecedents in Dutch (above) and German (below)

Animacy therefore has an effect on the resolution of pronouns in both German and Dutch, but the type of influence is different. In German, the d-pronoun robustly resolves towards the non-topical second-mentioned entity (and is marked for it), but the personal pronoun, being guided by semantic information, is disambiguated earlier when it refers to animate antecedents. To investigate whether such semantic information could override the first-mention-topical effect, future studies could test the influence using mixed AI and IA items. Although Dutch pronouns do not have a preference for animate or inanimate antecedents, animacy information seems to have an effect on comprehenders' discourse expectations; topics (or first-mentioned entities) are generally favored when two inanimate entities are presented. Thus, the presentation of two inanimate entities in a sentence initiates the top-down process to continue the discourse with the topical entity. This leads to a facilitatory effect on the resolution of the personal pronoun, because the update of its bottom-up process maintains this top-down interpretation. In contrast, the d-pronoun initiates a search for the non-topical antecedent; thus, the top-down interpretation needs to be modified.

This cross-linguistic difference in sensitivity towards animacy information between German and Dutch may be explained in terms of the differences in case marking between the two languages. As Lamers and de Hoop (2005) have pointed out in their *Incremental Optimization of Interpretation Approach*, upon encountering a first NP which is marked as accusative, German native speakers identify it as the object of the sentence and an expectation about the sentence continuation is triggered. Thus, case marking is a highly reliable cue to sentence processing in German.

(63)	Den Zaun	hat	der Junge	zerbrochen.
	The fence_{ACC}	has	the boy_{NOM}	broken.
	"The fence, the boy broke."			

In contrast, Dutch marks case information only on pronouns. Thus, in the example (see (64), translated into Dutch), no case information is available on the full NP. In that sense, inanimacy-animacy information may well be a more important cue for disambiguation in Dutch, as it makes the first-mentioned entity more likely to be an object antecedent.

(64)	Het hek	heeft	de jongen	gebroken.
	The $\text{fence}_{SUB/OBJ}$	has	the $\text{boy}_{SUB/OBJ}$	broken.
	"The fence, the boy broke./The fence broke the boy."			

In our materials (*The doctor is friendlier than the cook./The cupboard is heavier than the table.*) both entities were either animate or inanimate; thus animacy information could not disambiguate the relationship between the two NPs. When two relatively "non-salient" inanimate antecedents were presented, it appeared to create the expectation that the discourse would continue with the topical entity, and therefore it triggered the early first-mentioned target looks. When the discourse continued with a personal pronoun, the discourse expectation was met and its integration was facilitated, whereas in the case of the d-pronoun it was not met resulting in a relatively late resolution (1200 ms) towards the second-mentioned entity.

Summing up, the reanalysis of the experiments of chapter 3.1 revealed that when personal and d-pronouns appear after canonical antecedent structures topicality information is used to disambiguate the relationship. Although in the initial set of experiments, it was not possible to disentangle order of mention effects from topicality effects, it is likely that the order of mention effects arose due to the pragmatic function of the antecedents. This is because in the experiments of chapter 3.2, the pragmatic distinction between the antecedents (focus vs. non-focus) was shown to affect the resolution of personal and d-pronouns, rather than the order of mention information. This preference was so strong that when the focus information highlighted one of the potential antecedents, the two pronominal forms showed an overlapping co-reference function: they were both resolved towards this focused antecedent. But when the two antecedents were presented after topic-comment antecedent structures, the pronouns had different functions. Moreover, particularly in German, more ambiguity was observed for the personal pronoun in contrast to the d-pronoun which was attributed to the unmarkedness-markedness distinction between the two forms. Specifically, while d-pronouns are marked for non-topical co-reference, the

personal pronoun is unmarked and thus more ambiguous. The findings of the animacy analysis support this assumption, since personal pronouns were shown to be affected by the animacy information of the antecedents while d-pronouns did not show such sensitivity. The Dutch pronouns were relatively unaffected by the animacy information of the antecedent candidates; but an effect was found on their discourse expectation. Interestingly, although for German the unmarkedness-markedness distinction for personal and d-pronouns accounted for the eye-tracking results found in chapter 3.1 and 3.3, the off-line results showed that despite its unmarkedness, the personal pronoun elicited a strong first-mentioned final interpretation as shown by the ceiling effect. This was probably due to the nature of the off-line task in which the participants were forced to take a decision; even in cases of uncertainty a decision on the direction of the preferences needed to be made. However, this does not explain why the marginally significant second-mention preference for the marked d-pronoun was not as strong in the off-line interpretations as it had been on-line. One possibility is that the material following the window of analysis (the two seconds after the onset of the pronoun) may have been semantically biased towards the more plausible first-mentioned antecedents for some of the items, weakening the second-mention preference for the d-pronoun, and strengthening the first-mention preference for the personal pronoun. On the other hand, if the German d-pronoun is so strongly marked for non-topicality, it is still puzzling that the semantic context had such a strong effect on the off-line results, given that German d-pronouns were shown to be unaffected by the semantic animacy information. It was also surprising that the off-line results between German and Dutch were so similar. To further investigate these issues, we conducted a set of off-line control experiments with German and Dutch native speakers. These experiments are reported in the next section.

3.4. Individual Differences in Pronoun Resolution – Comparing on- and off-line results

In chapter 3.1, some differences between the off-line results and the on-line results for both languages were discussed. Although the final interpretation preferences pointed in the same directions, i.e. the personal pronouns were resolved towards the topical entity, while the d-pronouns were resolved towards the non-topical entity, in the German as well as in the Dutch off-line questionnaire this preference (G: 56%; D: 57%) was only marginally different from chance (G: χ^2 (1) = 2.64, p=.061; D: χ^2 (1) = 2,893, p=.052). Was this discrepancy between on- and off-line task due to the fact that in general d-pronouns are resolved differently in real-time resolution than in off-line resolution, in that final interpretation preferences diverge from on-line preferences? To further investigate this issue, we presented only the off-line task to a German and a Dutch control group.

3.4.1. Off-line Control Experiments

3.4.1.1. German control group

The participants of the German native speakers control group were twenty teachers at the Goethe Gymnasium Karlsruhe. They were given the exact same materials as the German experimental group in chapter 3.1[54].

Figure 3.16 and Table 3.21 show the results for the control group. Strikingly, they revealed a strong second-mention preference for *der* (91%) as well as a first-mention preference for *er* (93%). The association between the type of pronoun and the choice of an antecedent was significant, χ^2 (1) = 333.91, p<.001.

Separate Chi Square tests on each pronoun investigated whether the pattern of results were different from chance. For the personal pronoun, the pattern of results obtained was significantly different from chance, χ^2 (1) = 107.10, p<.001. The first-mentioned entity was 13 times more likely to be chosen for *er* than the chance level would predict. For *der* the distribution of responses was also significantly different from chance, χ^2 (1) = 95.55, p<.001. The second-mentioned entity was 9.86 times more likely to be chosen for *der* than the chance level would predict

[54] Due to a mistake in the distribution of the questionnaires, eleven participants received version 1 and nine participants received version 2 of the questionnaire. But each questionnaire contained 12 items in the personal pronoun condition and 12 items in the d-pronoun condition, this was not considered to be a problem.

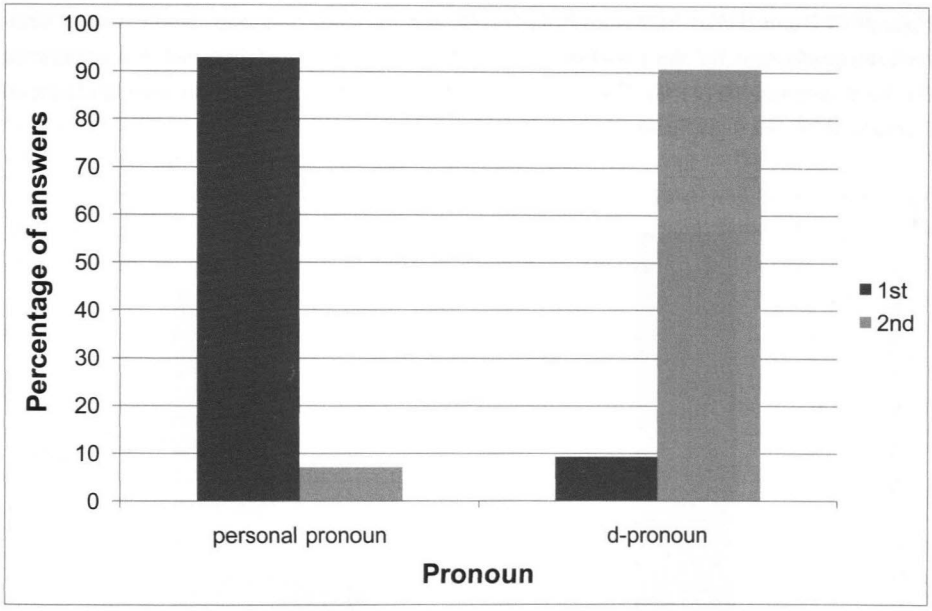

Figure 3.16: L1 German Control Results on the Forced Choice Questionnaire; pronoun being coreferential either with the first- or the second-mentioned entity (1st, 2nd) in the two conditions (personal pronoun, d-pronoun)

	1st	2nd
personal pronoun	92.86% (221)	7.14% (17)
d-pronoun	9.21% (22)	90.79% (217)

Table 3.21: L1 German Control Results on Forced Choice Questionnaire; pp = 20; items = 24; 2 *er*-answers eliminated because of choosing both antecedents, 1 *der*-answer missing

In summary, the German control experiment revealed both a highly significant first-mention preference for the personal pronoun *er* and a comparably high second-mention preference for the d-pronoun *der*.

The same off-line experiment was conducted with a Dutch control group.

3.4.1.2. Dutch control group

The participants of the Dutch control group were thirty-eight students at the Vrije Universiteit Amsterdam. As the German control group, the Dutch control group was given the same off-line questionnaire as the participants of the Dutch experiment in chapter 3.1.4. Half of the participants received version 1 and the other half received version 2 of the questionnaire.

Figure 3.17 and Table 3.22 show the results for the control group. There was a first-mention preference for the personal pronoun *hij* (93%), and a second-mention preference for the d-pronoun *die* (81%). The association between the type of pronoun and the choice of an antecedent was significant, χ^2 (1) = 496.13, p<.001.

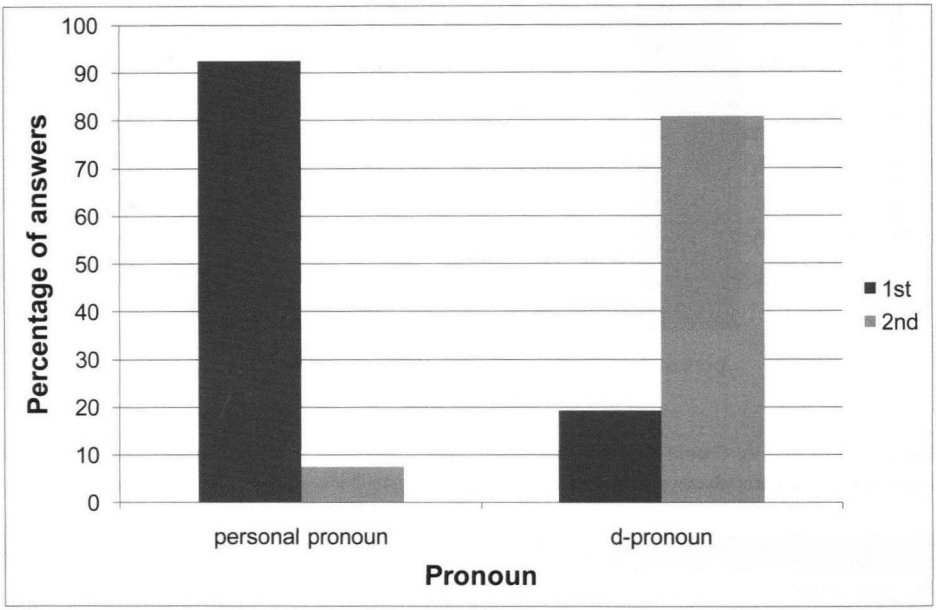

Figure 3.17: L1 Dutch Control Results on the Forced Choice Questionnaire; pronoun being coreferential either with the first- or the second-mentioned entity (1st, 2nd) in the two conditions (personal pronoun, d-pronoun)

	1st	2nd
personal pronoun	92.53% (421)	7.47% (34)
d-pronoun	19.64% (87)	80.84% (367)

Table 3.22: L1 Dutch Control Results on Forced Choice Questionnaire; pp = 38; items = 24; 1 *hij*-answer missing, 2 *die*-answers missing

Separate Chi Square tests on each pronoun investigated whether the pattern of results were different from a 50%-chance-level. For the personal pronoun, the pattern of results obtained was significantly different from chance, χ^2 (1) = 201.04, p<.001. The first-mentioned entity was 12.38 times more likely to be chosen for *hij* than the chance level would predict. For the d-pronoun *die* the distribution of responses was also significantly different from chance, χ^2 (1) = 95.42, p<.001. The second-mentioned entity was 4.22 times more likely to be chosen for *die* than the chance level would predict.

In summary, the Dutch control experiment revealed both a highly significant first-mention preference for the personal pronoun *hij* and highly significant second-mention preference for the d-pronoun *die*. The direction of the interpretation preferences were thus the same as for the German control group. However, the second-mention preference for the d-pronoun was not as strong which may be due to the German d-pronoun being marked for this type of co-reference relation. Interestingly, both the findings for the German and the Dutch control group are different to what was found for the experimental groups in chapter 3.1. While the experimental groups showed a marginally significant second-mention effect for the d-pronouns during the off-line questionnaire, the control groups showed a highly significant second-mention preference in both languages.

To better understand this difference in interpretation patterns for the d-pronouns, the German and Dutch data from chapter 3.1 were reanalyzed.

3.4.2. Reanalyzing the results of the experimental groups

The off-line results from chapter 3.1 were looked at in more detail for each participant (see Table 3.23). It was discovered that while 27 of the 28 German native participants chose the first-mentioned antecedent for at least 10 of the 12 items in the personal pronoun condition, only half of the participants chose the second-mentioned antecedent for the d-pronoun equally often. In other words, there were individual differences in final interpretation strategies for the d-pronoun, and 24 out of 28 (86%) participants were categorized as either using an *off-line-1st* or an *off-line-2nd* strategy for the d-pronoun. For the Dutch speakers, the distribution was weaker, yet still 18 of the 28 (64%) applied such a strategy.

Experimental group	Pronoun	Off-line-strategy	
		first-mentioned	second-mentioned
German natives (28 participants)	personal pronoun	27 (96%)	0
	d-pronoun	10 (36%)	14 (50%)
Dutch natives (28 participants)	personal pronoun	26 (93%)	0
	d-pronoun	8 (29%)	10 (36%)

Table 3.23: Overview of the amount of participants who consistently chose either the first- or the second-mentioned antecedent for a specific pronoun at least for 10 out of 12 items (83%) in the off-line task in chapter 3.1. Percentages from the total amount of participants in parentheses.

In order to find out whether these inter-individual differences had also affected the eye-tracking results, those of chapter 3.1 were reanalyzed, according to the off-line interpretation choices of the participants. The variable *off-line* (first- or second-mentioned)

was thus entered in the eye-tracking data.

3.4.2.1. German experimental group

As stated above, each on-line fixation data point was coded with respect to the participant's antecedent choice in the off-line task (either 1 or 2). Figure 3.18 shows the plot of the results according to the off-line antecedent choices for the German d-pronoun (looks were at ceiling for the personal pronoun, so this condition was not examined). In the plot in Figure 3.18, the black lines represent the looks of those who had chosen the second-mentioned referent in the off-line task, and the grey lines represent those who had chosen the first-mentioned antecedent.

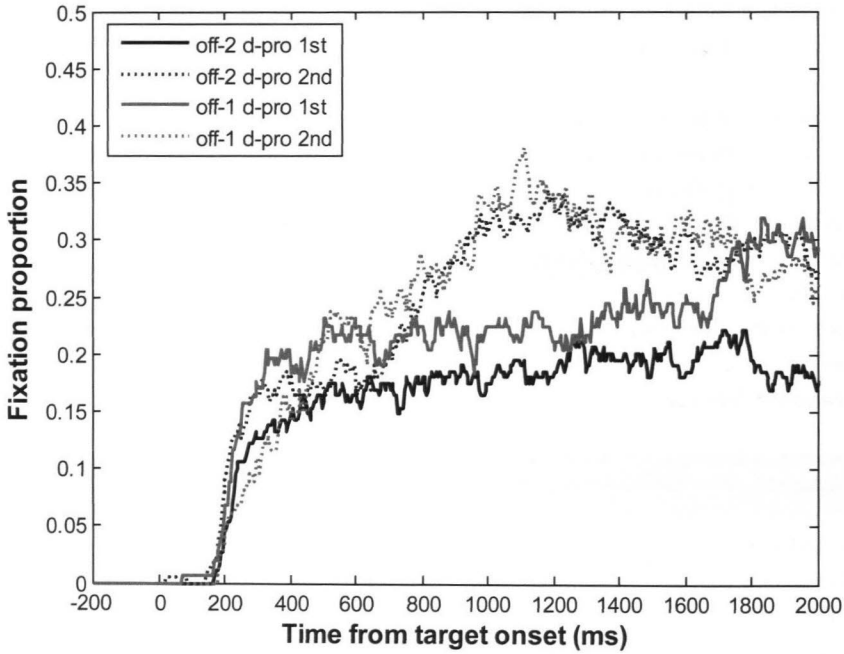

Figure 3.18: Probability of fixating the first-and second-mentioned referent in the d-pronoun condition; plotted separately according to the offline interpretation (gray = offline-*1st*; black = offline-*2nd*)

The similarity of the plots is striking. They both clearly show a second-mention preference for the d-pronoun which emerged around 800 ms after pronoun onset and reached a 35%-level at around 1200 ms.

A loglikelihood test was calculated to test whether including the off-line choice as a

predictor variable in the mixed model would improve the goodness-of-fit of the model to the data. Thus, the full model contained a three-way interaction term, namely *condition x mention x off-line* (intercept = condition: *d-pronoun,* mention: *1st*, off-line: *1st*). We examined whether the model explained the variance in the data better than the *simple model* or the *interaction model*. It was predicted that the off-line preferences should be visible during on-line resolution, in that participants who chose the first-mentioned entity for the d-pronoun were expected to also trigger more first-mentioned target looks during the eye-tracking task. This could mainly lead to two possible on-line resolution patterns:

- either a first-mention preference
- or no preference (which then turned into a first-mention preference during the off-line task, as the questionnaire forced participants to make a choice to either first-mentioned or second-mentioned NP)

However, the *off-line model* did not fit the data better than the *simple* or the *interaction model* for any of the time windows. This means that although the final interpretation preferences differed, the on-line resolution of the German d-pronoun towards the second-mentioned non-topical antecedent was always present, even for the items for which a first-mentioned antecedent was chosen during the off-line task. This underlines the observation that the d-pronoun is preferentially resolved towards the non-topical entity, even when other cues push the ultimate resolution towards another entity. Moreover, this observation highlights the importance of using on-line measures in the study of pronoun resolution as it would not have been possible to see this strong resolution preference for the d-pronoun, had we only looked at the final interpretations.

3.4.2.2. Dutch experimental group

To gain more insight into this on-line/off-line resolution behavior, we analogously reanalyzed the Dutch data from chapter 3.1.4. As can be seen in Figure 3.19, the second-mention effect for the d-pronoun was stronger for the cases where this entity was also finally chosen as being co-referential. However, in the cases where the first-mentioned entity was chosen off-line for the d-pronoun, there was no clear effect.

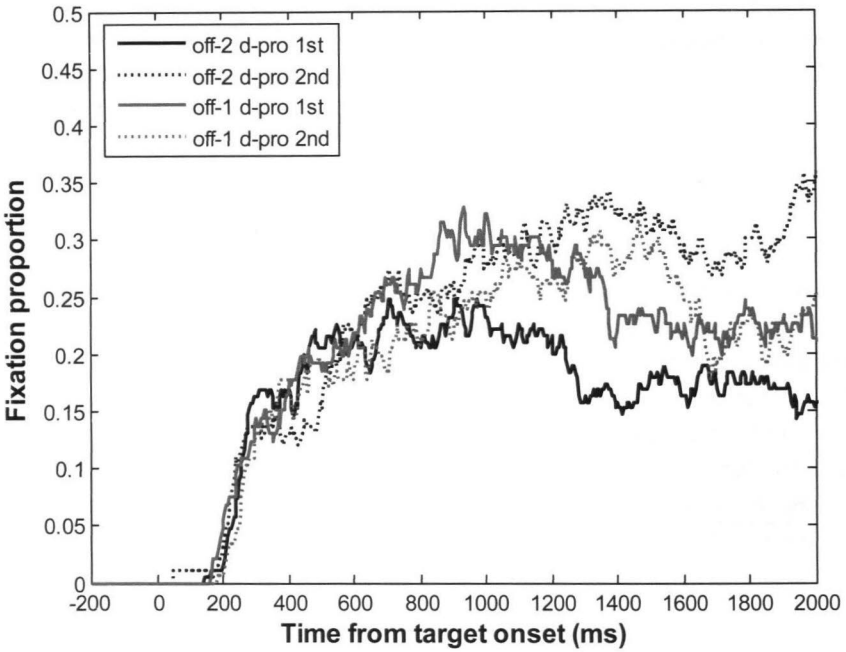

Figure 3.19: Probability of fixating the first-and second-mentioned referent in the d-pronoun condition; plotted separately according to the offline interpretation (gray = offline-*1st*; black = offline-*2nd*)

The loglikelihood test (see Table 0.11 in the Appendix) which evaluated the goodness-of-fit of the full model with the three-way interaction term of *condition x mention x off-line* compared to the *interaction model* and *simple model*, revealed that the *off-line model* explained the data marginally significantly better than the interaction model during time window 7 (1200 – 1400 ms). Therefore, the fixed effects for the *off-line model* were analyzed and revealed a significant three-way interaction during time window 7 (see Table 3.24). The two-way interactions between *condition x off-line* and *mention x off-line* were also significant; however, the *condition x mention* interaction was not. The main effects of pronoun type and off-line choice were also significant. This suggests that the final

Time window	in ms	Fixed predictors						
		main effects			2-way interactions			3-way interaction
		pronoun condition	order of mention	off-line	condition x mention	condition x off-line	mention x off-line	condition x mention x off-line
1	0-200	-0.07 (-1.135)	-0.08 (-1.187)					
2	200-400	-0.17 (-0.997)	-0.38 (-2.252)*					
3	400-600	0.23 (0.832)	-0.09 (-0.32)		-0.63 (-1.651)†			
4	600-800	0.5 (1.716)†	-0.08 (-0.286)		-0.98 (-2.388)*			
5	800-1000	0.49 (1.672)†	0.05 (0.161)		-1.08 (-2.588)**			
6	1000-1200	0.58 (1.927)†	0.52 (1.707)†		-1.34 (-3.126)**			
7	1200-1400	-2.36 (-1.844)†	-0.86 (-0.865)	-0.76 (-1.705)†	2.71 (1.529)	2.81 (2.61)*	1.26 (2.074)*	-3.68 (-2.458)*
8	1400-1600	0.82 (2.74)**	0.98 (3.278)**		-1.45 (-3.435)***			
9	1600-1800	0.84 (2.849)**	0.53 (1.797)†		-1.14 (-2.747)**			
10	1800-2000	1.09 (3.689)***	1.01 (3.422)***		-1.82 (-4.345)***			

Table 3.24: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*hij* vs. *die*), order of mention (1st vs. 2nd) and offline (off-1 vs. off-2). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

interpretation preference was reflected in the eye movements. However, we need to be cautious when interpreting the effects, because the analysis also contained the looks for the personal pronoun. To understand how the factor *off-line choice* affected the resolution of the d-pronoun *on-line* individual analyses were conducted for the resolution of the d-pronouns depending on the off-line choice (see Table 3.25).

These analyses revealed that there was no on-line effect for the d-pronoun for which the first-mentioned entity was chosen off-line (*off-line-1st*), whereas for the d-pronouns for which the second-mentioned entity was chosen off-line (*off-line-2nd*) there was an on-line second-mention effect starting at 1000 ms and lasting till the end of the analysis period (2000 ms).

Time window	**in ms**	**Fixed predictor: order of mention**	
		off-line-1st	off-line-2nd
1	0-200	-0.26 (-1.651)	0.16 (1.308)
2	200-400	-0.02 (-0.054)	-0.36 (-1.123)
3	400-600	-0.17 (-0.414)	-0.03 (-0.071)
4	600-800	-0.49 (-1.097)	0.23 (0.583)
5	800-1000	-0.61 (-1.334)	0.55 (1.393)
6	1000-1200	-0.26 (-0.55)	1.11 (2.766)**
7	1200-1400	0.4 (0.843)	1.65 (4.29)***
8	1400-1600	0.43 (0.946)	1.39 (3.589)***
9	1600-1800	-0.07 (-0.169)	0.99 (2.552)*
10	1800-2000	0.13 (0.3)	1.68 (4.345)***

Table 3.25: Results of the separate time course analyses for the d-pronouns which subsequently resulted in a different offline interpretation (*offline-1st* and *offline-2nd*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

This pattern of results is highly interesting as it suggests that when there is no on-line preference, thus when the ambiguity cannot be resolved on-line (at least during the first 2 s after pronoun onset), the topical antecedent is chosen off-line. Furthermore, this differs from the German pattern reported above: the Dutch d-pronoun seems not to be as marked for non-topicality as is the German d-pronoun. This is consistent with the idea discussed earlier that the two languages may differ with regard to the markedness feature of the d-pronoun (chapter 3.1), even though both languages have personal and d-pronouns with similar functions, they do not behave in an identical manner. This observation is not

captured by theories of reference and it underlines the importance of the Form-Specific Multiple-Constraints Approach, extending it to cross-linguistic resolution differences between forms and the degree to which they are sensitive towards different resolution constraints.

3.4.3. Discussion

Based on the reanalyses of the German and the Dutch native on-line data in relation to the participants' final interpretation of the d-pronouns (first- or second-mentioned referent), it is assumed that although the overall preference to resolve the d-pronoun towards the non-topic is a cross-linguistic preference, the two forms are differently marked for this preference. The German d-pronoun is strongly marked to refer to the second-mentioned entity because this resolution preference even emerges when it is the first-mentioned antecedent that is ultimately chosen. In contrast, the Dutch d-pronoun is less marked: it can either be clearly resolved towards the second-mentioned entity, in which case the ultimate preference is the same, or when ambiguity cannot be resolved on-line but an off-line choice is required, it is resolved towards the first-mentioned entity. The Dutch results show that the stronger the on-line second-mention effect, the higher the chance that this entity will be chosen ultimately, while this is not the case for the German.

In future research, it would be particularly interesting to investigate systematically the relationship between on- and off-line results. It might be the case that on-and off-line tasks are not only two different methodologies which measure the same phenomenon, but that the combined analysis of their results allow us an even deeper understanding of the involved processes. This may be so especially in the case of ambiguity. Here, it may be useful to look at longer time windows than 2000 ms, to better capture when final interpretations are made. For example, one could semantically bias the final resolution of the d-pronoun towards the first-mentioned entity by the post-pronominal discourse context. If there exists an initial preference to resolve the d-pronoun towards the second-mentioned entity, it should ultimately switch to a first-mention preference, observable in the on-line eye-tracking data. Finally, off-line results for d-pronoun interpretation should be obtained testing spoken language comprehension for two reasons. First of all, the modality to the on-line eye-tracking task should be the same across tasks and thus account for better comparability, and secondly, since the d-pronoun mainly appears in spoken language, a spoken language off-line comprehension task would be more valid.

3.5. General Discussion

This chapter has investigated the resolution preferences for personal and d-pronouns in ambiguous discourses in German and Dutch. In the first section, the question addressed was whether the order of mention of the antecedent candidates would have an influence on the resolution of the two pronominal forms following canonical comparative antecedent structures (e.g. *The cupboard is heavier than the table*). It was found that the personal pronouns were preferentially resolved towards first-mentioned topical antecedents and the d-pronouns towards second-mentioned non-topical antecedents in both languages. This is in line with the assumption that the two pronouns have different co-reference preferences (Ariel, 1990, 1994, 2001; Bosch, et al., 2003; Comrie, 1994; Diessel, 1999; Gundel, 2003; Gundel, et al., 1993; Lambrecht, 1994; Levinson, 1987, 1991). However, the visual-world eye-tracking task also revealed a cross-linguistic difference. In German, the personal pronoun was resolved relatively late compared to the d-pronoun which indicated a higher degree of ambiguity. This finding is in line with the assumption that the d-pronoun has a disambiguation function, thus is marked for non-topical co-reference relations, while the personal pronoun is unmarked (Ahrenholz, 2007; Bosch & Umbach, 2007b; Zifonun, et al., 1997). Such an effect was not found for the Dutch pronouns. Thus, the direction of the preferences for personal and d-pronouns was generalizable across languages, but the languages differed with regard to the markedness-unmarkedness feature of the pronouns.

Since the first-mentioned antecedents in the materials were always topical and the second-mentioned antecedents non-topical, the two factors order of mention and topicality could not be disentangled. In the second part of this chapter, the question of whether the same resolution preferences for the two pronominal forms would be found after non-canonical topic-focus antecedent structures (e.g. *Heavier than the table is the cupboard*) was investigated. It was argued that if the order of mention of the antecedent candidates guided the pronoun resolution, then the personal pronoun should still prefer first-mentioned antecedents, while the d-pronoun should prefer second-mentioned antecedent. On the other hand, with regard to the results of previous studies on pronoun resolution after non-canonical OVS structures, it was assumed that the focus information would make the second-mentioned entity particularly prominent for future reference, and therefore it was predicted that if the pragmatic focus–non-focus distinction had an effect, then the two pronominal forms should prefer the second-mentioned focused entity. Alternatively, with regard to theories of reference (Ariel, 1990, 2001; Gundel, 2003; Gundel, et al., 1993; Levinson, 1987, 1991) it was predicted that the two pronominal forms would differ in their resolution preferences, in that the personal pronoun should prefer the discourse salient focused entity, while the d-pronoun would favor the less prominent non-focused entity. The

results showed that in German as well as in Dutch, the two pronominal forms were resolved towards the second-mentioned focused entity, which suggests that in certain contexts where two potential antecedents appear, personal and d-pronouns may show overlapping functions. More importantly, however, the results indicate that discourse pragmatic information guided the resolution process. The resolution effect emerged so early that it was argued that the ambiguity had been resolved by the antecedent structure, i.e. upon encountering the pronoun, the first-mentioned non-focused entity had been backgrounded so that it was in fact not available for resolution at all. This finding was particularly interesting with regard to the results of previous pronoun resolution studies after non-canonical antecedent structures, because it suggests that the information structural cues of such structures might account (at least partly) for the findings.

In the third part of this chapter, the influence of the inherent semantic factor animacy was investigated. Since both German and Dutch pronouns allow reference to animate as well as inanimate entities, half of the items was constructed with two animate antecedents, and the other half with two inanimate antecedents. When animacy was included as a predictor into the model of the data, it significantly better predicted the outcome for both the German and Dutch data. However, a more detailed analysis of this factor on each type of pronoun revealed cross-linguistic differences of its influence. In German, it was found that animacy did not change the direction of the observed preferences, but it did have an effect on the timing of the first-mention preference for the personal pronoun: such that the resolution took place considerably earlier following two animate antecedents. This was not observed for the d-pronoun, which in fact underlines the robustness of its preference. In Dutch, the resolution preferences for the pronouns were not affected by animacy although there was an early first-mention effect after inanimate items across pronouns. Rather than an effect of pronoun resolution, this was argued to reflect the general discourse expectations in Dutch to continue the discourse with the first-mentioned inanimate entity.

In the fourth part of this chapter the on- and off-line results which were obtained in chapter 3.1 were compared. Although both tasks showed the same direction of results, the second-mention effect for the d-pronoun was only marginally significant in the off-line task. The question addressed was whether the observed variation across the tasks was due to the materials or individual participant differences. A German and a Dutch off-line control experiment revealed no influence of the experimental materials on the preferences, i.e. the control groups showed a strong preference to resolve the d-pronoun towards the second-mentioned non-topical entity. A closer look at the off-line results of the experimental groups revealed inter-individual response behaviors to the d-pronoun: 85% of the German participants and 64% of the Dutch either robustly chose the first-mentioned or the second-

mentioned antecedent[55]. To explore whether these inter-individual preferences had also affected the on-line resolution of the d-pronoun, off-line response as a predictor was included in the model. In other words, the question of whether the final choice of antecedent for the pronoun was already visible in the on-line eye-tracking data was addressed. The German data was not better explained when the off-line choice was included as a predictor: surprisingly, the second-mentioned on-line preference was also present even when the first-mentioned antecedent was ultimately chosen during the off-line task. The results of the additional analyses for German again point to the fact that the second-mentioned non-topical on-line effect for the d-pronoun was robust which underlines that the German d-pronoun is a marked form. The Dutch findings were different. For those items for which the second-mentioned entity was chosen off-line, a second-mention preference was also observed on-line. However, in the Dutch data, there was no preference on-line for those items for which the first-mentioned antecedent was chosen off-line. Thus, when the ambiguity could not be resolved on-line, but a decision needed to be made off-line, the Dutch participants chose the first-mentioned candidate. The findings further indicate the need of a systematic investigation of on- and off-line effects, first, because we would not have detected the robust second-mention preference for the German d-pronoun, had we only looked at the off-line results; and second, because it shows that in cases in which the ambiguity is not resolved, off-line tasks might push participants to make a decision which initially was not present.

[55] Alternatively, it could have been the case that the participants interpreted the d-pronoun as co-referring to the first-mentioned entity in some of the items and to the second-mentioned entity in the other items.

Resolving Ambiguous Pronouns in L2 German and Dutch

Chapter 4

4.1. Pronoun Resolution after Canonical Antecedent Structures[56]

4.1.1. Introduction

This chapter is devoted to the resolution preferences for Dutch L2 learners of German and German L2 learners of Dutch. As in the case of native speakers, I will approach this question with two tasks, one measuring final interpretation preferences and the other measuring on-line preferences as the speech signal unfolds. A central question in the study of second language processing has been whether observed differences between L1 and L2 processing should be attributed to contrasting linguistic properties between source and target languages, to different processing strategies between the first and second language, or both. German and Dutch are typologically very similar, and it was shown in chapter 3 that native speakers show very similar resolution preferences for personal and d-pronouns. Thus, from this point of view it can be hypothesized that L2 learners of both languages should be able to process the input in a target-like manner, as they would profit from L1 transfer. However, if in general, processing procedures differ between native speakers and L2 learners, then despite the similarity between the L1-L2 language pairs, processing differences might be observed. In this section, I review some basic issues in the area of L2 processing and pronoun resolution.

In the area of L2 processing, it has been shown that L2 learners are influenced by their L1 at the lexical level (Frenck-Mestre & Pynte, 1997; Juffs, 1998). For syntactic processing, the results are mixed. For example, relative clause attachment studies in SLA research are based on the observation that in sentences such as *Someone shot the servant of the actress who was on the balcony* some languages such as English and Swedish show a preference to attach the relative clause low, to the second NP of the antecedent (*the actress*) while languages such as German, Dutch, Spanish and Greek show a preference to attach the relative clause high, to NP1 (*the servant*). Such constructions allow for an investigation of cross-linguistic processing preferences. Examining the question of whether or not L2 learners would transfer their L1 parsing preference to their processing of the L2, some SLA studies reported an influence of the L1 (Dussias, 2003; Frenck-Mestre, 1997, 2002), but others did not (Felser, Roberts, Gross, & Marinis, 2003), even when native speakers of the

[56] Parts of the research presented in this chapter will appear in Ellert, M., Roberts, L. & Järvikivi, J. (in press). Verarbeitung und Disambiguierung pronominaler Referenz in der Fremdsprache Deutsch: Eine psycholinguistische Studie [Processing and disambiguating pronominal reference in German as a foreign language: A psycholinguistic study]. In Krafft, A. & Spiegel, C. [eds.] *Sprachliche Förderung und Weiterbildung - Transdisziplinär*. Frankfurt am Main: Peter Lang.

source-target language pairs had the same parsing preferences. For instance, Papadopoulou and Clahsen (2003) investigated relative clause attachment resolution in Spanish, German and Russian learners of Greek. All four languages show a preference to attach the relative clause high (*the servant*). Nevertheless, all three learner groups diverged from the native pattern of results in that they showed no preference in their L2 Greek. More evidence comes from the study of processing *wh*-dependencies, like those in long-distance questions. While some studies have found that the L2 learners perform in a native-like way (Juffs & Harrington, 1995; Williams, Möbius, & Kim, 2001), others have found a processing difference between L2 learners and native speakers (Marinis, Roberts, Felser, & Clahsen, 2005), irrespective of whether the learners' L1 forms *wh*-constructions in a similar (German, Greek) or a different way (Japanese, Chinese) from the L2 (English).

Clahsen and Felser (Clahsen & Felser, 2006a, 2006b) have suggested that differences in L1 and L2 results in grammatical processing studies are attributable to a special L2 learner processing mechanism. They argue that L2 learners rely more on lexical-semantic, pragmatic and contextual cues and less on structural ones when processing input in their L2. This *Shallow Processing Hypothesis* claims that during L2 processing, learners compute representations which contain less syntactic detail, and so are shallower, than L1 representations due to inadequacies in their L2 grammar. This effect is said to appear in on-line processing tasks, because these tasks "are believed to reduce the degree to which participants are able to draw on 'explicit' grammatical knowledge during processing", whereas native-like performance in off-line tasks "does not imply that the nature and extent of their [the learners] grammatical knowledge was [is] native-like" (Clahsen & Felser, 2006b, p.120). Thus, L2 learners are unable to compute a full syntactic representation of the L2 input due to their incomplete L2 grammar.

According to Sorace and Filiaci (2006), the observed L1-L2 processing differences can be explained by the fact that L2 learners have more difficulty with linguistic phenomena that lie at the *interfaces* (or boundaries between linguistic levels), for instance, at the interface between the syntax and the lexicon. In one study, Sorace and Filiaci (2006) investigated whether L2 learners were able to discriminate between the functions of two pronominal forms in their L2: null (or zero) pronouns and overt pronouns (= personal pronouns). Speakers of pro-drop languages, such as Italian, Spanish, Greek and Turkish, frequently use null pronouns, i.e. in finite clauses the pronoun is phonologically silent, as in the Italian example in (65)a (Carminati, 2002, p.1-2), as compared to when it is phonologically realized by the use of overt pronouns (as in (65)b). The authors tested English L2 learners of Italian, since the learners did not have this distinction in their L1, i.e. their L1 is a non-pro-drop language as opposed to Italian which is a pro-drop language.

(65)

a. Ø Mangia la mela.
He/she/it [null pronoun] eats the apple.

b. Lui/Lei mangia la mela.
He/she [overt pronoun] eats the apple.

These two pronominal forms have been claimed to have different functions in discourse, comparable to the claim for personal and d-pronouns in Germanic languages: null pronouns prefer topical antecedents whereas overt pronouns prefer non-topical antecedents; moreover that overt pronouns are marked for co-reference to non-topical entities. They have a [+Topic Shift] feature, compared to the null pronouns which are lacking this feature (Sorace, 2000). Interestingly, the English L2 learners of Sorace and Filiaci's (2006) study did not discriminate between the functions of null and overt pronouns, in that they incorrectly accepted the use of overt pronouns for co-reference to the topical entity. In sum, the learners correctly analyzed the syntactic target dependencies when resolving null pronouns (co-referring with intrasentential subjects), but they differed in the resolution of the overt pronouns, because additional discourse-based constraints were required for the resolution of overt pronouns. The authors argue that the resolution of the overt pronouns required more processing resources in comparison to null pronouns, causing problems at the Syntax-Discourse-Interface for the L2 learners.

More evidence on L1-L2 differences in the area of pronoun resolution comes from the area of L1 attrition studies which have reported an influence of the non-pro-drop language English on the use of overt pronouns in the L1 pro-drop languages Greek, Italian, Spanish and Turkish (Gürel, 2003, 2004; Kim & Montrul, 2003; Montrul, 2004; Sorace, 2000; Tsimpli, Sorace, Heycock, & Filiaci, 2004). The learners have an L2 preference for overt pronouns to co-refer with topical antecedents. Interestingly, this has been found to be the case even for a +pro-drop/+pro-drop language pair by Bini (1993), who investigated Spanish L2 learners of Italian. This finding is particularly pertinent to the study presented here, because we also test the resolution of two pronominal forms in L1-L2 language pairs that are typologically similar German-Dutch and Dutch-German. Given the results noted above, it might be that the learners nevertheless show resolution preferences which are different from those of native speakers. On the other hand, one must keep in mind that Bini's (1993) study found these non-native-like co-reference preferences in a production task. It may very well be the case that the learners are able to discriminate between the functions of null and overt pronouns in on-line comprehension, but in production, often being more demanding, they may rely on different strategies. Furthermore, it is not clear

whether the suggested differences between null and overt pronouns in pro-drop-languages are comparable to the distinction between personal and d-pronouns in Germanic languages.

Among the psycholinguistic studies on L2 pronoun resolution, very few have investigated the resolution preferences in a non-pro-drop target language. One such study is by Roberts, Gullberg and Indefrey (2008), who investigated the resolution of personal pronouns in the non-pro drop language Dutch, by highly proficient Turkish L2 learners and German L2 learners. Turkish being a pro drop language, they asked whether Turkish learners of Dutch showed different on- and off-line resolution patterns for personal pronouns (*hij/zij*), when compared to L2 learners from a non-pro drop language background like German. They used two off-line measures, an acceptability judgment task and a comprehension questionnaire, and an on-line measure, eye-tracking during reading. They presented their materials in one ambiguous condition presenting two potential antecedents (66) and two unambiguous conditions to which the reading measures of the ambiguous condition could be compared.

(66) Peter en Hans zitten in het kantoor. Terwijl Peter aan het werk is, eet hij een boterham. Het is een rustige dag.
Peter and Hans are in the office. While Peter is working, he is eating a sandwich. It is a quiet day.

Comprehension Questionnaire
Er wordt een boterham gegeten door _____ .
A sandwich is eaten by _____ .

The results of the off-line comprehension questionnaire revealed an L1 influence on the L2 pronoun interpretation: the German learners patterned like the natives choosing the intra-sentential NP (*Peter*) as co-referring to a subsequent subject pronoun in more than 90% of the cases, the Turkish learners showed no preference between the intra- and extra-sentential NP (*Peter = Hans*). However, when measuring eye-movements during reading, Roberts et al. found that both learner groups patterned together; they did not show a preference (*Peter = Hans*) while the native Dutch readers showed the same intra-sentential NP (*Peter*) preference as in the off-line task. Although, both learner groups arrived at distinct final interpretations, they both showed the same L2 processing pattern. Since there was syntactic ambiguity in this condition, the authors argued that the learners needed more time than native Dutch speakers to coordinate both syntactic and discourse information on-line.

With a visual-world eye-tracking task, Wilson (2009) investigated the influence of grammatical role and order of mention information on the resolution of German personal

and d-pronouns in highly advanced English-speaking L2 learners (European Reference Frame C1). To disentangle order of mention and grammatical role, she presented the pronouns either after SVO or OVS antecedent structures, such as in (67)[57]. The primary focus will be on the results obtained for pronoun resolution after the canonical antecedent structures, as the studies presented in this section also investigate pronoun resolution preferences after canonical antecedent structures (for a more detailed discussion on the non-canonical OVS results, see chapter 4.2).

(67)

a. SVO
   Der Kellner erkennt den Detektiv, als das Bier umgekippt wird. **Er/Der** ist offensichtlich sehr fleißig.
   The waiter$_{NOM}$ recognizes the detective$_{ACC}$ as the beer is tipped over. ***He*** *[P/D] is clearly very hard working.*

b. OVS
   Den Kellner erkennt der Detektiv, als das Bier umgekippt wird. **Er/Der** ist offensichtlich sehr fleißig.
   The waiter$_{ACC}$ recognizes the detective$_{NOM}$ as the beer is tipped over. ***He*** *[P/D] is clearly very hard working.*

Wilson found that the L2 learners differed from native German speakers in the resolution of the d-pronoun: while the native speakers resolved it towards the second-mentioned entity after SVO antecedent structures (*the detective$_{ACC}$*), the learners showed no preference at all. For the personal pronoun the learners had a first-mention preference (*the waiter$_{NOM}$*) while the native Germans had none (see discussion chapter 3.1). These results are in line with the findings of the studies discussed above on the interpretation of null and overt pronouns in a pro-drop target language (e.g. Sorace & Filiaci, 2006) which found non-target-like topic co-reference relations for overt pronouns, and which were explained due to integration difficulties at the syntax-discourse interface.

Similarly, Roberts, Järvikivi, Ellert and Schumacher (in prog.), investigated the L2 resolution preferences in highly proficient Finnish L2 learners of German (European Reference Frame C1-C2). The use of demonstrative pronouns in Finnish is comparable to the German use of d-pronouns (Kaiser & Trueswell, 2008). Thus, the learners were assumed to be familiar with the difference in functionality of personal and d-pronouns. The Finnish L2 learners did indeed treat the two pronouns differently, but they did not perform

[57] The example is the same as in (58), repeated here for convenience.

identically to the native speakers. After canonical SVO antecedent structures, the learners showed a target-like second-mention preference for the d-pronoun, but they showed no preference for the personal pronoun. These results indicate that the L2 learners had no difficulty in assigning the non-topical entity to the d-pronoun but that the personal pronoun remained ambiguous. This supports the claim that the d-pronoun is marked for co-reference with the non-topical antecedent, while the personal pronoun is more ambiguous in general, as it is neutral in this regard, as was found for German native speakers (chapter 3.1), and in this way, the learner preferences are understood to reveal linguistic properties of the target language. Furthermore, the results are particularly important to the experiments presented here, because they suggest that when the L2 learners are familiar with the different discourse functions of the two pronominal forms because the source language shares this feature (unlike Wilson's (2009) English L2 learners of German), they might be able to engage in target-like pronoun resolution, and differentiate between the functions of personal and d-pronouns.

Summing up, the findings of previous studies on L2 preferences in pronominal co-reference relations, point into two directions. First, studies investigating the L2 co-reference preferences of null and overt pronouns in pro-drop target languages by L2 learners from non-pro drop and pro-drop languages suggest that L2 learners may in general have difficulty discriminating between the functions of two pronominal forms. Comparable L2 learner processing effects were found for Turkish and German learners of Dutch in that they patterned together and showed no resolution preferences when reading personal pronouns in ambiguous contexts, in contrast to the target-like behavior of the German learners of Dutch during off-line comprehension. These results suggest that L2 learners may not show native-like behavior in L2 pronoun resolution even when source and target language preferences pattern together. On the other hand, the findings of previous studies on the L2 resolution of German personal and d-pronouns showed that Finnish learners resolved d-pronouns in a target-like manner unlike English L2 learners of German. This difference may be due to possible L1 influences. Given the lack of definitive results on this issue, the present study investigates the resolution preferences for personal and d-pronouns in Dutch learners of German and German learners of Dutch. This bidirectional perspective will enable us to draw better conclusions about the generalizability across L1-L2 language pairings.

4.1.2. The present study

The aim of this study is to investigate the resolution of personal and d-pronouns in L2 German and Dutch. Do Dutch L2 learners of German and German L2 learners of Dutch process these pronouns in a target-like manner? Do they use the order of mention

information to disambiguate co-reference relationships when resolving personal and d-pronouns in discourse? Have they acquired the difference in functions between personal and d-pronouns in ambiguous contexts? Since there is also evidence that the processing of L2 discourse dependencies may differ from L1 processing even when the source and target languages pattern together, we may also ask whether both L2 learner groups (Dutch-German, German-Dutch) show comparable resolution preferences across languages.

As with the native German and Dutch experiments reported above, a visual-world eye-tracking task as well as an off-line forced choice comprehension questionnaire was employed. The application of the visual-world paradigm to the study of L2 acquisition is particularly valuable, in that the learners' final interpretation preferences as well as their on-line resolution preferences can be studied. As has been mentioned by Clahsen and Felser (2006b), even when L2 off-line preferences are target-like, processing behavior might differ. Furthermore, compared to eye-tracking during reading, the visual world paradigm measures the processing of spoken language as it unfolds. Thus, moment-by-moment inferences become measureable. Furthermore, as the pictures provide additional non-linguistic information, the task might be easier for L2 learners than a reading task (Tanenhaus, Magnuson, Dahan, & Chambers, 2000 for a discussion on the effect of non-linguistic contextual cues affecting online speech processing; Tanenhaus, et al., 1995).

The following research questions will be investigated in this section:

1. How does the order of mention of the antecedent candidates influence the resolution of personal and d-pronouns in L2 German and Dutch?
2. Do L2 listeners show the same resolution preferences for personal and d-pronouns as native speakers? (And if not, in what ways do they differ?)
3. Do Dutch L2 learners of German and German L2 learners of Dutch resolve personal and d-pronouns in the same way?

4.1.3. Pronoun resolution in Dutch L2 learners of German

The first experiment investigates the resolution of personal and d-pronouns in Dutch learners of German.

4.1.3.1. Methods

Participants

Thirty-two Dutch learners of German (25 female, 7 male) took part in the experiment. They were paid for their participation. All participants had normal or corrected-to-normal vision. The L2 learners took the standardized German Placement Test from the Goethe Institute

and filled out a language background questionnaire in which they provided information on their Dutch knowledge.

The participants had learned German at high school, on average between 3 and 5 years[58] and lived in the Netherlands at the time of testing. They were all students at the Radboud University Nijmegen and were aged between 18 and 29 years (mean = 21.69; SD = 3.04) and had started learning German at the age of 12 (mean = 12.22; SD = 3.26). The participants rated their German skills on a 5-point-scale for speaking, comprehension, writing, reading, grammar and pronunciation. Taking all of these scores together, the median for self-reported proficiency was 3, indicating a mediocre self-reported language level.

	Speaking	Comprehension	Writing	Reading	Grammar	Pronunciation
Median	3.5	2	4	3	4	3

Table 4.1: Medians of the self-reported level of German skills for the Dutch learners of German (1 = very good, 2 = good, 3 = mediocre, 4 = little knowledge, 5 = poor)

The German Placement Test developed by the Goethe Institute tested the learners for their proficiency level according to the European Reference Frame (see Table 0.19 in the Appendix). On average, the Dutch learner group scored 17.16 (SD = 4.10) of a possible maximum score of 30, i.e. on average they were on level B2 which means that they were advanced learners of German.

Experimental Tasks

The experimental tasks were identical to those in the previous experiments. After the experiment, the L2 participants were asked to perform a naming task on the computer screen which presented a lexical item one at a time and three pictures additionally to a "don't know" button. The participants were asked to choose the picture which corresponded to the lexical item or otherwise to click on "don't know". This was done to control for their lexical knowledge of the antecedents presented in the experimental items. If they did not correctly identify either one of the antecedents belonging to one item, all fixations to that item were excluded from the analysis. On average, the Dutch learners of German correctly identified 45 of 48 (mean = 44.86; SD = 2.24) antecedent noun phrases. 528 fixations were removed from the analysis resulting in a total of 4754 fixations.

The German Placement Test and the language background questionnaire were completed

[58] The profile questionnaire was a web-based questionnaire, which asked the participants to indicate via a radio button box for how long they had been learning German. The following categories were possible answers: 0-1 year, 1-2 years, 2-3 years, 3-5 years, 5-10 years, more than 10 years. The median fell into the 3-5 years category.

after the experiment. The whole session took about 70 min.

Materials and design

The materials for the Dutch learners of German were the same as used with the native German speakers (chapter 3.1), here repeated for convenience.

(68) Conditions: Sample Item

a. Personal Pronoun Condition
Der Schrank ist schwerer als der Tisch. **Er** stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.
The cupboard is heavier than the table. ***It*** *[P] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

b. D-pronoun Condition
Der Schrank ist schwerer als der Tisch. **Der** stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.
The cupboard is heavier than the table. ***It*** *[D] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

Procedure

The procedure was identical to that of the German experiment with the native listeners. We used the same eye tracker as in the experiment with the German listeners, an SR Research EYELINK II eye tracker. The dominant eye was recorded. A sampling rate of 500-Hz was used which monitored gaze locations every 2 ms. The calibration of the camera which links the position of the eyes with a certain location on the screen, ensured that spatial accuracy was at least 0.5°.

The accuracy of the responses to the comprehension questions was very high with 89% correct answers (24 questions; mean correct answers = 21.25, SD = 2.02). This indicates two things: the participants were listening, and they were able to correctly understand the items.

4.1.3.2. Results

Forced Choice Questionnaire

Figure 4.1 and Table 4.2 show that surprisingly the Dutch learners of German had a preference to choose the first-mentioned antecedent for *er* (96%) as well as for *der* (84%). A Chi Square test revealed however that there was a significant association between the

type of pronoun and the antecedent that was chosen, χ^2 (1) = 31.19, p<.001 indicating that despite the overall first-mention preference, the quantitative distribution of responses was different. Based on the odds ratio, the odds of choosing the first-mentioned antecedent were 4.86 times higher for the personal pronoun than for the d-pronoun. The pattern of results for both pronouns were different from chance, for the personal pronoun χ^2 (1) = 210.18, p<.001, and for the d-pronoun χ^2 (1) = 102.50, p<.001. Compared to chance, the first-mentioned antecedent was 26.43 times more likely to be chosen when a personal pronoun was present and 5.44 when a d-pronoun was present.

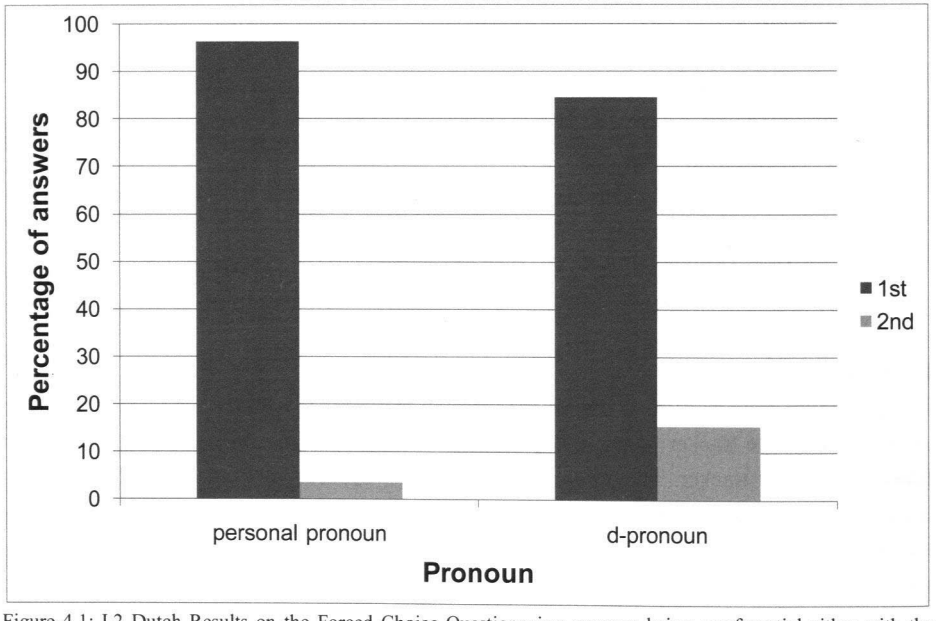

Figure 4.1: L2 Dutch Results on the Forced Choice Questionnaire; pronoun being coreferential either with the first- or the second-mentioned entity (1st, 2nd) in the two conditions (personal pronoun, d-pronoun)

	1st	2nd
personal pronoun	96.35% (370)	3.65% (14)
d-pronoun	84.47% (321)	15.53% (59)

Table 4.2: L2 Dutch Results on Forced Choice Questionnaire; pp = 32; items = 24; 4 *der*-answers eliminated from analysis because both referents or a post pronominal referent was chosen

Thus, for both pronouns we observe a first-mention preference in the learners' data; this preference is slightly stronger for personal pronouns than for d-pronouns.

Visual-World Eye-Tracking

Data Analysis

The data analysis was done as in the German native experiment (see chapter 3.1.3). 25 looks (< 1%) to either target picture that started before pronoun onset (-200 till 0 ms) were excluded from the analysis, resulting in a total of 4729 fixations which entered the analysis.

Overview of the eye movements

Figure 4.2 shows that all looks start to rise at 200 ms, which is the time needed to launch a saccade.

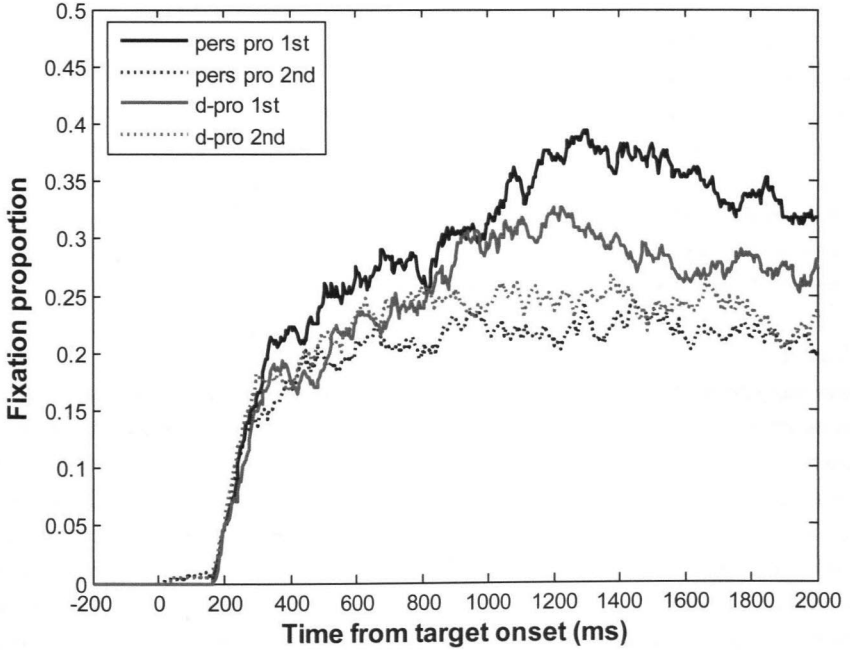

Figure 4.2: Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Personal pronoun. For the personal pronoun *er*, at 400 ms after pronoun onset there are slightly more looks to the first-mentioned target than to the second-mentioned target. At about 600 ms after pronoun onset, the looks to the first-mentioned target are still rising, while the looks to the second-mentioned entity remain at a 20%-level. At about 1300 ms the looks to the first-mentioned target start to decrease again, however the first-mention

preference remains visible until the end of the analysis window (2000 ms). Overall there is a first-mention preference for the personal pronoun *er*.

D-pronoun. For the d-pronoun *der*, we observe an increase in looks to both targets until about 900 ms after pronoun onset, after which the looks to the first-mentioned entity still increase while the looks to the second-mentioned target persist at a 25%-level. The looks to the first-mentioned target decrease again at about 1200 ms after pronoun onset, and from 1400 ms onwards there are only slightly more looks to the first-mentioned entity than to the second. This pattern reflects a relatively long period of ambiguity for the d-pronoun, after which, as for the personal pronoun, there is a tendency to resolve the d-pronoun to the first-mentioned target; however, the first-mention preference for *der* is not as clear.

Statistical Analysis of the eye movement patterns

Loglikelihood tests revealed that the *interaction model*, which contained the *condition x mention* interaction term, explained the data marginally significantly better in time windows 4 and 6 (600 – 800 ms; 1000 – 1200 ms), and significantly better in time windows 7, 8, and 9 (1200 – 1800 ms) than the *simple model*, containing the main effects only (see Table 0.12 in the Appendix). This means that although, as can be seen in Figure 4.2, the first-mentioned target was preferred overall, the pattern of results was differently affected by the type of pronoun in these three time windows.

The fixed effects analysis showed marginally significant interaction between type of pronoun and order of mention in time window 4 (600 – 800 ms), and a marginally significant main effect of condition (see Table 4.3). Given the low t-values for the effects (especially the main effect), and considering the fact that the interaction term predicted the data only marginally better (χ^2 (1) = 3.08, p = 0.079) than when it was not included in the model, these effects may not represent pronoun resolution preferences[59]. During time window 5 (800 – 1000 ms), there was a highly significant main effect of mention which was due to an overall first-mention preference. The interaction term was either marginally significant or significant during time windows 6, 7, 8 and 9 (1000 – 1800 ms) due to the higher amount of looks to the first-mentioned target picture for the personal pronoun. The main effect of mention was marginally significant or significant during time window 6 and

[59] Actually, the fixed effects which would be obtained by the *simple model* for this time window show no significant effect at all.

Time window	in ms	Fixed predictors		
		pronoun condition	order of mention	interaction
4	600-800	0.12 (0.565)	-0.32 (-1.553)	

7 (1000 – 1400 ms), and the main effect of pronoun condition was significant during time windows 7, 8 and 9 (1200 – 1800 ms). During the last time window (1800 – 2000 ms), there was a highly significant main effect of order of mention, which was again motivated by an overall first-mention preference for the two types of pronouns.

Time window	in ms	Fixed predictors		
		pronoun condition	order of mention	interaction
1	0-200	-0.02 (-0.434)	0.04 (0.805)	
2	200-400	-0.02 (-0.094)	-0.08 (-0.454)	
3	400-600	0.08 (0.415)	-0.26 (-1.322)	
4	600-800	0.47 (1.64)†	0.03 (0.116)	-0.72 (-1.754)†
5	800-1000	0.05 (0.224)	-0.63 (-2.987)**	
6	1000-1200	0.42 (1.399)	-0.6 (-1.997)*	-0.79 (-1.835)†
7	1200-1400	0.71 (2.299)*	-0.54 (-1.767)†	-0.96 (-2.214)*
8	1400-1600	0.8 (2.616)**	-0.37 (-1.239)	-1.06 (-2.465)*
9	1600-1800	0.63 (2.082)*	-0.33 (-1.118)	-0.9 (-2.103)*
10	1800-2000	0.21 (1.009)	-0.82 (-3.922)***	

Table 4.3: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*er* vs. *der*) and order of mention (1st vs. 2nd). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

As these analyses suggest, personal and d-pronouns were affected differently by the order of mention information of the antecedents (see Table 4.4) although both favor a first-mentioned interpretation. In order to get a better understanding on how order of mention influenced the resolution of the two pronominal forms, individual analyses were conducted with order of mention as a predictor. For the *er*-condition the analyses revealed a marginally significant main effect of order of mention during time window 3 (400 – 600 ms), which became significant in the subsequent time window and lasted until the last time window (600 - 2000 ms). The beta coefficient was negative, indicating the expected first-mention preference.

For the d-pronoun *der*, the mixed models revealed a significant main effect of order of mention during time window 6, which was marginally significant in time windows 7 and 10 (1000 – 1400 ms; 1800 – 2000 ms), the negativity of the beta coefficients of mention also indicated a first-mention preference for this pronoun.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	0.01 (0.099)	0.08 (1.036)
2	200-400	-0.25 (-1.03)	0.09 (0.369)
3	400-600	-0.52 (-1.813)†	-0.01 (-0.045)
4	600-800	-0.68 (-2.319)*	0.03 (0.115)
5	800-1000	-0.9 (-2.993)**	-0.37 (-1.247)
6	1000-1200	-1.39 (-4.528)***	-0.6 (-1.995)*
7	1200-1400	-1.5 (-4.803)***	-0.54 (-1.776)†
8	1400-1600	-1.44 (-4.619)***	-0.37 (-1.25)
9	1600-1800	-1.23 (-4.01)***	-0.33 (-1.125)
10	1800-2000	-1.15 (-3.78)***	-0.52 (-1.766)†

Table 4.4: Results of the individual time course analyses for each type of pronoun (*er* and *der*). Order of mention (1st vs 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

4.1.3.3. Discussion

The eye-movement measures of the Dutch learners of German showed a clear effect of order of mention. After hearing the personal pronoun, participants fixated more often pictures of the first-mentioned than the second-mentioned antecedent. Surprisingly, this was also true for the d-pronoun. The same pattern of results was obtained in the off-line questionnaire. There was a highly significant first-mention effect for *er* (96%), and a highly significant first-mention effect for *der* (85%). Overall, the pattern of results suggests that participants had a preference to interpret both pronouns as referring to the first-mentioned topical entity.

Unlike the German native speakers, the learners thus have a different resolution preference for d-pronouns. The timing of the effects was also different. While native listeners took a longer period to disambiguate the personal pronoun (1400 ms) due to its unmarkedness, the second language learners disambiguated the personal pronoun immediately (400 - 600 ms), and it was the d- pronoun which was more ambiguous.

The question arises as to whether the L2 learners were in fact able to resolve the d-pronoun, perhaps because they confounded it with other forms. In other words, in spoken language the personal and the d-pronoun have a similar phonological realization (*er* - *der*) which might be difficult for a second language listener to discriminate. Also, the pronominal form *der* is homophonous with the masculine singular definite article in German. There is a

possibility that the learners might have confounded it with the definite article. While the phonological realization of the forms is different, and thus should not be confounded by a native speaker, this minimal difference might not have been noticed by the learners who were waiting for the noun phrase yet to be heard in order to resolve the full referring expression. This could have resulted in a longer period of ambiguity. However, there are two reasons why it is unlikely that the d-pronoun was confounded with other forms. Firstly, the eye-movement results revealed a significant *condition x mention* interaction, thus indicating a qualitative resolution difference. Second, during the off-line task they read the full sentences, before making a choice. Yet, the off-line results showed the same preferences, reducing the likelihood of this explanation. Thus it is arguable that the learners did indeed discriminate between the two pronominal forms, but not between their functions: that is, they had two forms for one function, i.e. to establish co-reference with the first-mentioned topical entity.

Another possibility is that the d-pronoun was more ambiguous for the Dutch learners because they were not so familiar with the form. All had learned the German language at High School and were living in the Netherlands at the time of testing. The d-pronoun is mainly found in spoken language and its use is not a central issue in teaching German as a foreign language (Ahrenholz, 2007). Therefore the learners might have been lacking the form in their input, and were uncertain of how to resolve it. Note, however that Roberts et al.'s (in prog.) results provide evidence against this explanation, since the Finnish learners of German were tutored learners like the Dutch learners of this study, and yet they found the personal pronoun more ambiguous than the d-pronoun.

Choosing the most salient first-mentioned entity may have been a learner strategy which was applied to the resolution of the d-pronoun. Order of mention cannot be disentangled from topicality information with our materials; but it is very likely that the learners choose the first-mentioned entity by default, since it is the topical entity and it is more likely to be mentioned again in the subsequent discourse.

Our results show that Dutch L2 learners of German show non-target-like behavior when resolving pronouns in their L2, despite the typological similarity between source and target language. The findings indicate that the Dutch learners had two pronominal forms for one function, i.e. co-reference to the first-mentioned topical entity which is in line with previous results on L2 preferences in pronominal co-reference relations (Sorace & Filiaci, 2006; Wilson, 2009). The non-target-like L2 processing preference might have occurred due to learner difficulties at the syntax-discourse interface. Our results diverge from Roberts et al.'s (in prog.) findings who found target-like resolution behavior for d-pronouns by

Finnish L2 learners of German. This difference might be due to the higher proficiency level of their learners. While the Dutch learners of this study were on average on level B2 according to the European Frame of Reference, Roberts et al.'s Finnish learners of German were on levels C1-C2. Proficiency might thus have an influence on the resolution of pronouns in a second language. In order to further analyze this issue, in chapter 4.4 additional analyses are provided which include proficiency as a predictor in the mixed model.

The observation that L2 learners favor pronominal forms in general to co-refer to topical antecedents at least in production has been made by Klein and Perdue (1997). In the cross-linguistic longitudinal *Second Language Acquisition by Adult Immigrants* project, which took place between 1981 to 1988, L2 data were collected from 40 untutored learners. All learners used a language system which was called the *Basic Variety* and characterized as follows:

"[The system]

- seemed to be determined by the interaction of a small number of organizational principles,
- was largely (though not totally) independent of the specifics of source and target language organization,
- was simple, versatile and highly efficient for most communicative purposes. (p.303)"

Among the *Basic Variety* principles for utterance organization, Klein and Perdue (1997) mention pragmatic constraints which reflect principles of reference introduction or maintenance, and topic-focus organization. They state that in sentences containing two referents which are introduced by "names and lexical nouns, the topic/focus status of the referent is indicated solely by position. It follows from the observed distribution that reference maintenance in focus cannot be achieved by pronominal means (p.319)". A possible explanation for our L2-resolution pattern may be that our L2 learners were using principles of the *Basic Variety* language system. In our materials, the L2 learners heard an antecedent sentence containing two lexical noun phrases. The first noun phrase was the topical referent as indicated by its position. The Dutch L2 learners of German interpreted both pronouns as referring to the first-mentioned topical entity, since their language system did not license pronominal use to refer to the non-topical entity. From this perspective, it is interesting to re-analyze the data with regard to the learners' proficiency levels (chapter 4.4): it may enable us to tear apart the resolution preference of L2 learners following *Basic Variety* principles from L2 learners who might have passed to the next variety.

Summing up, Dutch L2 learners of German do not differentiate between the functions of personal pronouns and d-pronouns. This is most remarkable, since the Dutch learners come from a language background which also encodes this difference. Thus, they are not able to benefit from positive L1 transfer in L2 pronoun resolution. In the next section, we will report an experiment which used the same set-up and (translated) materials with German L2 learners of Dutch. It is predicted that the German L2 learners of Dutch will show the same resolution preferences as the Dutch L2 learners of German, namely an overall first-mentioned topical preference for the two types of pronouns.

4.1.4. Pronoun resolution in German L2 learners of Dutch

4.1.4.1. Methods

Participants

Thirty-two German learners of Dutch (16 female, 16 male) took part in the experiment. The participants were aged between 17 and 26 years (mean = 18.03; SD = 1.81). On average, the length of exposure to Dutch was 12 months (mean = 12.25; SD = 8.13). The participants were paid for their participation. All participants had normal or corrected-to-normal vision.

Seven learners were students at the Radboud University Nijmegen and lived in the Netherlands at the time of testing. They had learned Dutch during an intensive language course at the Radboud University Nijmegen which lasted about 5 weeks (mean age of onset = 20; SD = 1.41) and at the end of which they had all passed the NT2-exam (*Dutch as a second language*) attesting them a level of high proficiency in Dutch. The participants reported how many hours a week they spoke, heard, read and wrote Dutch. Table 4.5 provides an overview of the median amount of time spent with the Dutch language (because of outliers we will report the median).

	Speaking	Listening	Reading	Writing
Median	30	30	15	20

Table 4.5: Median amount of hours per week spent on Dutch by the German L2 students

The participants rated their Dutch skills on a 5-point-scale for speaking, comprehension, writing, reading, grammar and pronunciation. Taking all of these scores together, the median for self-reported proficiency was 3, indicating a mediocre self-reported language level.

	Speaking	Comprehension	Writing	Reading	Grammar	Pronunciation
Median	3	2	3	2	3	3

Table 4.6: Medians of the self-reported level of Dutch skills for the student German learners of Dutch (1 = very good, 2 = good, 3 = mediocre, 4 = little knowledge, 5 = poor)

Twenty-five German learners of Dutch were pupils at the "Freiherr-vom-Stein-Gymnasium", a high school in the German town Kleve (close to the Dutch border), who had been learning Dutch during school lessons for an average of 14 months (mean = 13.72; SD = 4.19). Their mean age of onset was 16 (mean = 16.2; SD = 0.63). The pupils spent a lower number of hours per week with the Dutch language than the students (see Table 4.7), because they mainly had language contact at school, while the students were living in the Netherlands at the point of testing. Nevertheless, the self-rated Dutch proficiency scores were the same across pupils and students with a mediocre self-reported language level (see Table 4.8).

	Speaking	Listening	Reading	Writing
Median	2	1.5	1	1

Table 4.7: Median amount of hours per week spent on Dutch by the German L2 pupils

	Speaking	Comprehension	Writing	Reading	Grammar	Pronunciation
Median	3	2	3	2	3	3

Table 4.8: Medians of the self-reported level of Dutch skills for the German learners of Dutch (1 = very good, 2 = good, 3 = mediocre, 4 = little knowledge, 5 = poor)

Experimental Tasks

The tasks were identical to those in the native experiment. Additionally, after the experiment, the German learners of Dutch were asked to perform a naming task on the computer screen. Due to not correctly identifying some of the antecedents, 342 fixations were removed resulting in a total of 5096 fixations to be analyzed. The German learners of Dutch knew 46 of the 48 experimental antecedent noun phrases (mean = 46.16; SD= 1.94). The students named 47 antecedents correctly (mean = 47.43; SD= 0.79), and the pupils 46 (mean = 45.80; SD= 2.02). That means both groups were similarly familiar with the lexical items. At the end of the session a language background questionnaire was completed. The whole session took about 60 minutes.

Materials and design

The materials for the German learners of Dutch were the same as used with the native Dutch speakers (chapter 3.1).

(69) Conditions: Sample Item

a. Personal Pronoun Condition
De kast is zwaarder dan de tafel. **Hij** is afkomstig uit een meubelwinkel in België. De sofa zal volgende week geleverd worden.
The cupboard is heavier than the table. ***It*** *[P] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

b. D-pronoun Condition
De kast is zwaarder dan de tafel. **Die** is afkomstig uit een meubelwinkel in België. De sofa zal volgende week geleverd worden.
The cupboard is heavier than the table. ***It*** *[D] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

Procedure

The procedure was identical to that of the Dutch experiment with the native listeners. We used the same eye tracker as in the experiment with the Dutch listeners, a portable SR Research EYELINK II eye tracker. The dominant eye was recorded. A sampling rate of at least 250-Hz was used which monitored gaze locations every 4 ms. The calibration of the camera which links the position of the eyes with a certain location on the screen, ensured that spatial accuracy was at least 0.5°.

The accuracy of the responses to the comprehension questions was very high with 88% correct answers (24 questions; mean correct answers = 21.03, SD =1.86). This percentage of correct responses was comparable to the Dutch learners of German (89%).

Nevertheless, there was a difference between both groups. We calculated an independent t-test on the accuracy scores[60]. On average, the student participants scored higher on the content questions (mean = 22.57, SE = 0.43) than the pupils (mean = 20.60, SE = 0.37). Although, this difference was significant $t(16.032) = 3.483$, $p<.01$, and represented a large-sized effect $r = .66$, both groups scored very high on the content questions. This means that all participants were able to correctly understand the items, with the students being even more accurate than the pupils.

[60] The results should be regarded with caution, since we only had 7 participants in the students group.

4.1.4.2. Results

Forced Choice Questionnaire

Figure 4.3 and Table 4.9 show a first-mention preference for the personal pronoun *hij* (95%), and a tendency to choose the second-mentioned antecedent for the d-pronoun *die*. But this preference is not very strong (59%). The association between type of pronoun and antecedent choice was significant, χ^2 (1) = 255.68, p<.001. The odds ratio showed that choosing the first-mentioned antecedent was 29.94 times more likely for the personal pronoun than for the d-pronoun. The patterns of results for both pronouns were different from chance, for the personal pronoun χ^2 (1) = 192.76, p<.001, and for the d-pronoun χ^2 (1) = 6.50, p<.01. The first-mentioned antecedent was 18.20 times more likely to be chosen when a personal pronoun was present compared to chance. In case of a d-pronoun, the odds of choosing the second-mentioned entity were 1.65 times higher than predicted by chance.

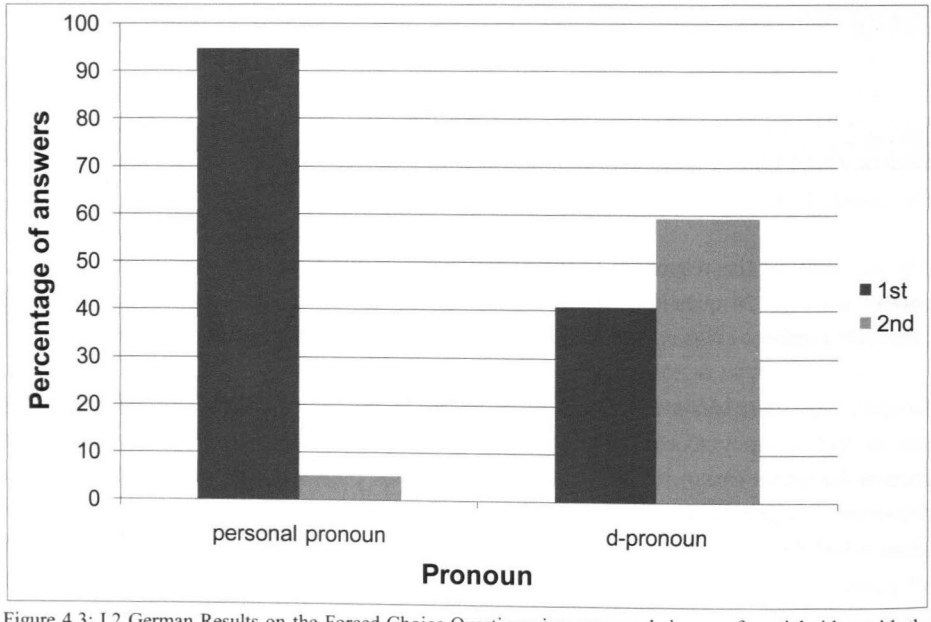

Figure 4.3: L2 German Results on the Forced Choice Questionnaire; pronoun being coreferential either with the first- or the second-mentioned entity (1st, 2nd) in the two conditions (personal pronoun, d-pronoun)

	1st	2nd
personal pronoun	94.79% (364)	5.21% (20)
d-pronoun	40.79% (155)	59.21% (225)

Table 4.9: L2 German Results on Forced Choice Questionnaire; pp = 32; items = 24; 4 *die*-answers eliminated from analysis due to 3 missing answers and 1 answer where both referents were chosen

Summing up, the German learners of Dutch had a preference to resolve the personal pronoun towards the first-mentioned entity, while they showed a second-mention preference for the d-pronoun.

Visual-World Eye-Tracking

Data Analysis

The data analysis was conducted as in the German native experiment. 50 looks (< 1%) to either target picture that started before pronoun onset (-200 till 0 ms) were excluded from the analysis, resulting in a total of 5046 fixations which entered the analysis.

Overview of the eye movements

As can be seen in Figure 4.4, the looks to all target pictures rise at 200 ms. The looks to the first-mentioned targets rise more than to the second-mentioned targets until about 300 ms. However, this does not seem to be due to pronoun resolution, but rather suggests an anticipatory effect.

Personal Pronoun. For the personal pronoun, the looks to the first-mentioned entity continue rising until they reach a 35%-level at about 1200 ms. The looks to the second-mentioned target also rise, but do not cross the 20%-level, indicating that there is a first-mention effect for the personal pronoun which persists until the end of the analysis period (2000 ms).

D-Pronoun. For the d-pronoun, the looks to both target pictures are increasing until they cross the 20%-level at 1200 ms after pronoun onset. After this point the looks to the second-mentioned target are still increasing until they reach a 30%-level at about 1800 ms, while the looks to the first-mentioned entity decrease during the last time windows. As a whole, the pattern of looks suggests that the d-pronoun is disambiguated relatively late and towards the second-mentioned entity.

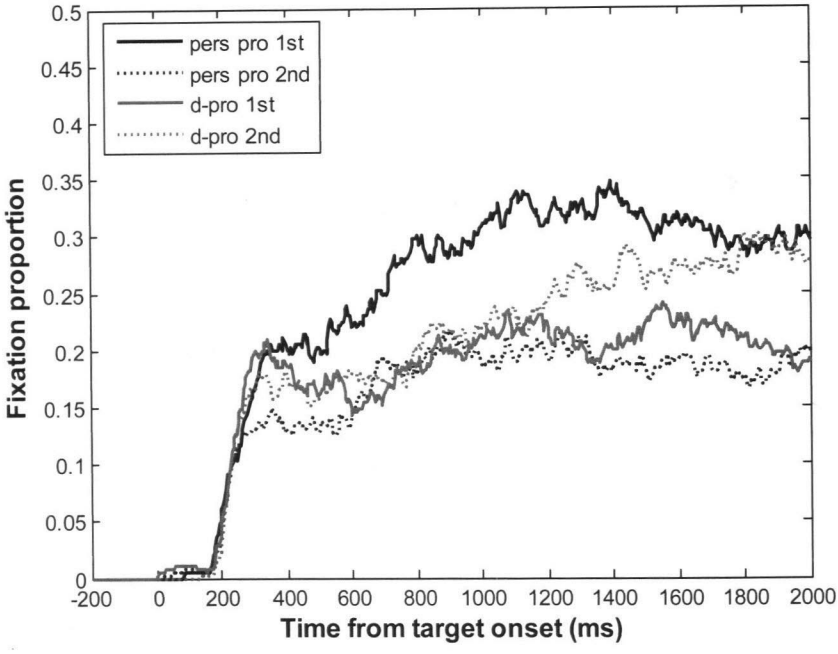

Figure 4.4: Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Statistical Analysis of the eye movement patterns

The *interaction model* containing the interaction term *condition x mention* explained the data better than the *simple model* containing condition and mention as fixed factors for time windows 4 to 10 (600 – 2000 ms) (see Table 0.13 in the Appendix).

The analysis of fixed factors (see Table 4.10) showed a marginally significant main effect of condition in the first time window (0 – 200 ms). Due to the low amount of overall looks in this time window (less than 5%), this rather seemed to be a spurious effect. During time windows 2 and 3 (200 – 600 ms), there was a marginally significant and then significant main effect of mention. There were more looks to the first-mentioned entity than to the second. Throughout time windows 4 to 8 (600 – 1600 ms), we observed a significant interaction between type of pronoun and order of mention, due to a first-mention preference for the personal pronoun. There was also a highly significant main effect of condition, with its beta coefficient indicating that there were more looks for the personal pronoun than for the d-pronoun. During time window 9 and 10 (1600 – 2000 ms), there was a marginally

significant and then significant main effect of mention. The positivity of the beta coefficient indicated that there were overall more looks to the second-mentioned than the first-mentioned entity.

Time window	in ms	Fixed predictors		
		pronoun condition	order of mention	interaction
1	0-200	0.1 (1.747)†	-0.07 (-1.162)	
2	200-400	-0.16 (-1.009)	-0.27 (-1.714)†	
3	400-600	0.02 (0.138)	-0.4 (-2.359)*	
4	600-800	0.66 (2.628)*	0.06 (0.237)	-0.78 (-2.177)*
5	800-1000	0.88 (3.32)**	-0.04 (-0.148)	-0.93 (-2.481)*
6	1000-1200	0.93 (3.415)***	-0.14 (-0.526)	-1.11 (-2.872)**
7	1200-1400	1 (3.646)***	0.27 (0.992)	-1.52 (-3.914)***
8	1400-1600	0.97 (3.488)**	0.4 (1.429)	-1.74 (-4.42)***
9	1600-1800	0.85 (3.114)**	0.53 (1.937)†	-1.73 (-4.459)***
10	1800-2000	0.93 (3.426)***	0.93 (3.408)***	-2.01 (-5.222)***

Table 4.10: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*hij* vs. *die*) and order of mention (1st vs. 2nd). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

A mixed model with only mention as a fixed factor was conducted for each pronoun type (see Table 4.11). For *hij*, the model revealed a highly significant main effect of order of mention between 400 ms and 2000 ms after pronoun onset. The negativity of the beta coefficient indicated a first-mention effect.

For the d-pronoun *die*, there was a marginally significant second-mention effect in time window 9 (1600 – 1800 ms) which became highly significant in the last time window (1800 – 2000 ms). This effect is rather late, showing that the d-pronoun initiates a relatively long period of ambiguity.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	-0.04 (-0.426)	-0.1 (-1.253)
2	200-400	-0.35 (-1.582)	-0.19 (-0.846)
3	400-600	-0.67 (-2.819)**	-0.13 (-0.541)
4	600-800	-0.72 (-2.758)**	0.06 (0.244)
5	800-1000	-0.97 (-3.538)**	-0.04 (-0.152)
6	1000-1200	-1.25 (-4.504)***	-0.14 (-0.534)
7	1200-1400	-1.24 (-4.465)***	0.27 (1.004)
8	1400-1600	-1.34 (-4.808)***	0.4 (1.426)
9	1600-1800	-1.2 (-4.327)***	0.53 (1.936)†
10	1800-2000	-1.08 (-4.005)***	0.93 (3.364)**

Table 4.11: Results of the individual time course analyses for each type of pronoun (*hij* and *die*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

Overall, the statistical analysis of the L2 data indicated an early and clear first-mention effect for the personal pronoun *hij*, and a relatively late second-mention effect for the d-pronoun *die*.

4.1.4.3. Discussion

The eye-movement data of the German learners of Dutch revealed a clear and early first-mention effect for the personal pronoun *hij* (400 ms) and a late second-mention preference for the d-pronoun *die* (1600 ms). This late effect for the d-pronoun indicated a high amount of ambiguity for the d-pronoun. The off-line results are consistent with the eye-tracking results, showing a strong first-mention preference for the personal pronoun (95%) and a weaker second-mention preference for the d-pronoun (59%); unlike the Dutch learners of German, the German learners of Dutch showed target-like resolution preferences. This is striking since for both learner groups the resolution preferences for personal and d-pronouns in source and target languages pattern together.

This divergence between the performance of the learner groups could have resulted from proficiency differences: at lower proficiency levels (Dutch L2s) the learners do not discriminate between the functions of personal and d-pronouns. Rather, they have two pronominal forms for one function (to co-refer to the first-mentioned topical antecedent). As the learners become more proficient (German L2s) this initial strategy changes into a preference to resolve the d-pronoun towards the non-topic, and the learners become target-

like. This implies that the German L2 learners of Dutch were more proficient than the Dutch learners of German. As we did not directly measure the proficiency levels of the German learners, we cannot fully answer this question. But this would be unlikely given the fact that the German learners of Dutch had a length of exposure of 12 months only, while the Dutch learners had learned German for between three and five years. Moreover, the proficiency level of the Dutch learners of German as was tested by the German Placement Test attested a high command of German in general (B2). We need to leave this issue open, as we cannot draw any definite conclusions from our data. For future work it would be interesting to test German learners of Dutch at different proficiency levels.

There is another potential explanation for the observed L2 differences. Although German and Dutch native resolution preferences go into the same direction (personal pronouns prefer first-mentioned topical antecedents, while d-pronouns prefer second-mentioned non-topical antecedents), there were subtle cross-linguistic differences with regard to the timing of the effects. In German the personal pronoun was resolved relatively late compared to the d-pronoun and this finding was not observed for the Dutch natives. This observation is in line with the view that while the d-pronoun is marked for non-topical co-reference relations, the personal pronoun is un-marked, i.e. it is more neutral, and therefore more ambiguous. The German learners of Dutch might have resolved the d-pronoun in a target-like way because of the markedness-unmarkedness distinction for co-reference relations for the two pronouns in their L1. The German d-pronoun is highly marked for non-topicality compared to the relatively unmarked personal pronoun. The assumption that learners can be quite sensitive to form markedness is in line with what has been suggested in previous research (Roberts, et al., in prog.) and could be explained by *Learner Variety* differences (Klein & Perdue, 1997). In production, utterances of the *Basic Learner Variety* are determined by a set of cross-linguistic organizational principals, such as semantic and pragmatic principles. The way in which these principles interact may be guided by source language preferences. In our study, the German L2 learners of Dutch might have interpreted the d-pronoun as co-referring with the non-topical entity according to the strong L1 principle of form-markedness; the Dutch learners of German did not show this preference as it does not exist in their L1.

In general, the order of mention of the antecedent candidates influenced the resolution of personal and d-pronouns. This shows that even when syntactic subject-object distinctions are not available in the input, order of mention information is used to resolve the pronouns. Although it is very likely that the first-mention preference was due to topicality, and the preference for the second-mentioned entity occurred because it was non-topical, the relative influence of these two factors cannot be torn apart, since in the materials of the experiments

reported here the first-mentioned entity was always topical and the second-mentioned entity non-topical. Therefore, a second set of German-Dutch L2 experiments was run, in order to test (a) whether the resolution preferences of personal and d-pronouns varied when following non-canonical antecedent structures, and (b) whether the information status of the antecedent candidates had an influence on the order of mention preferences for personal and d-pronouns.

4.2. Pronoun Resolution after Non-Canonical Antecedent Structures

4.2.1. Introduction

In this section we investigate how Dutch learners of German and German learners of Dutch resolve personal and d-pronouns after non-canonical structures. In chapter 3, we reported the German and Dutch native resolution preferences after both types of antecedent structures, and found that after canonical topic-comment structures the two pronominal forms had different functions in that personal pronouns co-referred to the first-mentioned topical entity and d-pronouns to the second-mentioned non-topical entity. However, after a non-canonical topic-focus antecedent structure, the functions of the two pronominal forms overlapped; personal pronouns as well as d-pronouns were resolved towards the second-mentioned entity in focus. Focus was thus suggested to be an important cue in pronoun resolution. The learner question which is of interest is whether focus information is a comparably strong cue in L2 pronoun resolution. This question is particularly pertinent with regard to the Dutch L2 learners of German. In the previous section (chapter 4.1), the Dutch learners showed non-target-like resolution patterns in that they did not discriminate between the co-referential functions of personal and d-pronouns, but rather had one function for the two pronominal forms. Here we ask whether the learners are able to resolve personal and d-pronouns in a target-like way when they do not have to discriminate between their functions. Native-like L2 resolution would show a preference for the focused second-mentioned rather than the topical first-mentioned entity.

Previous visual-world studies on the resolution of personal and d-pronouns in a second language (Roberts, et al., in prog.; Wilson, 2009) tested the resolution preferences for the pronouns following canonical SVO and non-canonical OVS sentences (see example (70)). This design was chosen to disentangle grammatical role from order of mention influences on the resolution of the two pronominal forms. In this section I will discuss their findings with a main focus on the OVS conditions (see chapter 4.1 for SVO).

			Antecedent Structures	
		Pronouns	**SVO**	**OVS**
Roberts et al. (in progress)	Finnish L2s	**personal pronoun**	no preference	no preference
		d-pronoun	2nd	no preference
Wilson (2009)	English L2s	**personal pronoun**	1st	2nd → 1st
		d-pronoun	no preference	no preference

Table 4.12: Overview of results from previous visual-world studies on the resolution of personal and d-pronouns in L2 German

(70)

a. SVO antecedent structure (NOM-ACC)
Der Zauberer wollte **den Arzt** umarmen, weil die Sonne schien. Aber **er/der** war viel zu klein.
The magician$_{NOM}$ wanted the doctor$_{ACC}$ to-hug, because the sun$_{NOM}$ was shining. But he [P/D] was much too small.

b. OVS antecedent structure (ACC-NOM)
Den Arzt wollte **der Zauberer** umarmen, weil die Sonne schien. Aber **er/der** war viel zu klein.
The doctor$_{ACC}$ wanted the magician$_{NOM}$ to hug, because the sun$_{NOM}$ was-shining. But he [P/D] was much too small.

Example stimuli of Roberts et al. (in prog.)

Wilson (2009) tested English L2 learners of German and found an initial second-mention preference for the personal pronoun which turned into a first-mention preference over time. Together with the first-mentioned topic preference after SVO sentences, she takes this as evidence that the learners have a topic preference. But since no analysis over time windows is given, it is difficult to interpret this switch in preferences, since we do not know when it occurred (see chapter 3.1). For the d-pronoun, there was no preference irrespective of antecedent structure; when Wilson included proficiency as a factor into the analysis she found that "the higher a score participants obtained on the test, the more likely they were to interpret the demonstrative [d-pronoun] as the post-verbal NP (p.150-151)". It is surprising that the factor proficiency only had an effect on the OVS-d-pronoun and not on the SVO-d-pronoun. It indicates that the markedness of the antecedent structure had an effect on L2 pronoun resolution which pushed the learners towards a second-mentioned interpretation at higher proficiency levels, while in canonical topic-comment structures this was not the

case. The interaction is difficult to interpret without any separate analysis split by proficiency groups. Roberts et al. (in prog.) tested Finnish L2 learners of German in a similar set-up. They found no resolution preference for the personal pronoun after SVO and OVS antecedent structures. For the d-pronoun, they found a second-mentioned non-topical preference after SVO structures and no effect after OVS structures. The fact that no preference was found for the personal pronoun was taken as indication that the ambiguity could not be resolved due to the unmarkedness of the personal pronoun. After OVS sentences this preference did not emerge, because the non-canonical OVS structure had a weakening effect on the preference.

As Table 4.12 (gray column) shows, there are no preferences observed for the two pronominal forms (except for the unclear switch in preferences for the personal pronoun after OVS sentence structures in Wilson's (2009) study). This could be due to possible non-canonical word-order processing difficulties in the L2 learners. The OVS structures in German above might be particularly difficult for L2 learners, since learners cannot rely on positional cues to understand them, but need to analyze morphosyntactic cues such as case marking. This might be quite a challenge to the L2 processing system, above all when presented in spoken language and in real time (Hopp, 2007; p.195-254), particularly as L2 learners have been shown in a variety of acquisition areas to over-rely on non-structural cues at the expense of a detailed syntactic analysis (Clahsen & Felser, 2006a, 2006b). The fact that Wilson (2009) observed a different behavior concerning the OVS-d-pronoun depending on the proficiency level, might further indicate that more proficient L2 learners can better process this type of antecedent structure.

In sum, our understanding of pronoun resolution after non-canonical antecedent structures in L2 learners is very limited. The only two previous studies have reported no preferences for personal and d-pronouns after OVS sentences. Instead of indicating L2 pronoun resolution preferences, the results may reflect processing difficulties of OVS antecedent structures which require more processing effort than SVO structures in German due to complex morphosyntactic analyses such as case marking. In this study the role of word order on pronoun resolution preferences in L2 learners is examined, after inverted double nominative comparative structures which do not require such elaborate morphosyntactic analyses. It is hoped that we will be able to disentangle L2 processing difficulties which stem from the processing of difficult sentence structures from true resolution preferences. Additionally, as has been shown for the natives, the materials provide a context where the functions of the personal and the d-pronoun overlap, and we ask whether co-reference relations can be resolved target-like in such contexts where the information status of the referents may be used as a cue to disambiguation.

4.2.2. The present study

The present study examines L2 learners' resolution preferences for personal and d-pronouns after non-canonical antecedent structures. We will use the same antecedent sentences as in the experiments of chapter 4.1, but inverse them as to present structures of the type *Comparative-NP1-verb-NP2*. The advantage of this type of inversion compared to OVS structures is that the participants do not need to engage in heavy morphosyntactic analyses, but can still rely on pragmatic topic-focus cues to determine the relationship between the constituents of the sentence. The question of particular interest is whether the learners show the same resolution preferences as native German and Dutch speakers (chapter 3.2), i.e. whether the information structural cue *focus* is an important predictor for L2 resolution preferences.

The following research questions will be addressed:

1. How does the factor information structure of the antecedent clause influence L2 pronoun resolution?
2. Do L2 listeners show the same resolution preferences for personal and d-pronouns after non-canonical antecedent structures as native listeners? (And if not, in what ways do they differ?)
3. Do Dutch L2 learners of German and German L2 learners of Dutch resolve personal and d-pronouns after non-canonical antecedent structures in the same way?

4.2.3. Pronoun resolution in Dutch L2 learners of German

The first experiment investigates the resolution of personal and d-pronouns in Dutch learners of German.

4.2.3.1. Methods

Participants

Twenty-eight Dutch learners of German (23 female, 5 male) took part in the experiment[61]. They were paid for their participation and had normal or corrected-to-normal vision. The L2 learners took the same proficiency test as the L2 learners of chapter 4.1; the standardized German Placement Test developed by the Goethe Institute. They also filled out a language background questionnaire in which they provided information on their Dutch knowledge and use.

[61] Note that we tested 32 Dutch learners of German in chapter 4.1. Since the results of the native German participants in chapter 3.2 had shown that there was little variation in the resolution pattern after non-canonical antecedent structures, in this experiment we tested 28 L2-learners.

The participants had learned German at high school, on average between 3 and 5 years[62] and lived in the Netherlands at the time of testing. They were all students at the Radboud University Nijmegen and were aged between 18 and 32 years (mean = 23.11; SD = 3.36). They had started learning German at the age of 12 (mean = 12.43; SD = 1.95). The participants rated their German skills on a 5-point-scale for speaking, comprehension, writing, reading, grammar and pronunciation. Taking all of these scores together, the median for self-reported proficiency was 3, indicating a mediocre self-reported language level.

	Speaking	Comprehension	Writing	Reading	Grammar	Pronunciation
Median	4	3	4	3	4	3

Table 4.13: Medians of the self-reported level of German skills for the Dutch learners of German (1 = very good, 2 = good, 3 = mediocre, 4 = little knowledge, 5 = poor)

On average, the Dutch learner group scored 15.54 (SD = 3.64) on the German Placement Test, which had a maximum score of 30, i.e. on average they scored at the interface between level B1-B2, which means that they were advanced learners of German.

Experimental Tasks

The tasks were identical to those in the experiment with Dutch learners of German in chapter 4.1. The participants took part in a naming task which tested whether they were familiar with the lexical noun phrases which were used as antecedents in the material. The Dutch learners of German on average knew 45 of 48 (mean = 45.39; SD = 2.34) antecedent noun phrases which indicates that in general they knew the lexical noun phrases. However, in cases where they did not identify the lexical noun phrases correctly, the fixations for this item were removed from the analysis (347 fixations removed) resulting in a total of 4059 analyzable fixations.

Materials and design

The materials for the Dutch learners of German were the same as used with the native German speakers (chapter 3.2.3).

[62] The profile questionnaire was a web-based questionnaire, which asked the participants to indicate via a radio button box for how long they had been learning German. The following categories were possible answers: 0-1 year, 1-2 years, 2-3 years, 3-5 years, 5-10 years, more than 10 years. The median fell into the 3-5 years category.

(71) Conditions: Sample Item

a. Personal Pronoun Condition
Schwerer als der Tisch ist der Schrank. **Er** stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.
Heavier than the table is the cupboard. ***It*** *[P] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

b. D-pronoun Condition
Schwerer als der Tisch ist der Schrank. **Der** stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.
Heavier than the table is the cupboard. ***It*** *[D] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

Procedure

The procedure was identical to that of the German experiment with the native listeners (chapter 3.2.3). We used an SR Research EYELINK II eye tracker. The dominant eye was recorded. A sampling rate of 500-Hz was used which monitored gaze locations every 2 ms. The calibration of the camera which links the position of the eyes with a certain location on the screen, ensured that spatial accuracy was at least 0.5°.

The accuracy of the responses to the comprehension questions was very high with 92% correct answers (24 questions; mean correct answers = 22.00, SD = 1.78). This means that the participants were able to correctly understand the items.

4.2.3.2. Results

Forced Choice Questionnaire

Figure 4.5 and Table 4.14 show that the L2 learners of German chose the second-mentioned antecedent as being coreferential with the pronoun in 98% of all cases. Quite surprisingly, this was true for both types of pronouns.

A Chi-Square test was conducted on the data, which as expected showed no significant association between the type of pronoun and type of antecedent chosen, χ^2 (1) = .34, p=.386[63]. Thus, there was a second-mention preference across pronoun type. Two subsequent Chi Square tests revealed a highly significant second-mention preference (personal pronoun: χ^2 (1) = 200.28, p<.001; d-pronoun: χ^2 (1) = 206.82, p< .001).

[63] p>.05

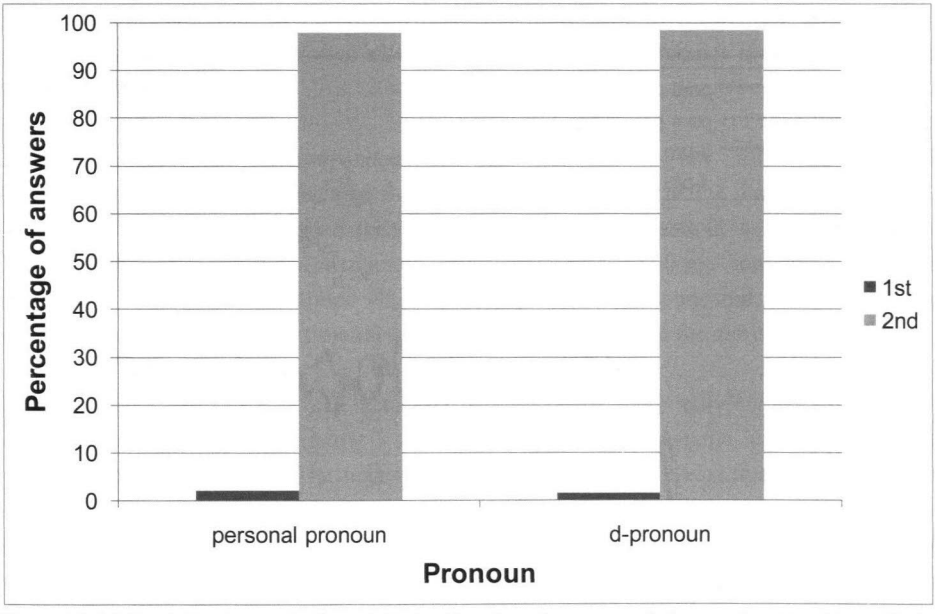

Figure 4.5: L2 Dutch Results on the Forced Choice Questionnaire; pronoun being coreferential either with the first- or the second-mentioned entity (1st, 2nd) in the two conditions (personal pronoun, d-pronoun)

	1st	2nd
personal pronoun	2.08% (7)	97.92% (329)
d-pronoun	1.49% (5)	98.51% (331)

Table 4.14: L2 Dutch Results on Forced Choice Questionnaire; pp = 28; items = 24

Visual-World Eye-Tracking

Data Analysis

The data analysis was the same as in the German native experiment (chapter 3.2.3). 24 looks (< 1%) to either target picture that started before pronoun onset (-200 till 0 ms) were excluded from the analysis, resulting in a total of 4035 fixations which entered the analysis.

Overview of the eye movement patterns

As can be seen in Figure 4.6, the amount of looks to all target pictures rises at 200 ms which is consistent with the finding that a saccade needs about 200 ms to be programmed.

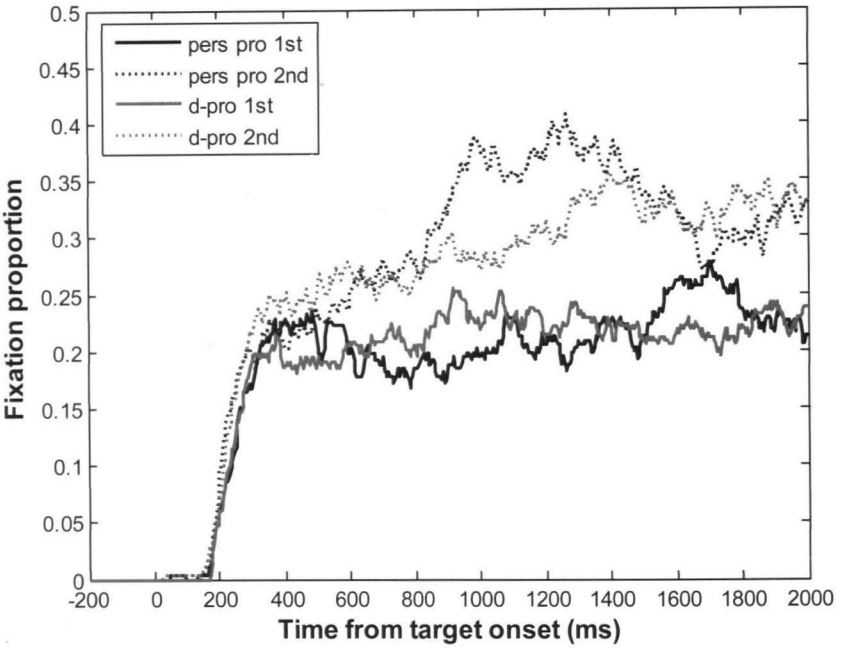

Figure 4.6: Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Personal pronoun. At about 500 ms after pronoun onset, the looks to the second-mentioned entity increase, with an even steeper rise from 800 ms onwards, while the looks to the first-mentioned entity stay at a 20%-level. At about 1200 ms after pronoun onset, the looks to the second-mentioned target reach a 40%-level, after which the looks to the second-mentioned entity start to decrease again until about 1700 ms. Until the end of the analysis window (2000 ms), the amount of looks to the second-mentioned entity rises again, and the amount of looks to the first-mentioned entity drops. Figure 4.6 thus shows a second-mention preference for the personal pronoun.

D-pronoun. As early as 400 ms after pronoun onset, there are more looks to the second-mentioned entity than to the first-mentioned entity for the d-pronoun. This difference remains, with an equal increase of looks to both targets until about 1200 ms, where we see a steeper increase in looks towards the second-mentioned entity, with a drop in looks to the first-mentioned entity, which persists until the end of the analysis period (2000 ms). So, this pattern of results also suggests a second-mention preference for the d-pronoun.

Taken together, the plot in Figure 4.6 reveals a preference to resolve the two types of pronouns to the second-mentioned antecedent, although the preference seems to be clearer for the personal pronoun.

Statistical Analysis of the eye movement patterns

The loglikelihood analysis revealed that the *interaction model* fit the data better than the *simple model* for time windows 5 and 6 (800 – 1200 ms) (see Table 0.14 in the Appendix). This corresponds to the above mentioned observations that at 800 ms there was a steep rise in looks to the second-mentioned entity for the personal pronoun, and at 1200 ms they decreased again. At the same time the second-mention preference for *der* became clearer.

Looking at the fixed factors (see Table 4.15), we found a marginally significant effect of order of mention in time window 3 (400 – 600 ms) which became significant in time window 4 (600 – 800 ms). The beta coefficient indicated that it was due to more overall looks to the second-mentioned entity. During time windows 5 and 6 (800 – 1200 ms), the interaction between type of pronoun and order of mention became significant, and in time window 6 there was a marginally significant second-mention main effect. In the last four time windows (1200 – 2000 ms), we observed again a significant main effect of order of mention with more looks to the second-mentioned target picture than the first.

Time window	in ms	Fixed predictors		
		pronoun condition	order of mention	interaction
1	0-200	-0.02 (-0.374)	0.08 (1.241)	
2	200-400	-0.05 (-0.251)	0.32 (1.682)	
3	400-600	0 (-0.02)	0.36 (1.668)†	
4	600-800	0.08 (0.383)	0.69 (3.14)**	
5	800-1000	-0.38 (-1.189)	0.5 (1.564)	0.96 (2.12)*
6	1000-1200	-0.21 (-0.639)	0.61 (1.88)†	1 (2.165)*
7	1200-1400	0.22 (0.932)	1.36 (5.842)***	
8	1400-1600	0.21 (0.903)	1.14 (4.955)***	
9	1600-1800	0.04 (0.156)	0.75 (3.264)***	
10	1800-2000	-0.11 (-0.458)	1.03 (4.427)***	

Table 4.15: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*er* vs. *der*) and order of mention (1st vs. 2nd). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

Individual analyses for each pronoun were conducted and order of mention was entered as a fixed factor (see Table 4.16). For the personal pronoun there was a significant main effect of order of mention between 600 and 1600 ms after pronoun onset and between 1800 and 2000 ms. The beta coefficient indicated more looks to the second-mentioned entity than the first.

For the d-pronoun, we observed an either marginally significant or significant main effect of mention for time windows 3 to 10 (400 – 2000 ms), except for time window 5 (800 - 1000 ms). As for the personal pronoun, there were more looks to the second-mentioned entity than to the first.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	0.07 (0.751)	0.09 (1.01)
2	200-400	0.2 (0.754)	0.43 (1.583)
3	400-600	0.05 (0.154)	0.67 (2.192)*
4	600-800	0.8 (2.577)*	0.58 (1.861)†
5	800-1000	1.46 (4.559)***	0.5 (1.555)
6	1000-1200	1.61 (4.911)***	0.61 (1.87)†
7	1200-1400	1.66 (4.987)***	1.07 (3.275)**
8	1400-1600	1.16 (3.568)***	1.12 (3.437)**
9	1600-1800	0.4 (1.184)	1.1 (3.452)***
10	1800-2000	0.95 (2.899)**	1.1 (3.343)**

Table 4.16: Results of the individual time course analyses for each type of pronoun (*er* and *der*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

4.2.3.3. Discussion

The on- and off-line results of the Dutch learners of German showed a target-like overall 3preference to resolve both personal and d-pronouns towards the second-mentioned entity. Contrary to the overall topic preference obtained after canonical antecedent structures (chapter 4.1), after non-canonical topic-focus antecedent structures, the learners had an overall preference for the entity in focus. In the eye movement results, this overall main effect of mention was marginally significant between 400 and 600 ms after pronoun onset, after which point it gained significance. This pattern indicates that the participants relied on the focus information provided by the antecedent sentence for future reference, as did the native German listeners (chapter 3.2.3).

In comparison to the previous L2 pronoun resolution studies (Roberts, et al., in prog.; Wilson, 2009) which mainly found no resolution preferences after non-canonical OVS antecedent structures the current findings indicate a clear sensitivity towards discourse-pragmatic information in L2 pronoun resolution. This difference is probably due to the fact that we used double nominative comparative structures instead of OVS antecedent structures. In chapter 3.2, the differences of our non-canonical pronoun resolution results in the natives compared to the native results of the above mentioned studies[64] were discussed. The differences may have been motivated by the difference in linguistic structure; that is comparative structures, especially marked ones, have a different information structure than OVS sentences and therefore influence pronoun resolution differently. However, second language learners face additional L2 difficulties when processing OVS-case marked sentence structures (Clahsen & Felser, 2006a, 2006b; Hopp, 2007). In the current L2 results obtained with non-case marked non-canonical antecedent structures, there was no hint at possible L2 difficulties in the processing of the non-canonical sentence structures. In comparison to the results of chapter 4.1.3, the opposite appears to be the case, that is, a facilitation of L2 pronoun resolution by the inverted non-canonical structures. The fact that inverted comparative structures are morphosyntactically less complex than OVS structures, shows that inverted comparative structures are a suitable alternative to study non-canonical processing in L2 learners, putting less load on their processing systems.

Nevertheless, the significant interaction between 800 and 1200 ms indicates that the pattern of results was different for personal and d-pronouns. The separate analysis for both pronouns showed that this interaction was due to the second-mention effect being highly significant for the personal pronoun, while for the d-pronoun it was temporarily not significant between 800 and 1000 ms and marginally significant between 1000 and 1200 ms, after which it gained significance again. One possible explanation is that there was temporal ambiguity for the d-pronoun which may be due to the German d-pronoun being homomorphous with the definite article. The ambiguity might have been motivated by the fact that the L2 learners expected a noun phrase to follow the d-pronoun, and it took processing effort and time to recompute when this expectation was not met. Alternatively, the variation for the d-pronoun as compared to the personal pronoun may be due to inter-individual processing differences due to the proficiency levels of the learners. This difference is not found in the off-line condition; when the learners have enough time, just like the natives they prefer to resolve the two pronouns towards the entity in focus.

This differs from what was observed in chapter 4.1.3. After canonical antecedent structures,

[64] Roberts et al.'s (in prog.) study on L2 pronoun resolution used the same materials as Schumacher et al.'s (in prog.) study on native resolution preferences in L1 German.

we found non-target-like resolution of the d-pronoun in the Dutch L2 learners of German. Why had the learners fewer problems after non-canonical marked antecedent structures? In contrast to the canonical structures, in non-canonical contexts the function of the personal and the d-pronoun overlap: the two pronouns both preferred the second-mentioned focused entity. In canonical contexts, however, the two pronominal forms have two different functions (as shown by the native results in chapter 3.1); the personal pronoun co-referred to the first-mentioned topical entity, while the d-pronoun co-referred towards the second-mentioned non-topical entity. Taken together, L2 learners have no difficulties resolving personal and d-pronouns in a target-like way when they have overlapping functions in discourse. Only when a discrimination of these functions is required, L2 learners behave non-target-like. This indicates that they have not learned to discriminate between both functions yet; however, they are able to use discourse-pragmatic cues to resolution, i.e. they are able to switch from their topical first-mention preference to a focused second-mention preference when the information structural context requires so.

In future research, it would be interesting to look at different alternatives of marking information structure in antecedent sentences (particularly focus), to understand which cues L2 learners rely on when resolving pronouns and processing discourse, and to get a better understanding of the importance of information structure in L2 pronoun resolution. For example, intonation would be an interesting factor to study, since canonical sentence structures could be kept, while marking different pragmatic antecedent functions, such as topic and focus and comparing them directly to a focus-unmarked (topic-comment) condition.

4.2.4. Pronoun resolution in German L2 learners of Dutch

4.2.4.1. Methods

Participants

Thirty-two German learners of Dutch (20 female, 3 male) took part in the experiment. The participants were aged between 19 and 28 years (mean = 23.04; SD = 2.39). They were paid for their participation. All participants had normal or corrected-to-normal vision.

The learners were students at the Radboud University Nijmegen and lived in the Netherlands at the time of testing. We tested undergraduate as well as graduate students (median length of exposure: 14 months; ranging from 1 to 50 months). All of them had learned Dutch during an intensive language course at the Radboud University Nijmegen (mean age of onset = 20.22; SD = 2.28) at the end of which they had all passed the NT2-

exam (*Dutch as a second language*) attesting them a level of high proficiency in Dutch. The German learners of Dutch self-reported a good level of Dutch, as indicated by the general median of 2 (see Table 4.17).

	Speaking	Comprehension	Writing	Reading	Grammar	Pronunciation
Median	2	2	2	2	2	2

Table 4.17: Medians of the self-reported level of Dutch skills for the German learners of Dutch (1 = very good, 2 = good, 3 = mediocre, 4 = little knowledge, 5 = poor)

Experimental Tasks

The tasks were identical to those in the native experiment. After the experiment, the German learners of Dutch were asked to perform a naming task on the computer screen. The German learners of Dutch knew 47 of the 48 experimental antecedent noun phrases (mean = 47.21; SD= 0.88) which showed in general a good understanding of the lexical NPs used in the materials. However, 101 fixations during items for which at least one of the antecedents was not correctly identified during the naming task were removed from the eye-tracking analysis, resulting in a total of 3729 fixations to be analyzed. At the end of the session a language background questionnaire was completed. The whole session took about 60 min.

Materials and design

The materials for the German learners of Dutch were the same as used with the native Dutch speakers (chapter 3.2.3) repeated here for convenience.

(72) Conditions: Sample Item

a. Personal Pronoun Condition
Zwaarder dan de tafel is de kast. **Hij** is afkomstig uit een meubelwinkel in België. De sofa zal volgende week geleverd worden.
Heavier than the table is the cupboard. ***It*** *[P] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

b. D-pronoun Condition
Zwaarder dan de tafel is de kast. **Die** is afkomstig uit een meubelwinkel in België. De sofa zal volgende week geleverd worden.
Heavier than the table is the cupboard. ***It*** *[P] originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.*

Procedure

The procedure was identical to that of the Dutch experiment with the native listeners. We used the same eye tracker as in the experiment with the Dutch listeners, an SR Research EYELINK II eye tracker. The dominant eye was recorded. A sampling rate of at least 500-Hz was used which monitored gaze locations every 2 ms. The calibration of the camera which links the position of the eyes with a certain location on the screen, ensured that spatial accuracy was at least 0.5°.

The accuracy of the responses to the comprehension questions was very high with 94% correct answers (24 questions; mean correct answers = 22.54, SD =1.22). This percentage of correct responses was comparable to the Dutch learners of German (92%).

4.2.4.2. Results

Forced Choice Questionnaire

Due to a ceiling effect, two cells (personal pronoun –1st, d-pronoun – 1st) had an expected count less than 5 (namely 1 and 3). Therefore, Fisher's exact test was conducted on the data, and as expected showed no significant association between the type of pronoun and the type of antecedent chosen, P(Fisher, 2-sided) = .312[65]. There was an overall second-mention preference. Two subsequent Chi Square revealed that the second-mention preference was highly significant (personal pronoun: χ^2 (1) = 181.59, p<.001; d-pronoun: χ^2 (1) = 188.47, p<.001).

[65] P(Fisher, 2-sided)>.05

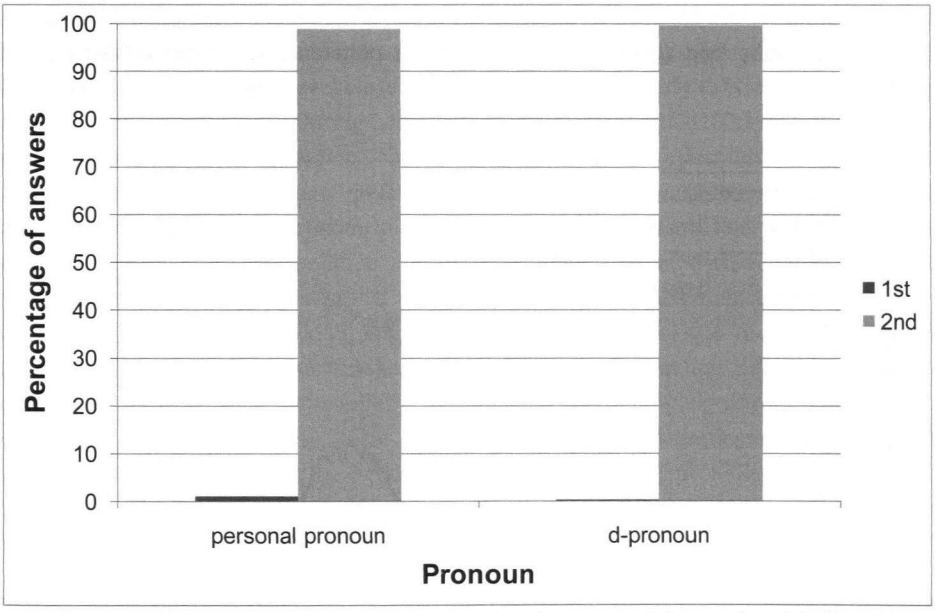

Figure 4.7: L1 Dutch Results on the Forced Choice Questionnaire; pronoun being coreferential either with the first- or the second-mentioned entity (1st, 2nd) in the two conditions (personal pronoun, d-pronoun)

	1st	2nd
personal pronoun	1.04% (3)	98.96% (285)
d-pronoun	0.35% (1)	99.65% (287)

Table 4.18: L1 Dutch Results on Forced Choice Questionnaire; pp = 24; items = 24

Visual-World Eye-Tracking

Data Analysis

The data analysis was as in the German native experiment. 23 looks (< 1%) to either target picture that started before pronoun onset (-200 till 0 ms) were excluded from the analysis, 7resulting in a total of 3706 fixations which entered the analysis.

Overview of the eye movement patterns

In Figure 4.8, we see a rise in looks to all target pictures at around 180 to 200 ms which is the time needed to plan an eye movement. After this there is an increase of target looks to the second-mentioned entity for the two types of pronouns, while the looks to the first-mentioned entity decrease. At 1200 ms and onwards, there are more looks to the first-mentioned target for the d-pronoun as compared to the personal pronoun. Nevertheless,

they do not rise above the 25%-level, while the looks to the second-mentioned entity are on a 35%-level, indicating an overall second-mention preference regardless of the type of pronoun.

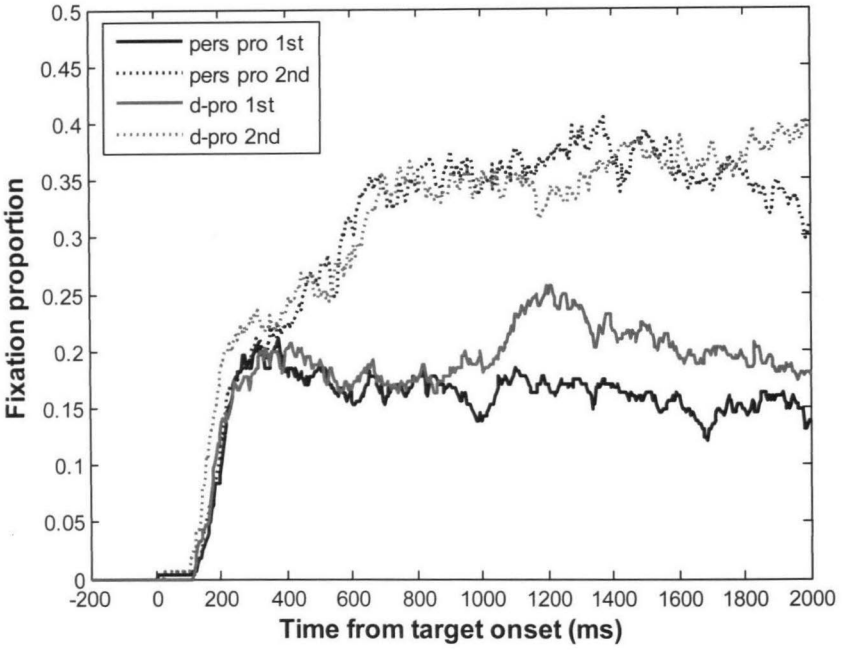

Figure 4.8: Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Statistical Analysis of the eye movement patterns

The loglikelihood test revealed that the *interaction model* which contained the *condition x mention* interaction term fit the data better during time window 7 (1200 – 1400 ms) than the *simple model* containing the main effect terms only (see Table 0.15 in the Appendix). This indicated an interaction which may be due to the above reported temporal increase in first-mentioned target looks for the d-pronoun in this time window, which nevertheless do not rise above a 25%-level.

The model estimates for the fixed effects and their significance levels (see Table 4.19) showed a significant main effect of condition and a marginally significant main effect of mention during the first time window (0 – 200 ms). There were more looks for the d-pronoun than for the personal pronoun, and an overall preference to look at the second-

mentioned entity. This was due to a slightly faster rise of looks to the second-mentioned entity for the d-pronoun. This slightly sooner rise had a relatively high impact on the data, since the looks arose from a level close to zero. They were too early to be attributable to the resolution of the pronouns. During time windows 3 to 10 (400 – 2000 ms), there was a significant main effect of mention. The beta coefficient of the effect indicated that there was an overall second-mention preference. In time window 7 (1200 – 1400 ms), the interaction between type of pronoun and order of mention was significant. We also found a significant main effect of condition due to more looks for the d-pronoun than for the personal pronoun. This indicated that the pattern of results differed among the pronouns, which was probably due to a slightly increased amount of looks to the first-mentioned entity for the d-pronoun for this time window. However, both pronouns still show a second-mention preference, since the main effect of mention was highly significant. During time window 9 (1600 – 1800 ms), there was also a marginally significant main effect of condition with more looks for the d-pronoun than for the personal pronoun. Taken together, the analysis showed a highly significant second-mention effect from 400 ms after pronoun onset onwards till the end of the analysis window for the two pronouns.

Time window	**in ms**	**Fixed predictors**		
		pronoun condition	order of mention	interaction
1	0-200	-0.19 (-2.249)*	0.16 (1.937)†	
2	200-400	-0.17 (-0.806)	0.26 (1.229)	
3	400-600	-0.04 (-0.199)	0.9 (4.091)***	
4	600-800	-0.01 (-0.028)	1.61 (6.965)***	
5	800-1000	0.04 (0.195)	1.77 (7.693)***	
6	1000-1200	-0.17 (-0.702)	1.58 (6.654)***	
7	1200-1400	-0.72 (-2.146)*	1.11 (3.3)**	0.94 (1.99)*
8	1400-1600	-0.14 (-0.598)	1.86 (7.938)***	
9	1600-1800	-0.38 (-1.636)†	1.84 (8.029)***	
10	1800-2000	-0.39 (-1.685)	1.85 (7.957)***	

Table 4.19: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*hij* vs. *die*) and order of mention (1st vs. 2nd **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † p<.1; * p<.05; ** p<.01; *** p<.001.

Individual analyses were conducted on each pronoun, and mention was entered as a fixed effect into the model (see Table 4.20). For the personal pronoun, the analysis revealed a highly significant second-mention effect between 400 and 2000 ms.

For the d-pronoun, we observed a significant main effect of mention already during the first time window (0 – 200 ms), which was due to a slightly steeper rise in looks to the second-mentioned entity than to the first entity in this time window. However, since all looks started rising from a level slightly above zero %, minimal differences had a maximal effect. Thus, this effect was most probably not due to pronoun resolution preferences. However, as for the personal pronoun, for the d-pronoun the second-mention effect becomes highly significant between 400 and 2000 ms.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	0.05 (0.451)	0.27 (2.16)*
2	200-400	0.12 (0.417)	0.39 (1.286)
3	400-600	0.96 (3.1)**	0.84 (2.686)**
4	600-800	1.67 (5.097)***	1.54 (4.703)***
5	800-1000	1.95 (6.01)***	1.59 (4.821)***
6	1000-1200	1.92 (5.758)***	1.24 (3.645)***
7	1200-1400	2.05 (6.172)***	1.11 (3.248)**
8	1400-1600	2.21 (6.824)***	1.51 (4.447)***
9	1600-1800	2.04 (6.357)***	1.65 (5.014)***
10	1800-2000	1.75 (5.417)***	1.95 (5.774)***

Table 4.20: Results of the individual time course analyses for each type of pronoun (*hij* vs. *die*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Thus, the individual analyses showed that there was an overall second-mention preference for the two types of pronouns starting at 400 ms and lasting till the end of the analysis window (2000 ms).

4.2.4.3. Discussion

The eye movement data showed an overall preference to resolve the two types of pronouns towards the second-mentioned entity, with the effect emerging as early as 400 ms after pronoun onset. The off-line task revealed the same pattern of results with an overall second-mention preference at ceiling (99 - 100%). The learners were able to use the information structural information, focus, to resolve the two pronouns in a target-like way. Furthermore, there was no hint at L2 processing difficulties for the non-canonical antecedent structures. The learners showed native-like resolution preferences for the two pronouns.

Since L2 pronoun resolution after non-canonical antecedent structures has never been studied in Dutch, it was particularly useful to use inverted double nominative comparative structures which enabled us to compare the results to the Dutch L2 learner patterns in German since these structures present the antecedents in the nominative in both languages.

In general, L2 learners have no difficulties resolving personal and d-pronouns in a native-like way, when the two pronouns have overlapping functions. Difficulties may arise when the pronouns have different functions. In the case of the German L2 learners of Dutch, we even observed target-like resolution preferences for the two pronouns when their functions needed to be discriminated (see chapter 4.1). Whether this is due to proficiency differences, or whether this is due to differences in the learner varieties cannot be resolved here. Future work should therefore test German L2 learners of Dutch at different proficiency levels, and learners from typologically more different languages to address this question.

In the next section, the question of whether the semantic factor animacy had an influence on the results is examined. The data from the experiments reported in chapter 4.1 is used to address this question, as there was more variability than in the data of this section. As the Dutch L2 learners showed non-target-like resolution preferences in chapter 4.1.3, it is generally assumed that they will also show non-target-like behavior with regard to the factor animacy. However, it will be particularly interesting to see the results for the German learners of Dutch who showed target-like resolution preferences in chapter 4.1.4, because the two native groups differed with regard to the influence of animacy.

4.3. The Influence of Animacy on L2 Pronoun Resolution

4.3.1. Background

German and Dutch natives differed in how animacy affected pronoun resolution (chapter 3.3). In German, the first-mention effect for personal pronouns was considerably earlier for animate than for inanimate antecedents. In Dutch, there was an early first-mention effect after inanimate antecedents for both pronominal conditions which emerged very early (at 200 ms); this effect was thus not attributed to pronoun resolution but rather to a general discourse expectation to continue with the first-mentioned entity. The discourse expectation facilitated the resolution of the personal pronoun which met the expectation to continue the discourse with the first-mentioned entity, but not for the d-pronoun which was later resolved to the second-mentioned entity (1200 ms).

	hij	*die*
animate	1st at 600 ms	2nd at 1000 ms
inanimate	1st at 200 ms 1st at 400 ms	1st at 200 ms 2nd at 1200 ms

	er	*der*
animate	1st at 600 ms	2nd at 1000 ms
inanimate	1st at 1400 ms	2nd at 600 ms

Table 4.21: Direction and emergence of effects of the personal and the d-pronoun for animate and inanimate antecedents in German (above) and Dutch (below)

For the L2 learners, with regard to the results reported in chapter 4.1.4, it was predicted that if the native-like behavior of the German learners of Dutch was due to their high proficiency level then they should be as sensitive towards animacy information as the Dutch native speakers. On the other hand, for the Dutch L2 learners of German who showed non-target-like resolution behavior in chapter 4.1.3, it was assumed that their resolution preferences might be affected differently by animacy information than those of the German native speakers. If however, the native-like resolution pattern found for the German learners of Dutch in comparison to the Dutch non-target-likeness is due to the fact that the Germans were able to discriminate the d-pronoun function from the personal pronoun function because in their native language the d-pronoun is marked for non-topicality, then they should show source language-like sensitivity towards animacy information.

Alternatively, animacy could have a different effect on L2 pronoun resolution than on L1

pronoun resolution. The second language acquisition literature has reported an influence of animacy on different aspects of acquisition, such as on thematic role assignments (Gass, 1987; Kempe & MacWhinney, 1998; Kilborn, 1989; LoCoco, 1987; McDonald, 1987; Sasaki, 1994) or relative clause attachment preferences (Jackson, 2008; Jackson & Roberts, accepted). This sort of influence would thus be due to a general L2 sensitivity towards this factor and should therefore be observable in both learner groups.

More specifically, some studies have found that less proficient L2 learners mainly rely on the highly salient cues of their L1 when interpreting agenthood in the L2 (LoCoco, 1987 for English L2 learners of German, word order versus case marking strategies). But studies using different language pairings have found that animacy is in fact a strong predictor for L2 learners, irrespective of its informational validity in their L1 (Gass, 1987; Sasaki, 1994). The evidence suggests that as the proficiency level increases, learners become more sensitive to the salient L2 cues (Kilborn, 1989; McDonald, 1987; Sasaki, 1994).

In a self-paced reading study, Jackson and Roberts (accepted) tested German L2 learners of Dutch and showed an effect of animacy on the L2 comprehension of relative clauses. They presented their materials in a 2x2-design (see example (73)) investigating the influence of clause type (Subject RC vs. Object RC) and animacy (animate subject vs. inanimate subject) on relative clause comprehension. All items were mixed animacy items, thus either appearing in AI or IA conditions. The auxiliary (*heeft*$_{SG}$ vs. *hebben*$_{PL}$) followed both antecedents (*de clown*$_{SG}$, *de taarten*$_{PL}$) and disambiguated the sentences.

(73) Experimental conditions (Jackson & Roberts, accepted)

Voor de kinderen ...
For the children...
"For the children..."

a.	... is de clown, die de taarten	heeft	gegooid, ...	(SubjRC, animate)
	...is the clown$_{SG}$ *that the pies*$_{Pl}$	*has*$_{SG}$	*thrown ...*	
	"...the clown, that threw the pies, ..."			
b.	... zijn de taarten, die de clown	heeft	gegooid, ...	(ObjRC, animate)
	...are the pies$_{PL}$ *that the clown*$_{SG}$	*has*$_{SG}$	*thrown ...*	
	"...the pies, that the clown threw, ..."			
c.	... zijn de taarten, die de clown	hebben	geraakt, ...	(SubjRC, inanimate)
	...are the pies$_{PL}$ *that the clown*$_{SG}$	*have*$_{Pl}$	*hit ...*	
	"...the pies, that hit the clown, ..."			

d. ... is de clown, die de taarten hebben geraakt, ... (ObjRC, inanimate)
...is the clown$_{SG}$ that the pies$_{PL}$ have$_{PL}$ hit ...
"...the clown, that the pies hit, ..."

... de hoogtepunt van de voorstelling.
... the highlight of the performance.
"...was the highlight of the performance."

Jackson and Roberts found that German learners of Dutch behaved like native Dutch speakers, in that no reading time differences were observed between subject and object relative clauses when the subject was animate and the object was inanimate. They also showed that the reading time patterns differed between L1s and L2s prior to the disambiguating region (*heeft$_{SG}$* vs. *hebben$_{PL}$*), namely on the ambiguous NPs (e.g.: ... *de taarten, die de clown* ...). While the native speakers showed no reading time differences in this region, the L2 learners showed longer processing times for inanimate antecedent nouns (e.g.: ... ***de taarten***, *die de clown* ...) than for animate antecedent nouns (e.g.: ... ***de clown***, *die de taarten* ...). No difference was found for the RC-internal nouns (e.g.: ... *de taarten, die* ***de clown*** ...). This observation indicated an interaction between topichood and animacy when assigning grammatical roles. The results suggest that the factor animacy may play an important role in L2 processing in general, and it is very likely that this influence may also be present in L2 pronoun resolution.

The research questions are as follows:

1. Does the factor animacy influence L2 pronoun resolution?
2. Does animacy information influence L2 pronoun resolution in a comparable way to L1 pronoun resolution?
3. Is L2 pronoun resolution in general affected differently by animacy information of the antecedents than L1 pronoun resolution? (The effect should be observable in both learner groups.)

4.3.2. Results: Dutch L2 learners of German

The data presented in chapter 4.1.3 were reanalyzed in the same way as the native data in chapter 3.3. In a first step, we tested with loglikelihood tests whether the model containing the factor animacy, the *animacy model*, could better explain the variance in the data. Thus, the goodness-of-fit of the *animacy model* which contained the *condition x mention x animacy* three-way interaction term, was evaluated against the *interaction model*, containing the *condition x mention* interaction, and the *simple model* containing only the main effects (see Table 0.16 in the Appendix).

Time window	in ms	Fixed predictors						
		main effects			2-way interactions			3-way interaction
		pronoun condition	order of mention	animacy	condition x mention	condition x animacy	mention x animacy	condition x mention x animacy
1	0-200	-0.02 (-0.434)	0.04 (0.805)					
2	200-400	-0.02 (-0.094)	-0.08 (-0.454)					
3	400-600	0.08 (0.415)	-0.26 (-1.322)					
4	600-800	0.47 (1.64)†	0.03 (0.116)		-0.72 (-1.754)†			
5	800-1000	0.05 (0.224)	-0.63 (-2.987)**					
6	1000-1200	0.97 (2.314)*	0.07 (0.17)	0.86 (2.028)†	-1.92 (-3.238)**	-1.13 (-1.865)†	-1.38 (-2.303)*	2.36 (2.753)**
7	1200-1400	0.71 (2.299)*	-0.54 (-1.767)†		-0.96 (-2.214)*			
8	1400-1600	0.82 (1.95)†	-0.42 (-0.999)	-0.38 (-0.884)	-1.76 (-2.956)**	-0.06 (-0.103)	0.09 (0.157)	1.49 (1.729)†
9	1600-1800	0.77 (1.851)†	0.77 (1.851)†	-1.65 (-2.802)	-0.33 (-0.793)**	-0.31 (-0.743)**	-0.32 (-0.525)	-0.01 (-0.013)
10	1800-2000	0.21 (1.009)	-0.82 (-3.922)***					

Table 4.22: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*er* vs. *der*), order of mention (1st vs. 2nd) and animacy (animate vs. inanimate). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

The analysis revealed that the *animacy model* explained the data marginally better than the *interaction model* in time window 6, 8 and 9 (1000 – 1200 ms; 1400 – 1800 ms). This means that *animacy* may have had an influence on learners' pronoun resolution.

Turning to the fixed effects (see Table 4.22), in time window 6 (1000 – 1200 ms) a highly significant three-way interaction was observed, and all of the two-way interactions were either marginally significant or significant. Interestingly, this suggests that the *condition x mention* interaction was resolved differently for animate and for inanimate items: animacy might have had an influence on the direction of the preferences. There was a significant main effect of condition and a marginally significant main effect of animacy. The positivity of the beta coefficients indicated that there were overall more looks for the personal pronoun as compared to the d-pronoun and more looks for inanimate items than for animate items. In windows 8 and 9 (1400 – 1800 ms), the three-way interaction was marginally significant, and the two-way interaction *condition x mention* was significant. The main effect of condition was marginally significant in both time windows, and the main effect of mention was marginally significant in time window 9 indicating that there were overall more looks to the second-mentioned entity. To further evaluate the effects found for the *animacy model*, separate analyses will be provided for animate and inanimate items.

4.3.2.1. Animate Items

Figure 4.9 shows the resolution of personal and d-pronouns separately for the animate items (2515 looks; 53%) by the Dutch learners of German. The resolution pattern for the animate items is different from the overall resolution pattern (see chapter 4.1.3), in that the d-pronoun only displays a very weak first-mention tendency which cannot be determined as a significant preference on the basis of the figure.

	er	*der*
animate & inanimate	1st at 600 ms	1st at 1000 ms

Table 4.23: Direction and emergence of effects of the personal and the d-pronoun over animate and inanimate items (see chapter 4.1.1)

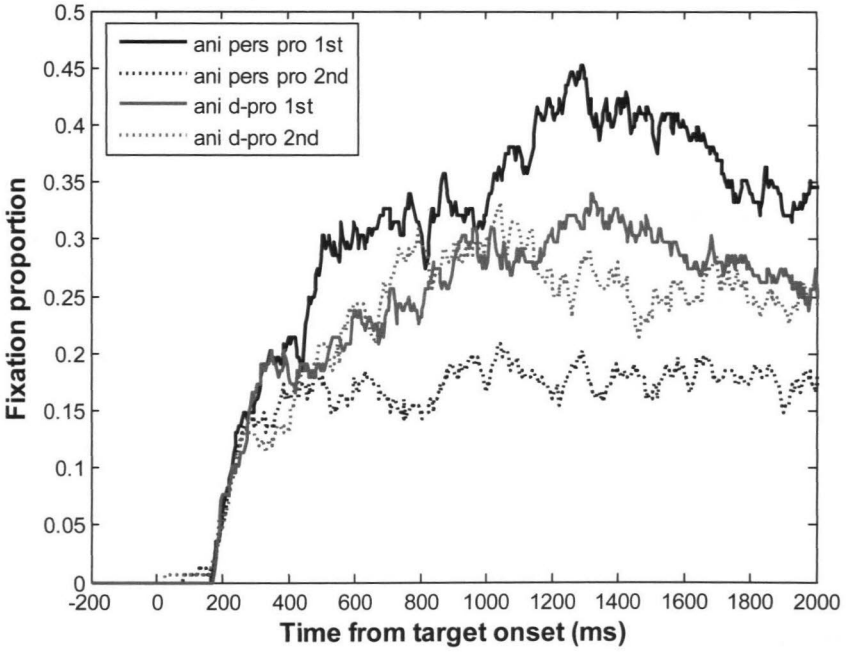

Figure 4.9: Animate items. Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Individual comparisons were conducted for the two types of pronouns with order of mention as a predictor (see Table 4.24). For the personal pronoun, the analysis revealed a significant main effect of order of mention for time windows 3 to 10 (400 – 2000 ms). The negativity of the beta coefficient for mention indicated a first-mention preference. The fact that the intercept was not significant in windows 7 and 8 (1200 – 1600 ms) indicated that there was variation in the data which was probably due to the fact that there is more variability in L2 learner data in general. Therefore, when the analyses were run only on half of the items, this variability had a greater impact. However, since the first-mention effect for the personal pronoun was very clear and the intercept term was significant in all of the other time windows, this arguably does not violate the conclusions.

The individual analysis revealed no preference for the d-pronoun. This is interesting when compared to the overall results from chapter 4.1 where the d-pronoun showed a significant first-mention preference. It suggests that the learners had a different preference between animate and inanimate items. This is analyzed in the next section.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	0 (0.018)	0.04 (0.322)
2	200-400	-0.25 (-0.742)	-0.31 (-0.934)
3	400-600	-0.95 (-2.398)*	-0.28 (-0.743)
4	600-800	-1.34 (-3.37)**	0.12 (0.295)
5	800-1000	-1.36 (-3.347)***	-0.05 (-0.115)
6	1000-1200	-1.85 (-4.431)***	0.07 (0.169)
7	1200-1400	-2.17 (-5.122)***	-0.38 (-0.887)
8	1400-1600	-2.18 (-5.269)***	-0.42 (-0.986)
9	1600-1800	-1.98 (-4.808)***	-0.33 (-0.776)
10	1800-2000	-1.69 (-4.169)***	-0.49 (-1.184)

Table 4.24: Results of the individual time course analyses for each type of pronoun (*er* and *der*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Summing up, there was a first-mention effect for the personal pronoun and surprisingly no effect for the d-pronoun. The first-mention preference for the personal pronoun emerged quite early at 400 ms and lasted until the end of the analysis window (2000 ms). Thus, concerning the time windows which had previously been identified to be sensitive towards the factor animacy (loglikelihood analysis), the personal pronoun showed a first-mention effect in all three time windows (1000 – 1200 ms, 1400 – 1800 ms), and the d-pronoun showed no effect. This non-preference for the d-pronoun sharply contrasts with the resolution pattern which was found over all items and in which *der* was preferentially resolved towards the first-mentioned entity. This first-mention preference was overridden by animacy in the learner data. Thus, in comparison to animacy effects on the timing of L1 pronoun resolution (chapter 3.3) in L2 pronoun resolution, animacy switches the direction of the preferences. The following section investigates how both pronouns were resolved following two inanimate antecedents.

4.3.2.2. Inanimate Items

The same type of analysis was conducted on the inanimate items. Figure 4.10 shows the resolution of personal and d-pronouns separately for the inanimate items (2214 looks; 47%) in the Dutch learners of German. Strikingly, the d-pronoun seems to show a clear first-mention preference for the inanimate items.

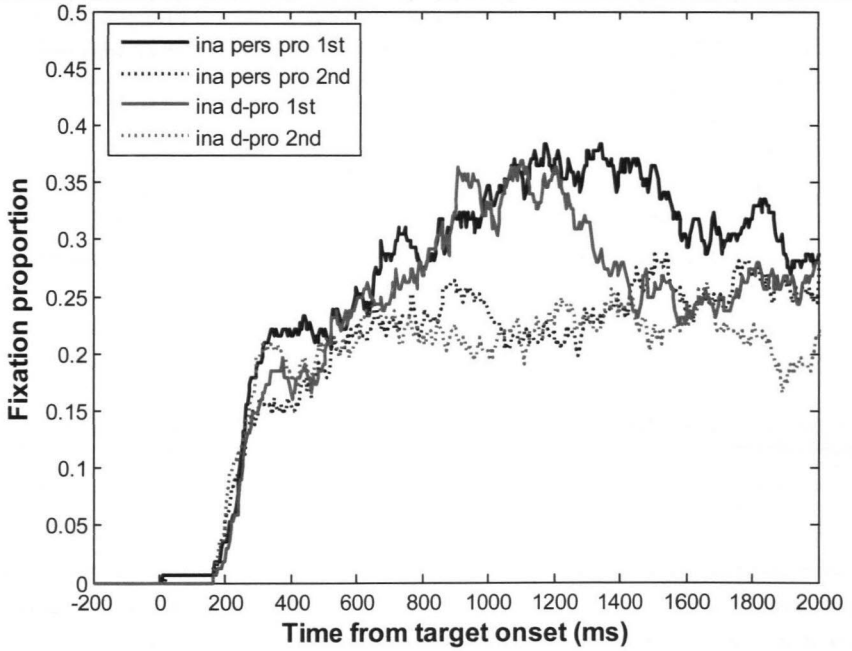

Figure 4.10: Inanimate items. Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

The individual analyses for the two pronouns revealed that following two inanimate antecedents, the personal pronoun was preferentially resolved towards the first-mentioned antecedent (see Table 4.25). This was indicated by the significant main effect of order of mention in time window 6 (1000 – 1200 ms) which was marginally significant in time window 7 (1200 – 1400 ms).

For the d-pronoun, there was a marginally significant main effect of order of mention in time window 5 (800 – 1000 ms) which became significant in window 6 (1000 – 1200 ms) and again marginal in window 7 (1200 – 1400 ms). As for the personal pronoun, the negativity of the beta coefficient reflected a first-mention preference for the d-pronoun.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	0.01 (0.126)	0.12 (1.374)
2	200-400	-0.24 (-0.71)	0.5 (1.482)
3	400-600	-0.03 (-0.069)	0.27 (0.702)
4	600-800	0.06 (0.128)	-0.06 (-0.134)
5	800-1000	-0.38 (-0.855)	-0.71 (-1.678)†
6	1000-1200	-0.87 (-1.92)*	-1.31 (-3.089)**
7	1200-1400	-0.75 (-1.622)†	-0.7 (-1.662)†
8	1400-1600	-0.6 (-1.278)	-0.33 (-0.776)
9	1600-1800	-0.38 (-0.837)	-0.34 (-0.816)
10	1800-2000	-0.53 (-1.174)	-0.54 (-1.317)

Table 4.25: Results of the individual time course analyses for each type of pronoun (*er* and *der*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

For inanimate items we found a first-mention preference for both types of pronouns which emerged roughly around the same time (1000 ms). The personal pronoun was thus resolved later for inanimate than for animate items (400 ms). Concerning the time windows which had been shown to be affected by the factor animacy, namely windows 6, 8 and 9, we observed a first-mention effect for the two pronouns during window 6, and no effects otherwise. This contrasts with the observation for animate items where a significant effect for the personal pronoun was found during all three time windows and no effect for the d-pronoun.

4.3.3. Discussion

The analysis of the factor animacy revealed an influence on the L2 pronoun resolution by Dutch learners of German. Interestingly, for the d-pronoun there was a first-mention preference following two inanimate NPs, and no preference after two animate NPs. This result is highly interesting as it answers some of the questions raised in chapter 4.1.3, where L2 Dutch learners of German were shown to have a non-target-like first-mention preference for the d-pronoun over all items. It was suggested that the learners had two forms but only one function, namely a topicality/first-mention reference function. This explanation cannot account for the no-preference pattern after animate items, however. The fact that the d-pronoun was treated differently across animate and inanimate items shows that the learners distinguished at least to some degree between the functions of personal and d-pronouns. Thus, the explanation needs to be extended. The learners may know that the d-pronoun

refers to the second-mentioned entity, as shown by the distribution after animate items. But if there are two inanimates, the discourse expectations might make the second-mentioned antecedent less available. It is perhaps more likely that the discourse will continue with an animate antecedent; the second-mentioned antecedent qualifies well as a topic of the subsequent utterance. In the case of inanimates this choice is more ambiguous since neither antecedents fit well with the expectation to become discourse topics. In fact, the second-mentioned antecedent is even more backgrounded; thus, it becomes less available, and the topicality default is applied. This explanation is supported by the timing differences of the first-mention preference for personal pronouns. They were resolved very early after animate antecedents (400 ms), but remained ambiguous for a longer period following two inanimate antecedents (1000 ms). This also indicates that there was a preference for animate antecedents to continue the discourse.

		er	*der*
L1 German	**animate**	1st at 600 ms	2nd at 1000 ms
	inanimate	1st at 1400 ms	2nd at 600 ms
L2 Dutch learners of German	**animate**	1st at 400 ms	-
	inanimate	1st at 1000 ms	1st at 800 ms

		hij	*die*
L1 Dutch	**animate**	1st at 600 ms	2nd at 1000 ms
	inanimate	1st at 200 ms 1st at 400 ms	1st at 200 ms 2nd at 1200 ms

Table 4.26: Direction and emergence of effects of the personal and the d-pronoun for animate and inanimate antecedents in L1 Germans (top), Dutch learners of German (center), and L1 Dutch (below)

It is not clear what underlies the no-preference pattern for d-pronouns after animate antecedents. In particular, the question remains as to whether the pattern was due to the participants switching their looks between the two potential antecedents, as a characteristic gaze behavior during ambiguity, or whether it was due to inter-individual differences. In other words, some of the participants may have resolved the pronoun towards the first-mentioned entity, while others resolved it towards the second. A possible inter-individual factor might be proficiency. As we do not have enough data to include proficiency in the *animacy model*, this issue cannot be resolved here. However, it may be an interesting starting point for future work.

4.3.4. Results: German L2 learners of Dutch

As for the Dutch learners of German, we conducted loglikelihood tests on the model fit to the data with animacy included as a factor and compared it to the models where animacy was not included. The *animacy model* explained the data marginally better than the *simple model* in time window 3 (400 – 600 ms) and significantly better than the *interaction model* in windows 4, 5 and 6 (600 – 1200 ms) (see Table 0.17 in the Appendix). Therefore, we analyzed the fixed effects of the *animacy model* for time windows 3, 4, 5 and 6.

The analysis of the fixed effects (see Table 4.27) revealed a significant three-way *condition x mention x animacy* interaction in all four time windows with all of the two-way interactions being significant (and the *condition x animacy* interaction being marginally significant in window 6). The two main effects of animacy and condition were always significant. The main effect of mention was marginally significant in windows 4 and 6, and significant in window 5. This indicates a strong influence of animacy on the resolution patterns of the pronouns, and might reflect a change in preferences due to animacy. The fact that all interactions were significant in the third time window, including a significant *condition x mention* interaction, which would not have been obtained if we did not include animacy in the model (as indicated by the poorer goodness-of-fit results of the *interaction model* compared to the *simple model* in chapter 4.1) is further evidence for the influence of animacy on pronoun resolution.

Time window	in ms	Fixed predictors						
		main effects			2-way interactions			3-way interaction
		pronoun condition	order of mention	animacy	condition x mention	condition x animacy	mention x animacy	condition x mention x animacy
1	0-200	0.1 (1.747)†	-0.07 (-1.162)					
2	200-400	-0.16 (-1.009)	-0.27 (-1.714)†					
3	400-600	0.82 (2.43)*	0.41 (1.212)	0.75 (2.193)*	-1.42 (-2.95)**	-1.07 (-2.22)*	-1.09 (-2.263)*	1.76 (2.589)**
4	600-800	1.32 (3.734)***	0.64 (1.805).	0.84 (2.356)*	-2.14 (-4.256)***	-1.33 (-2.643)**	-1.17 (-2.318)*	2.72 (3.833)***
5	800-1000	1.67 (4.509)***	0.89 (2.389)*	1.06 (2.853)**	-2.9 (-5.524)***	-1.59 (-3.026)**	-1.86 (-3.532)**	3.94 (5.308)***
6	1000-1200	1.38 (3.605)***	0.67 (1.736).	0.81 (2.101)*	-2.47 (-4.553)***	-0.91 (-1.67).	-1.63 (-2.988)**	2.73 (3.557)***
7	1200-1400	1 (3.646)***	0.27 (0.992)		-1.52 (-3.914)***			
8	1400-1600	0.97 (3.488)**	0.4 (1.429)		-1.74 (-4.42)***			
9	1600-1800	0.85 (3.114)**	0.53 (1.937)†		-1.73 (-4.459)***			
10	1800-2000	0.93 (3.426)***	0.93 (3.408)***		-2.01 (-5.222)***			

Table 4.27: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*hij* vs. *die*), order of mention (1st vs. 2nd) and animacy (animate vs. inanimate). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

4.3.4.1. Animate Items

Figure 4.11 shows the distribution pattern of looks to either target referent for the animate items. 2633 looks (52%) entered the analysis. As is shown in the figure, for animate antecedents there is a first-mention preference for the personal pronoun and a second-mention preference for the d-pronoun.

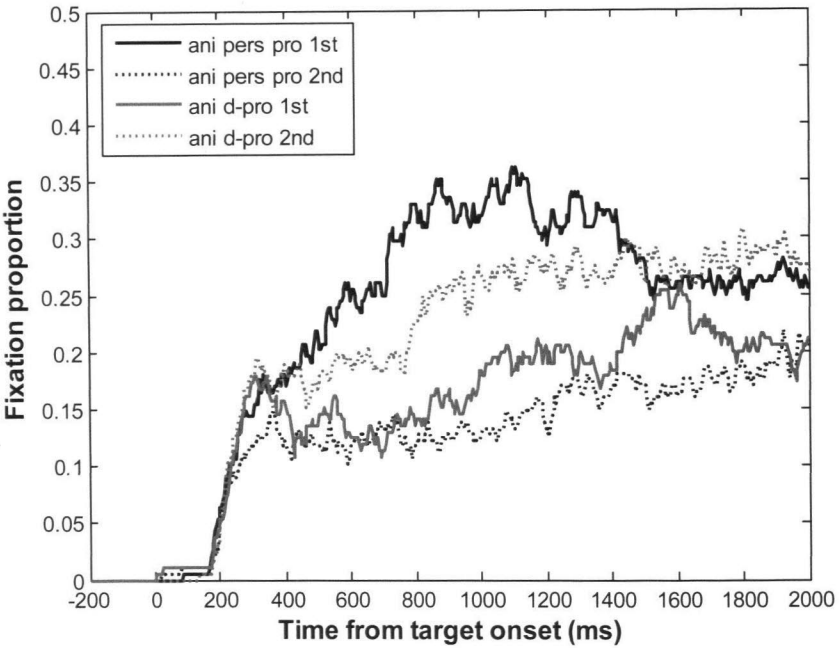

Figure 4.11: Animate items. Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

The individual analyses for the two pronouns revealed that following two animate antecedents, the personal pronoun was preferentially resolved towards the first-mentioned antecedent (see Table 4.28). This preference emerged at 400 ms after pronoun onset and lasted until the end of the analysis period (2000 ms) as indicated by the significant main effect of order of mention, which was marginally significant in the last time window.

For the d-pronoun, there was a marginally significant main effect of order of mention that emerged during time window 4 (600 – 800 ms), which was significant during the next time window (800 – 1000 ms), and which was again marginally significant during windows 9 (1600 – 1800 ms), and significant during time window 10 (1800 – 2000 ms). The positivity

of the beta coefficient reflected a second-mention preference for the d-pronoun.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	-0.03 (-0.234)	-0.09 (-0.751)
2	200-400	-0.33 (-1.059)	0.22 (0.735)
3	400-600	-1 (-2.947)**	0.41 (1.234)
4	600-800	-1.49 (-4.18)***	0.64 (1.918)†
5	800-1000	-2.01 (-5.365)***	0.89 (2.422)*
6	1000-1200	-1.81 (-4.694)***	0.67 (1.753)†
7	1200-1400	-1.47 (-3.788)***	0.63 (1.649)
8	1400-1600	-1.12 (-2.924)**	0.49 (1.209)
9	1600-1800	-0.8 (-2.107)*	0.65 (1.646)†
10	1800-2000	-0.68 (-1.768)†	0.86 (2.167)*

Table 4.28: Results of the individual time course analyses for each type of pronoun (*hij* vs. *die*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

The results of the animate items thus show a quite early and long-lasting first-mention effect for the personal pronoun (400 – 2000 ms), and a second-mention effect for the d-pronoun (emerging at 600 ms). Compared to the results obtained in chapter 4.1.4, the d-pronoun is resolved earlier for animate items, suggesting that animacy had an effect on its resolution.

	hij	*die*
animate & inanimate	1st at 400 ms	2nd at 1600 ms

Table 4.29: Direction and emergence of effects of the personal and the d-pronoun over animate and inanimate antecedents in German learners of Dutch

With regard to the time windows which were previously identified to be influenced by the factor animacy, namely windows 3, 4, 5, and 6, we observed a significant first-mention effect for the personal pronoun during all four time windows and a significantly second-mention effect for the d-pronoun in window 5, which is marginally significant in windows 4 and 6, but no effect for window 3.

4.3.4.2. Inanimate Items

Figure 4.12 shows the distribution of target looks for inanimate antecedents (2413 looks; 48%). We observe a first-mention preference for the personal pronoun which appears rather

late (around 1400 ms). For the d-pronoun the distribution looks less clear as there are initially more looks to the first-mentioned entity and ultimately more looks to the second-mentioned entity.

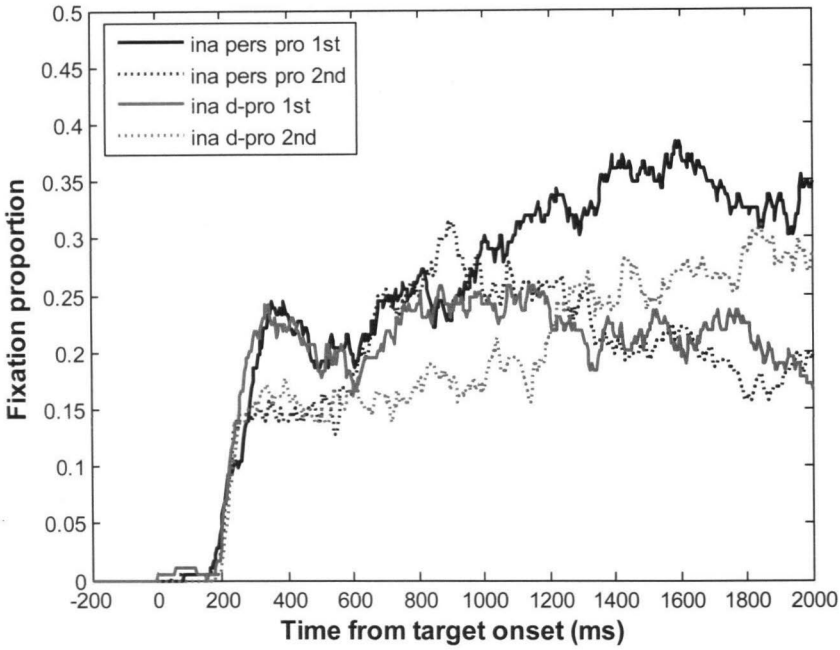

Figure 4.12: Inanimate items. Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

The individual analyses after inanimate items showed a first-mention effect for the personal pronoun which was marginally significant in time window 6 (1000 – 1200 ms) after which it became significant (1200 – 2000 ms) (see Table 4.30).

In case of the d-pronoun, there was an early marginally significant main effect of order of mention emerging at 200 ms after pronoun onset and lasting until 600 ms. During time windows 5 and 6 (800 – 1200 ms), the main effect became significant. The negativity of the beta coefficient indicated that the d-pronoun was initially resolved towards the first-mentioned entity. During the last time window (1800 – 2000 ms), a significant second-mention effect for the d-pronoun was again observed, as indicated by the positivity of its beta coefficient. So, interestingly, two different kinds of effects were in evidence. While there is an early first-mention effect for the d-pronoun, ultimately it is resolved towards the

second-mentioned entity. The first-mention preference starts to become visible at 200 ms. This effect is thus probably not caused by the pronoun, since it takes about 200 ms to launch a saccade and the pronoun takes about 200 ms to be uttered. But when confronted with an effect of inanimacy on discourse expectation, such that the first-mentioned topical antecedent is expected to be named (thus prior to pronoun resolution), we would expect this effect to show up in the *hij*-condition as well. The effect is thus difficult to interpret. Nevertheless the analysis showed that there was an earlier first-mention preference for the d-pronoun since the effect became significant during time windows 5 and 6. Ultimately, the preference changes into a second-mention preference.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	-0.04 (-0.382)	-0.11 (-1.07)
2	200-400	-0.36 (-1.156)	-0.6 (-1.796)†
3	400-600	-0.33 (-1.001)	-0.68 (-1.901)†
4	600-800	0.06 (0.16)	-0.52 (-1.461)
5	800-1000	0.07 (0.182)	-0.97 (-2.656)*
6	1000-1200	-0.7 (-1.726)†	-0.96 (-2.536)*
7	1200-1400	-1.02 (-2.551)*	-0.09 (-0.237)
8	1400-1600	-1.56 (-3.863)***	0.31 (0.797)
9	1600-1800	-1.59 (-3.976)***	0.42 (1.081)
10	1800-2000	-1.49 (-3.853)***	1 (2.575)**

Table 4.30: Results of the individual time course analyses for each type of pronoun (*hij* vs. *die*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

For the inanimate items, a first-mention effect was found for the personal pronoun which emerged 1200 ms and an earlier first-mention effect for the d-pronoun (800 ms) which ultimately switched into a second-mention effect (1800 ms). With regard to the time windows which had been shown to be affected by animacy, there was a marginally significant first-mention effect for the personal pronoun in time window 6, and a marginally significant first-mention effect for the d-pronoun during time window 3 which became significant in windows 5 and 6. This observation contrasts with the observation for animate items, where the first-mention effect of the personal pronoun was significant during all four time windows, and the d-pronoun showed a second-mention preference during windows 4, 5 and 6.

4.3.5. Discussion

Animacy had an influence on the resolution preferences of German learners of Dutch. Individual analyses for animate and inanimate items revealed an overall first-mention effect for the personal pronoun. For the d-pronoun, after two animate antecedents, a clear second-mentioned (target-like) preference was in evidence. However, after two inanimate entities, there was an early first-mention preference where the topical first-mentioned entity was preferred and this switched to a second-mention preference in a later time window, where the factor order of mention overrode this first-mention preference, leading to a target-like second-mention preference. The resolution of the personal pronoun also showed sensitivity towards animacy information; when resolved towards an animate first-mentioned entity the effect was immediate (400 ms), while following inanimate antecedent structures there was more ambiguity as indicated by the late disambiguation (1200 ms). This effect can be explained in two ways:

- A comparable timing difference was observed in the German native speakers due to the unmarkedness of the personal pronoun, and for the d-pronoun a switch in preferences was observed, such that the d-pronoun was ultimately resolved in a target-like way. The German L2 learners transferred resolution features from their L1, caused by the markedness-unmarkedness distinction of personal and d-pronouns in German.
- The results are due to general L2 processing effects. The Dutch L2 learners showed comparable resolution behavior regarding the timing differences observed for the personal pronoun, as well as an influence of animacy on the direction of the preferences for the d-pronoun.

		er	***der***
L2 Dutch learners of German	**animate**	1st at 400 ms	-
	inanimate	1st at 1000 ms	1st at 800 ms

		hij	***die***
L2 German learners of German	**animate**	1st at 400 ms	2nd at 600 ms
	inanimate	1st at 1200 ms	1st at 800 ms 2nd at 1800 ms

Table 4.31: Direction and emergence of effects of the personal and the d-pronoun for animate and inanimate antecedents in L2 German and Dutch

With regard to the Dutch learners of German, the direction and emergence of the first-mention preferences for personal pronouns were quite comparable. For the personal pronoun, both groups showed an earlier first-mention preference after animate items

compared to inanimate items. Interestingly, for the d-pronoun a change in preferences was observed for both groups. For the Dutch learners of German, there was no preference for the d-pronoun after animate items, whereas the German learners showed a target-like second-mention preference. After inanimate antecedents the Dutch learners showed a first-mention preference as did the German learners but only at earlier time windows. In the case of the Dutch learners of German, this preference had not yet ultimately switched. The no-preference pattern found after animate antecedents in the Dutch L2 learners hints at a possible trade off as the learners develop away from a first-mention-topic-only resolution function towards a more target-like resolution where the two pronouns have two different functions. In sum, animacy has a strong influence on L2 pronoun resolution.

In this section, we hoped to disentangle possible L1 transfer effects from proficiency effects. The patterns of results for both learner groups still do not allow us to decide whether the target-like second-mention preference for the d-pronoun in the German L2 learners was due to the markedness-unmarkedness distinction between the two pronouns in L1 German or the potentially higher proficiency of the German learners. The conclusion is that despite the target-like behavior of the German L2 learners of Dutch, there is at least one (non-target-like) learner component, i.e. an early first-mention preference for the d-pronoun following inanimate antecedents.

As discussed here and in chapter 4.1, the German learners of Dutch might have been more advanced than the Dutch learners of German. They showed a target-like second-mention preference for the d-pronoun, because they had already acquired the distinction of the functions between the two pronominal forms. The Dutch learners were less proficient and therefore still showed a first-mentioned topical preference for the d-pronoun after inanimate items. The fact that they showed no preference after two animate antecedents shows that they were also developing away from the two forms–one function system. In chapter 4.4, the potential role of the factor proficiency is investigated in more detail.

4.4. Individual Differences in L2 Pronoun Resolution - L2 Proficiency

4.4.1. Background

While the Dutch learners of German showed an overall preference to resolve the two types of pronouns towards the first-mentioned topical entity, the German learners of Dutch showed native-like resolution preferences, in that the d-pronoun was resolved towards the second-mentioned non-topical entity (see chapter 4.1). It is surprising that the d-pronoun was so difficult to resolve, when we consider that both source and target languages discriminate the functions of personal and d-pronouns in a similar way. The conclusion was that the observed difference between the learner groups was either due to

- language differences: the d-pronoun is marked for non-topical co-reference in German, while this is not the case in native Dutch
- proficiency differences: the L2 German learners of Dutch were more proficient than the Dutch L2 learners of German.

In the preceding section, it was shown that despite the overall differences between the groups, the learners showed a similar (and non-target-like) sensitivity towards animacy. Animacy affected the direction of the resolution preferences for the d-pronoun in that its resolution was more difficult to inanimate than animate entities. This indicates two things:

- Although the German L2 learners performed in a target-like way overall, we found non-target-like L2 processing of animacy.
- Despite the overall non-target-likeness of the Dutch L2 learners of German, the animacy analysis revealed that the previous assumption, namely that the learners had two forms for one function, was not true for all learners. Since their resolution preferences for personal and d-pronouns was not the same after animate antecedents, it must be the case that the Dutch learners do indeed distinguish between the two functions of personal and d-pronouns, at least to some degree. No preference was found for the d-pronoun following two animate antecedents. The variability which led to this non-preference may be explained in two ways:
  1. alternating gaze behavior between the first- and second-mentioned target indicating ambiguity
  2. inter-individual differences between participants: more proficient learners favor second-mentioned targets, and less proficient learners favor first-mentioned targets.

In this section, the question of whether proficiency had an influence on the Dutch learners' resolution preferences in German is examined. The Dutch learners were therefore split according to their proficiency levels as indicated by the German Placement Test Score

(according to the European Reference Frame) into a lower proficient (although still advanced) B-level group, and a highly proficient C-level group. The data of the experiment from chapter 4.1.3 were reanalyzed to see whether proficiency was a valid predictor for the L2 pronoun resolution preferences.

4.4.2. Split into proficiency groups

The Dutch learners were coded on the basis of their proficiency levels as revealed by the German Placement Test. The cut-off point to belong to a group of higher proficient learners versus lower proficient learners was to have a score of at least 21 points on the German Placement Test indicating a C1-level according to the European Reference Frame. Six participants were classified as belonging to the higher proficient C-level group, and 26 participants entered the lower proficiency B-level group (see Table 4.32). One participant entered the B-level group although the score actually indicated an A2-level.

B-level group		**C-level group**	
level	**N**	**level**	**N**
A2	1	C1	5
B1	12	C2	1
B2	13		
$\sum$	**26**	$\sum$	**6**

Table 4.32: L2 Dutch group split into the high proficiency C-Level-group (6 participants) and the lower proficiency B-level-group (26 participants) according to the German Placement Test Score

4.4.3. Results

4.4.3.1. Forced-Choice Questionnaire

The split into proficiency groups revealed for both groups that the association between type of pronoun and antecedent choice was significant, χ^2 (1) = 6.81, $p<.01$ for the B-level group, and χ^2 (1) = 31.42, $p<.001$ for the C-level group. For the C-level group, the odds ratio showed that choosing the first-mentioned antecedent was 10.49 times more likely for the personal pronoun than for the d-pronoun. The B-level group showed that the chance of choosing the first-mentioned entity when a personal pronoun was present was only 3.05 times higher than that for a d-pronoun.

	B-level group		C-level group	
	1st	2nd	1st	2nd
personal pronoun	97.76% (306)	2.24% (7)	90.14% (64)	9.86% (7)
d-pronoun	93.49% (287)	6.51% (20)	46.57% (34)	53.42% (39)

Table 4.33: L2 Dutch Results split into the high proficiency C-Level-group (6 participants) and the lower proficiency B-level-group (26 participants)

Individual analyses on each pronoun showed that the patterns of results were different from chance. In the C-level group, the test revealed for the personal pronoun χ^2 (1) = 27.39, p<.001 that the likelihood to choose the first-mentioned antecedent was 9.14 times higher than chance, while for the d-pronoun the pattern of results did not significantly differ from chance, χ^2 (1) = .17, p = .401[66]. In the B-level group, the results were different. For the two types of pronouns we found a first-mention preference, with χ^2 (1) = 185.14, p<.001 for the personal pronoun and χ^2 (1) = 143.31, p<.001 for the d-pronoun.

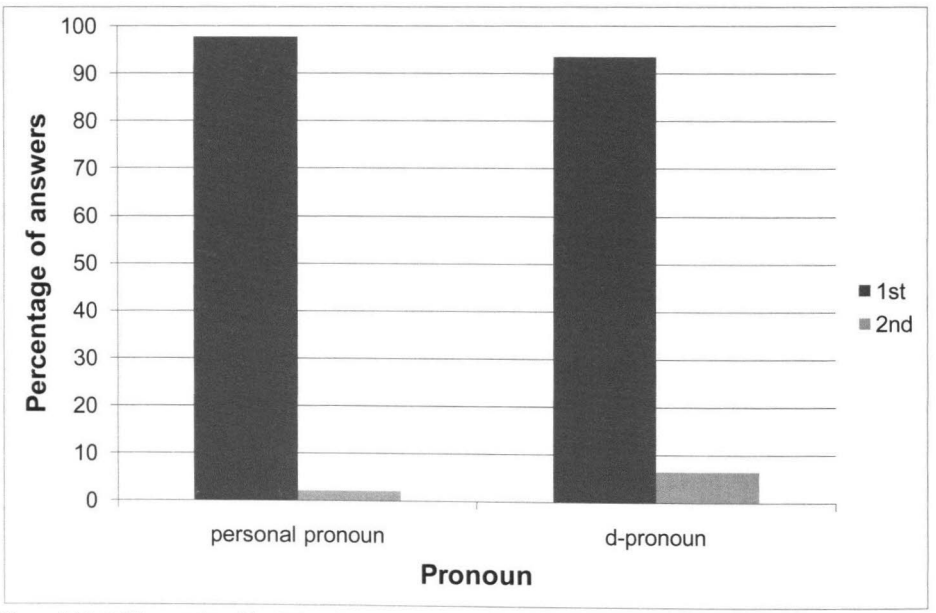

Figure 4.13: Offline results of the B-Level group reveal a topic-default resolution preference

[66] p>.05

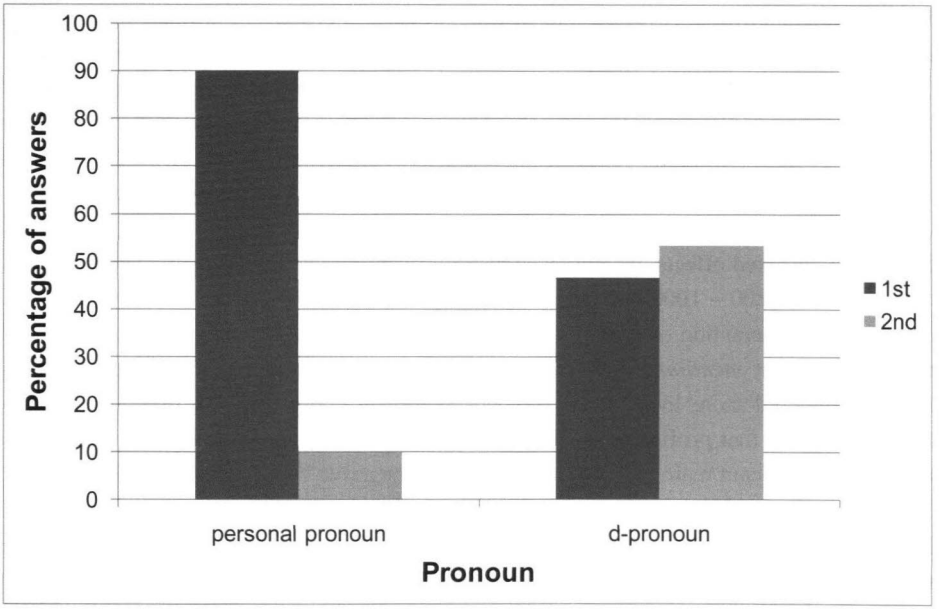

Figure 4.14: Offline results of the C-Level group reveal a targetlike second-mentioned tendency for the d-pronoun

There was thus a significant difference between the two proficiency groups. As expected, both groups showed a preference to resolve the personal pronoun towards the first-mentioned entity. But for the d-pronoun, a first-mention preference was only observed in the lower proficient B-level group.

4.4.3.2. Visual-World Eye-Tracking

To analyze the influence of proficiency on the eye-tracking results, we tested whether adding *proficiency* as a fixed factor to the linear mixed model would explain the data (of chapter 4.1.3) better than without this factor. Thus, the goodness-of-fit of the following three models was calculated for each of the ten 200 ms time slices from pronoun onset till 2 seconds after the pronoun onset:

1. *simple model:* containing the terms for the main effects of order of mention (*mention*) and type of pronoun (*condition*)
2. *interaction model:* the interaction term between order of mention and type of pronoun was added
3. *proficiency model:* the three-way interaction term between order of mention, type of pronoun and proficiency was added (intercept = condition: *d-pronoun*, mention: *1st*, proficiency: *B-level*)

The log-likelihood tests revealed that in time window 9 (1600 – 1800 ms) the data was significantly better explained by the *proficiency model* than by the *interaction model*. In time windows 5 (800 – 1000 ms) and 10 (1800 – 2000 ms), for which the *interaction model* did not explain the data better than the *simple model*, we however found that the *proficiency model* fit the data marginally better in window 5 and significantly better in window 10 than the *simple model*. This means that *proficiency* also had an effect on the eye-tracking data.

Turning to the fixed effects, we observed no significant effects for the *proficiency model* in time window 5 (800 – 1000 ms) (see Table 4.34). But in time window 9 (1600 - 1800 ms), the three-way interaction *condition x mention x proficiency* became significant. There was also a significant *mention x proficiency* interaction, indicating that the more proficient learners triggered more looks to the second-mentioned entity than the first. Overall this pattern indicated that proficiency had an effect on how the pronouns were resolved. There was also a significant main effect of *mention* with a negative beta coefficient indicating that overall more looks to the first-mentioned entity were made. In the last time window (1800 – 2000 ms), all of the above effects stayed significant. Additionally, we observed a significant *condition x proficiency* interaction. The main effect of *proficiency* was marginally significant with more overall looks for the less proficient B-level group.

To further evaluate the effects found for the *proficiency model*, individual analyses were conducted for the B-level and the C-level group. Note that while the B-level group consisted of 26 participants, the C-level group consisted of 6 participants only. Keeping in mind the relatively low number of fixations that therefore entered the analysis for the C-level group, we still calculated the resolution preferences separately for the two types of pronouns to better understand how proficiency influenced the resolution of personal and d-pronouns in the learners.

Time window	in ms	Fixed predictors						
		main effects			2-way interactions			3-way interaction
		pronoun condition	order of mention	proficiency	condition x mention	condition x proficiency	mention x proficiency	condition x mention x proficiency
1	0-200	-0.02 (-0.434)	0.04 (0.805)					
2	200-400	-0.02 (-0.094)	-0.08 (-0.454)					
3	400-600	0.08 (0.415)	-0.26 (-1.322)					
4	600-800	0.47 (1.64).	0.03 (0.116)		-0.72 (-1.754)†			
5	800-1000	0.17 (0.498)***	-0.26 (-0.796)	0.64 (1.178)	-0.28 (-0.603)	0.74 (0.991)	-0.54 (-0.74)	-1.23 (-1.167)
6	1000-1200	0.42 (1.399)	-0.6 (-1.997)*		-0.79 (-1.835)†			
7	1200-1400	0.71 (2.299)*	-0.54 (-1.767)†		-0.96 (-2.214)*			
8	1400-1600	0.8 (2.616)**	-0.37 (-1.239)		-1.06 (-2.465)*			
9	1600-1800	0.45 (1.348)	-0.74 (-2.216)*	-0.45 (-0.853)	-0.43 (-0.912)	0.87 (1.157)	2.01 (2.706)**	-2.3 (-2.163)*
10	1800-2000	0.19 (0.581)***	-0.97 (-2.974)	-0.85 (-1.644)**	0.15 (0.329)†	1.67 (2.252)	2.27 (3.108)*	-3.9 (-3.731)**

Table 4.34: Results of the time course analyses for the time segments following the onset of the pronoun for the fixed factors pronoun condition (*er* vs. *der*), order of mention (1st vs. 2nd) and proficiency (B-level vs. C-level). **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

B-level group

Figure 4.15 shows the resolution of personal and d-pronouns separately for the less proficient Dutch learners of German (3764 looks; 80%). The resolution pattern for personal and d-pronouns does not differ very much from the overall resolution pattern reported in chapter 4.1.3. This is probably due to the fact that the majority of fixations in the collapsed analysis come from less proficient learners. In the following, the statistical analysis for the B-level group which separately evaluated the factor order of mention for each type of pronoun is provided (see Table 4.35).

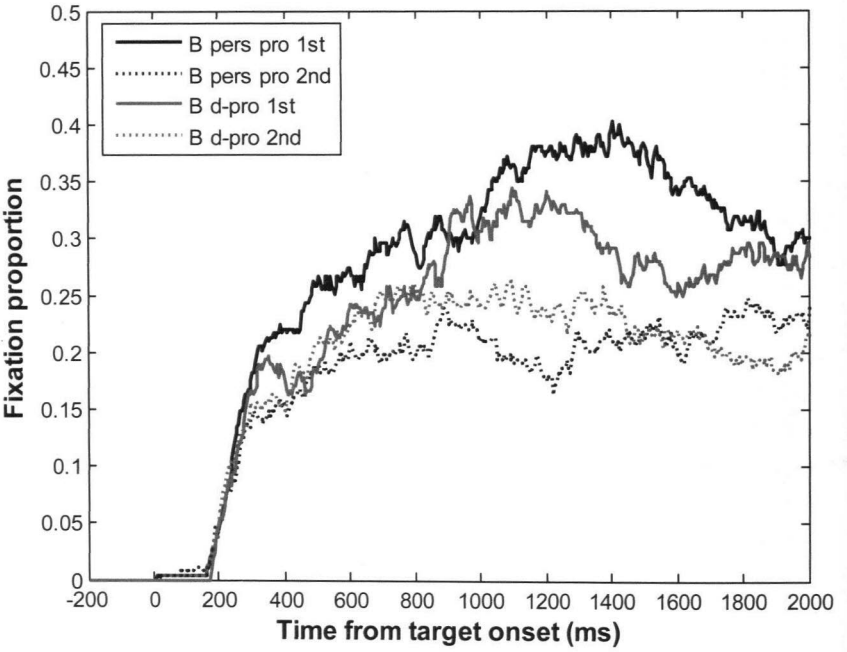

Figure 4.15: Online pronoun resolution of B-level group. Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

As expected, for the personal pronoun, a marginally significant main effect of order of mention emerged during time window 5 (800 - 1000 ms) which subsequently became significant (1000 – 2000 ms). The negativity of the beta coefficient for mention reflected the first-mention preference for the personal pronoun.

For the d-pronoun, there was a significant main effect of order of mention in window 6 (1000 – 1200 ms) which was marginally significant in time window 7 (1200 – 1400 ms),

and again significant in time windows 9 and 10 (1600 – 2000 ms). As for the personal pronoun, the negativity of the beta coefficient for mention reflected a first-mention preference for the d-pronoun.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	0.01 (0.117)	0.08 (0.938)
2	200-400	-0.32 (-1.2)	0.03 (0.116)
3	400-600	-0.21 (-0.977)	0.05 (0.174)
4	600-800	-0.12 (-0.529)	0.13 (0.398)
5	800-1000	-0.4 (-1.724)†	-0.26 (-0.799)
6	1000-1200	-0.98 (-4.08)***	-0.78 (-2.334)*
7	1200-1400	-0.95 (-3.901)***	-0.6 (-1.78)†
8	1400-1600	-0.94 (-3.91)***	-0.48 (-1.455)
9	1600-1800	-0.95 (-4.029)***	-0.74 (-2.279)*
10	1800-2000	-0.9 (-3.826)***	-0.97 (-2.996)**

Table 4.35: Results of the individual time course analyses for each type of pronoun (*er* and *der*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

These analyses showed that for the less proficient B-level group, there was a first-mention effect for the personal and for the d-pronoun. In time windows 9, and 10 which had previously been identified to be affected by the factor proficiency, we found significant effects for the two types of pronouns; in window 5 the first-mention effect for the personal pronoun was marginally significant. The preferences were not different from the overall preferences over both proficiency groups (chapter 4.1.3).

C-level group

The same analysis was conducted for the C-level group. Figure 4.16 shows the resolution of personal and d-pronouns separately for the more advanced Dutch L2 learners of German (965 looks; 20%). As expected, the direction of the preference for the personal pronoun was the same as for the B-level group, in that the personal pronoun was resolved towards the first-mentioned entity. Surprisingly however, a difference for the resolution of the d-pronoun was observed. While a first-mention preference for the d-pronoun was in evidence in the lower proficient B-level group, the more proficient C-level group showed a preference for the second-mentioned entity.

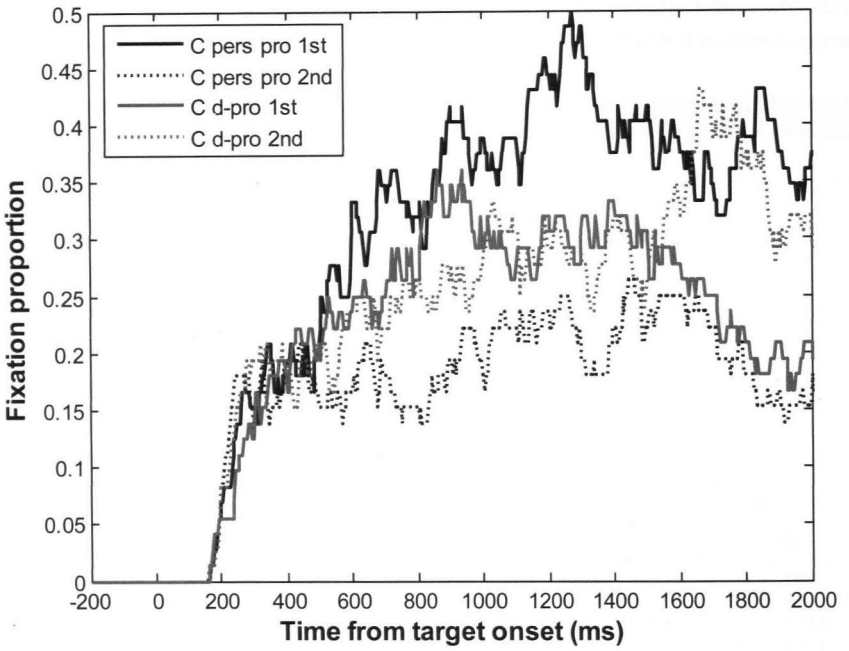

Figure 4.16: Online pronoun resolution of C-level group. Probability of fixating the first-and second-mentioned referent as a function of time in each of the two conditions (personal pronoun, d-pronoun)

Individual analyses were conducted for the two types of pronouns with order of mention as a predictor (see Table 4.36). For the *er*-condition, the analysis revealed a significant main effect of order of mention between 600 and 2000 ms. The negativity of the beta coefficient for mention indicated a first-mention preference. The intercept term was not significant for all of the significant time windows, which was probably due to the low number of fixations which entered the analysis and to the general high amount of variability in L2 learner data.

For the d-pronoun *der*, there was a marginally significant main effect of order of mention between 1600 and 1800 ms which subsequently became significant (1800 – 2000 ms). Interestingly, the positivity of the beta coefficient for mention reflected a second-mention preference for the d-pronoun *der*.

Time window	in ms	Fixed predictor: order of mention	
		personal pronoun	d-pronoun
1	0-200	0 (-0.012)	0.08 (0.537)
2	200-400	0.03 (0.061)	0.31 (0.557)
3	400-600	-0.62 (-1.012)	-0.27 (-0.426)
4	600-800	-1.92 (-2.886)**	-0.34 (-0.529)
5	800-1000	-2.32 (-3.448)***	-0.81 (-1.153)
6	1000-1200	-2.22 (-3.274)**	0.12 (0.171)
7	1200-1400	-2.29 (-3.238)**	-0.29 (-0.422)
8	1400-1600	-1.56 (-2.171)*	0.05 (0.078)
9	1600-1800	-1.47 (-2.128)*	1.27 (1.82)†
10	1800-2000	-2.46 (-3.735)***	1.29 (1.983)*

Table 4.36: Results of the individual time course analyses for each type of pronoun (*er* and *der*). Order of mention (1st vs. 2nd) was entered as a fixed predictor. **Note:** First numbers are coefficients. Numbers in parentheses are t-values. † $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$.

The resolution preferences found for the C-level group differed from those observed for the B-level group with regard to the d-pronoun. The B-level group showed a second-mention preference for the d-pronoun in the on- and off-line task. The C-level group, although showing no preference in the off-line task, resolved the d-pronoun towards the second-mentioned entity on-line.

4.4.4. Discussion

This section explored whether proficiency is a valid predictor for L2 pronoun resolution preferences. The findings showed that L2 learners at lower proficiency stages resolve d-pronouns differently than more highly proficient learners. For the lower proficient B-level group, an overall first-mention preference for the personal and the d-pronoun was observed, while there was a second-mention preference on-line for the d-pronoun in the more highly proficient C-level learners. The fact that there was no observed preference for the high proficient L2 learners in the off-line task cannot be elaborated on here; it may be due to the relatively low number of responses (6 participants) and the general variability in L2 data. We would need a larger data set to address this issue.

The findings are particularly interesting with regard to possible learner variety differences (Klein & Perdue, 1997). The interpretational differences for the d-pronoun between the low and high proficient L2 learners might stem from the fact that the lower proficient L2 learners have a *Basic Variety* language system while the higher proficient learners have

passed to the next learner variety. While the lower proficient learners follow the organizational principle that only topical antecedents are encoded by pronominal means, the resolution principles of the higher proficient learners may be influenced by preference patterns from the target language.

In sum, the conclusion is that proficiency is a crucial factor in the L2 resolution of personal and especially d-pronouns. Although Dutch L2 learners of German could in principle rely on their source language resolution preferences and apply them to L2 resolution, this is not what we found. The learners instead seemed to rely on general L2 processing strategies which were different from those of their L1. At lower proficiency levels, they have two pronominal forms for one function (to refer to the topic) and later they discriminate between the two functions, in that they resolve the d-pronoun towards the second-mentioned non-topical entity.

low proficient L2 learners	→	**high proficient L2 learners**
two forms, one function		*two forms, two functions*
topic co-reference		personal pronouns co-refer to topical entity
		d-pronouns co-refer to non-topical entity

4.5. General Discussion

The question of how Dutch learners of German and German learners of Dutch resolve personal and d-pronouns in their second languages was investigated in this chapter. German and Dutch were chosen because of the typological similarity between the two: personal pronouns are preferentially resolved towards first-mentioned topical antecedents, while d-pronouns preferred second-mentioned non-topical antecedents. We thus asked whether the learners were able to use their L1 knowledge to resolve the two pronouns in their L2.

As regards the effects of order of mention and topicality after canonical antecedent structures, it was found that the German learners of Dutch indeed showed target-like resolution preferences: a first-mentioned effect for personal pronouns and a second-mentioned effect for d-pronouns. The Dutch learners of German however showed a different pattern: they resolved both personal and d-pronouns towards the first-mentioned topical entity. The difference between the learner groups was discussed with respect to influences of the L1; German learners might have shown a second-mention preference for the d-pronoun because it is a marked form in German. On the other hand, proficiency differences between the groups to account for the observed difference were also considered.

The second part of the chapter focused on whether the learners would be equally sensitive towards disambiguating pragmatic information as the native speakers, i.e. we asked whether changing the word order of the antecedent structure would have an effect on their resolution preferences. Non-canonical comparative structures (e.g. *Heavier than the table is the cupboard*) had been shown to push the resolution of the two forms towards the focused second-mentioned entity in the native speakers. Thus, in contrast to canonical antecedent structures, the pronouns had shown to have overlapping functions when following non-canonical structures. Target-like resolution preferences in both learner groups were observed, which suggests that when the two pronouns have overlapping functions, the learners have no difficulty in resolving them. Only when discrimination between the functions of the two pronominal forms was required, such as when following canonical antecedent structures, we observed a difference.

In the third part of this chapter, the influence of animacy on L2 pronoun resolution was investigated. Two possible explanations of the differences between the learner groups in chapter 4.1 were considered: they could be due to learner variety differences or to proficiency differences. It was predicted that there would be no influence of animacy on the Dutch L2 learners of German, since they were assumed not to differentiate between the functions of personal and d-pronouns. Therefore they were expected to show their overall

first-mentioned topical preference across animate and inanimate items. On the other hand, the German L2 learners of Dutch had previously shown target-like resolution preferences. If these preferences were due to their relatively high degree of proficiency, then they were predicted to also show this target-like behavior with the factor animacy (no effect). If however their target-like preference was due to the fact that the d-pronoun is a marked form in German, then they were predicted to show source language-like sensitivity towards animacy (personal pronoun affected, d-pronoun unaffected). Surprisingly, it was found that the learner groups patterned together. They were both highly affected by animacy, in that it changed the direction of their resolution preferences for the d-pronoun. For the Dutch learners, there was a first-mention preference for the d-pronoun after inanimate antecedents, and no preference after animate antecedents. This finding was striking, as it revealed that the Dutch learners in fact discriminated at least to some extent between the two pronominal functions. Interestingly, for the Germans, after inanimate items there was an initial preference to resolve the d-pronoun towards the first-mentioned entity which later switched into a target-like second-mention preference. A second-mention effect was obtained after the animate antecedents. This suggested that although ultimately the learners had target-like resolution preferences, there was still a learner component to the process. Moreover, it is interesting that both learner groups disfavored the resolution of the d-pronoun towards second-mentioned inanimate antecedents. This could be due to the fact that they rely more on general discourse expectations which make it more likely that a second-mentioned animate antecedent will be topicalized than a second-mentioned inanimate antecedent.

In the fourth part of this chapter, the question of the effect of proficiency differences on L2 pronoun resolution was examined. Since the Dutch L2 learners had shown non-target-like resolution preferences, we asked whether the variation in the data could be accounted for by individual proficiency differences. The data from chapter 4.1 was reanalyzed to include proficiency as a predictor in the statistical model. The analysis revealed a significant effect of proficiency. While the lower proficient Dutch learners showed an overall first-mentioned topical preference, the higher proficient L2 learners showed a target-like non-topical preference for the d-pronoun. It was therefore concluded that learners at lower proficiency levels have two pronominal forms for one function (co-reference towards the topical entity) while learners at higher proficiency levels discriminate in a target-like fashion between the functions of personal and d-pronouns.

Summary and Conclusion

Chapter 5

In this dissertation I investigated how native speakers and L2 learners of German and Dutch resolve personal and d-pronouns in ambiguous discourse. More specifically the influence of the order of mention, the information status and the animacy of the antecedent candidates on the resolution preferences for the two pronouns was examined. In addition, individual differences concerning pronominal resolution preferences across participants were investigated by looking at differences in on- and off-line pronoun resolution and the learners' levels of proficiency in their respective target languages.

Previous psycholinguistic studies on the resolution of personal and d-pronouns have tried to answer the question how order of mention and grammatical role information influence the interpretation of the two pronominal forms, and therefore directly compared resolution preferences towards first-mentioned subjects and second-mentioned objects in SVO word orders versus first-mentioned objects and second-mentioned subjects in OVS word orders. But changing the word order of the antecedent structure also affects the information status of the antecedent candidates which in turn could have had an influence on pronoun resolution. Therefore, in the reported experiments, the pronouns were presented after comparative antecedent structures of the type *"The cupboard is heavier than the table. It..."*. This entailed that both antecedents were in nominative case, and so the question of how order of mention of the antecedent candidates affected the resolution of the two pronouns could better be addressed. Furthermore, the effect of presenting non-canonical antecedent word orders on pronoun resolution was examined, by use of comparative antecedent structures like *"Heavier than the table is the cupboard. It..."*. Furthermore, while previous studies have only looked at animate antecedents, the role of the animacy of the antecedent candidates was also taken into account. Two types of tasks were used: a visual-world eye-tracking task revealing on-line resolution preferences, and a forced-choice comprehension questionnaire which measured the final interpretation preferences. With this set-up, it was hoped that a better picture would be obtained of the participants' moment-by-moment inferences, while at the same time, making the results of the current experiments comparable to those of previous studies.

5.1. Summary

Chapter 3.1 presented experiments investigating the resolution of personal and d-pronouns after canonical comparative antecedent structures in German and Dutch. The eye-tracking results showed no difference in general resolution preferences between German and Dutch (personal pronoun to first-mentioned and d-pronoun to second-mentioned entity). However, the two languages differed as regards timing of the effects: in German, the resolution of the personal pronoun occurred late in comparison to the d-pronoun, whereas in Dutch resolution was immediate. This indicates that the German personal pronoun was more ambiguous than the Dutch personal pronoun, a finding which is in line with the suggestion that the German d-pronoun is marked for non-topicality while the German personal pronoun is more neutral in this regard. The results show that this appears not to be the case for Dutch.

The results of the off-line comprehension questionnaire were slightly different from what was found for the on-line eye-tracking task. For both languages, the resolution of the personal pronoun towards the first-mentioned entity was at ceiling, whereas the second-mentioned interpretation preference for the d-pronoun was only marginally significant. This finding is different from what has been found in previous off-line completion and acceptability judgment tasks after canonical SVO antecedent structures. While some found no preference for the personal pronoun and a clear second-mention effect for d-pronouns (for German: Bosch, et al., 2007a; Wilson, 2009), others have reported both a strong first-mention preference for personal pronouns and a second-mention preference for d-pronouns (for Dutch: Kaiser & Trueswell, 2004; for German: Roberts, et al., in prog.). The lack of definitive results on the off-line interpretation of personal and d-pronouns may be due to subtle differences in the materials and tasks. The off-line results obtained in this study might be different from others, because first our sentence materials were different, in that the antecedent sentence was a comparative structure (and without a potentially biasing main verb). Second, in contrast to the previous studies, in our comprehension questionnaire the participants had to make a first- or second-mentioned decision even if they were uncertain.

Chapter 3.2 presented the investigation into the question of whether the resolution preferences for personal and d-pronouns varied after non-canonical antecedent structures; in other words, whether the information structure of the antecedent clause had an influence. This issue seemed particularly important with regard to previous studies which directly compared pronoun resolution after SVO with OVS antecedent structures attempting to disentangle order of mention and grammatical role effects. Resolution preferences were elicited after non-canonical comparative antecedent structures of the type *"Heavier than the*

table is the cupboard. It...". The findings show that changes in the antecedent sentence structure affected the resolution of the two pronouns. Specifically, compared to the experiments in chapter 3.1 which used topic-comment antecedent structures, following a topic-focus antecedent structure, the second-mentioned entity was pragmatically marked and so the asymmetric resolution pattern for personal pronouns (first-mentioned) and d-pronouns (second-mentioned) following canonical antecedent structures was not replicated, in either German and Dutch. For both the personal and the d-pronoun, resolution to the second-mentioned referent was in evidence, an effect that occurred very soon after the presentation of the pronoun in the eye-tracking results. This indicated that the information structural information disambiguated the relationship immediately and suggested that focus information was a very strong cue in pronoun resolution. This is in line with research conducted on English personal pronouns which states that not only are topical entities salient for pronoun resolution, but that personal pronouns may also prefer entities in contrastive focus (Arnold, 1998, 1999, 2001; Cowles, 2003; Cowles, et al., 2007). Focus information has an influence on the resolution of both personal and d-pronouns, and leads to overlapping resolution behaviors of the two pronominal forms, not only when there is only one available antecedent, but even in the case where two potentially matching antecedents are available. This overlapping behavior cannot be explained by theories of reference which highlight the asymmetric distributional properties for the two pronouns.

Given that both pronominal forms in German and Dutch can refer to animate as well as inanimate entities, the effect of animacy on pronoun resolution was investigated (3.3). Previous studies have only presented animate antecedents in their materials; thus, it was not clear how generalizable the resolution preferences to inanimate entities may be. Surprisingly, the results revealed that in German pronoun resolution was affected by animacy, but in Dutch it was not. In both languages, the direction of the results was the same for animate and inanimate items; the personal pronoun was resolved towards the first-mentioned and the d-pronoun towards the second-mentioned entity. However the German personal pronoun was resolved relatively early following animate antecedents (at 600 ms; compared to 1400 ms on the collapsed items) suggesting that animacy information facilitated its resolution.

The last part of chapter 3 addressed the influence of individual differences on pronoun resolution. As mentioned above, while the direction of the off-line findings in chapter 3.1 patterned with the on-line findings, there was still a qualitative difference in the final interpretations; the first-mention preference for the personal pronoun was at ceiling while the second-mention preference for the d-pronoun was only marginally significant. By reanalyzing the off-line data, individual differences in interpretation preferences for the d-

pronoun were found. For the d-pronoun, there appeared to be participants who favored co-reference with the first and others with the second-mentioned entity. The data were reanalyzed to see whether such individual preferences had an effect on on-line resolution processing. Interestingly, we found for German speakers that even when the final interpretation for the d-pronoun was resolved towards the first-mentioned entity, there was a second-mention effect during the processing of the utterance (at 800 ms). This was argued to underline the markedness of the d-pronoun to refer to non-topical antecedents. In Dutch, the results were slightly different. Whenever a second-mentioned antecedent had been ultimately chosen during the off-line task, this preference was reflected in the on-line task. When a first-mentioned entity was chosen, there was no on-line preference; thus, when participants were uncertain about co-reference-relations, they chose the first-mentioned topical entity.

In the second part of the dissertation, I investigated how second language learners resolve personal and d-pronouns in their L2. More specifically, the German experiments were administered to Dutch L2 learners of German, and the Dutch experiments to German learners of Dutch. We examined whether L2 pronoun resolution would be affected by the order of mention of the antecedent candidates (4.1). For the Dutch learners of German, the on-line as well as the off-line results showed that this was not the case: they showed an overall preference to resolve the personal and the d-pronoun towards the first-mentioned topical entity. Two explanations were possible: the Dutch confounded the two pronominal forms, or they had two forms for one function, namely to refer to the topical antecedent. In contrast, the German L2 learners of Dutch indeed discriminated between the functions of the two pronominal forms, in that they resolved the personal pronoun towards the first-mentioned topical entity, and the d-pronoun towards the second-mentioned non-topical entity, and therefore showed target-like resolution preferences. Since the general rule, namely that personal pronouns prefer topics while d-pronouns are preferentially resolved towards non-topical antecedents, applies to both languages, it was unclear why the Dutch learners of German showed non-target-like behavior, while the German learners of Dutch showed target-like behavior. The following explanations were possible:

a) phonological similarity: the German personal and d-pronoun have a quite similar realization (*er* - *der*) compared to the distinct Dutch realization (*hij* - *die*) leading to a confound of the German personal and d-pronoun
b) input: the German d-pronoun might have appeared less frequently in the input of classroom settings than the Dutch d-pronoun (since it is also a proper demonstrative form and appears in written language)

c) proficiency: the Dutch learners of German were less proficient than the German learners of Dutch
d) form markedness and L1 transfer: since the German d-pronoun is marked to refer to the non-topic, the German learners of Dutch might have transferred this property from their L1

The first explanation was rejected with regard to the quality of the on-line effects and the off-line task. Although both forms were eventually resolved towards the first-mentioned entity, they showed timing differences: the d-pronoun was resolved later than the personal pronoun, thus suggesting that it was more ambiguous. The interaction between order of mention and pronoun type was significant, further indicating that the pronouns were resolved differently, although both towards the first-mentioned entity. During the off-line task the participants received the materials in written language and thus saw that the forms were different. We therefore excluded the possibility that the two pronominal forms were confounded, but rather pursued the possibility that the Dutch learners had two forms for one function. The second explanation, namely that the Dutch learners were lacking the d-pronoun from their language input seems more likely. However, Roberts et al. (in prog.) tested the L2 resolution preferences for personal and d-pronouns in Finnish L2 learners of German, who had also learned German in a classroom setting and found a robust second-mention preference for the d-pronoun after canonical antecedent structures. So, the second explanation does not seem likely, either. The role of proficiency was addressed in a special study (chapter 4.4, see also the summary below). With regard to learner variety language systems, a possible explanation for the observed resolution patterns is that the resolution pattern of the German L2 learners of Dutch was influenced by source language principles. In German, the d-pronoun is marked to co-refer to the non-topical entity. The German learners of Dutch might have transferred this feature from their L1. Thus, although German and Dutch are typologically very closely related and share the overall resolution preferences for personal and d-pronouns, the subtle "markedness" difference on the d-pronoun may induce the observed differences in the learner performances.

In chapter 4.2 the focus was on the information structure of the antecedent sentence, operationalized by presenting the comparative antecedent structure in a non-canonical order. Earlier studies on L2 German (Roberts, et al., in prog.; Wilson, 2009) have found no clear resolution preferences after non-canonical OVS antecedent structures and it may be hypothesized that the processing of OVS structures posed quite a high load on the processing system which might have affected subsequent pronoun resolution patterns. One indicator of such difficulties is the fact that when Wilson (2009) split up her learners into a more and a less proficient group, there was a difference in d-pronoun resolution preferences

(low = first-mention preference, high = second-mention preference), whereas there was no effect on the resolution of the d-pronoun after canonical SVO structures (low and high = no preference). Since in our study, both referents were presented in nominative case, we assumed that the possible rise in processing load of the non-canonical comparative antecedent structures was kept to a minimum. Furthermore, since there was cross-linguistic evidence from the native data that focus was a stronger cue than order of mention, we were interested whether this was also true for the learners. The results supported this claim. Both L2 learner groups were shown to be sensitive towards this information structural cue. The focus information overrode possible order of mention effects and the L2 learners had target-like second-mentioned resolution preferences. The off-line results patterned with the on-line results: the second-mentioned focused entity was chosen for the two types of pronouns. When the overall picture is examined, it must be concluded that the L2 learners have less difficulty with this type of antecedent structure—when the two pronouns had overlapping functions—than both the canonical sentences used in the current experiments and the non-canonical OVS structures used in earlier studies, since both learner groups behaved in a target-like fashion when no differentiation in functionality was necessary. The non-target-likeness seems to emerge when L2 comprehenders must discriminate between the functions of the two pronominal forms, as in canonical contexts where personal pronouns prefer topics whereas d-pronouns prefer non-topics. It is in these contexts, that the Dutch L2 learner preferences diverged from the native preferences.

In chapter 4.3 the effect of animacy on L2 pronoun resolution was examined. In the second language acquisition literature it has been suggested that L2 learners rely primarily on non-structural lexical-semantic and pragmatic cues when comprehending their second language (Clahsen & Felser, 2006a), therefore, animacy might be an important cue to L2 learners. For both groups of L2 learners, the direction of the resolution preference for the personal pronoun was the same for animate items as for inanimate items, only the timing differed in that the personal pronoun was resolved earlier for animate items. This suggests that animacy had a facilitatory influence on the first-mention effect which may be due to the fact that animate entities are more likely to be topics. Surprisingly, a change in preferences for the d-pronoun in both L2 learner groups was observed. The analysis of the Dutch L2 learners of German results revealed that the first-mention preference for the d-pronoun was only present following inanimate antecedents; there was no such preference after animate antecedents. The German L2 learners of Dutch showed a second-mention preference for the d-pronoun after animate items. However, after inanimate items they first showed a first-mention preference which later switched into a second-mention preference. This pattern of results is highly interesting. While Dutch L2 learners of German had previously shown an overall topic-default resolution strategy, regardless of type of pronoun, the animacy

analysis revealed that they did not treat the two pronouns identically; they have begun to discriminate between the two functions, at least when they followed animate antecedents. Secondly, it is very interesting that the German learners of Dutch, who showed a target-like second-mention preference for the d-pronoun when the items were collapsed (chapter 4.1), had an early first-mention preference for the d-pronoun when following inanimate antecedents; it was only later, that this switched to a second-mention preference. This finding indicates that across L2 learner groups, the second-mention preference for the d-pronoun is facilitated when it follows animate antecedents. This may be due to the explanation mentioned above that in cases of a topic change, the discourse is more likely to be continued with an animate entity, as it is more likely to function as a subsequent topic. The learners seem to rely greatly on these lexical-semantic and pragmatic cues when they are resolving pronouns in their second language.

Chapter 4.4 also examined the idea of proficiency differences in the learners. Since the German learners showed quite a homogenous target-like distribution, this made only sense for the Dutch L2 learners of German who had shown a non-target-like behavior in the experiment in chapter 4.1. The L2 learners were split into a low (level B) and a high proficiency (level C) group, and it was shown that the factor proficiency was a good predictor for their resolution preferences. Both groups showed a first-mention preference for the personal pronoun on-line and off-line. They differed however in their on-line resolution of the d-pronoun: the less proficient group resolved it towards the first-mentioned antecedent, while the more highly proficient group resolved it towards the second-mentioned antecedent. The off-line results reflected the on-line results for the low proficient group; the high proficient group showed no preference off-line, which might be due to the relatively low number of cases in the high proficient group (only 6 participants). However, the proficiency analysis showed that as L2 learners become more proficient, they develop from a system where they have two pronominal forms for one function towards a system where they discriminate between two functions, in a target-like manner.

5.2. The research questions

In chapter 1, 9 research questions were raised and each is answered below.

1. How does the *order of mention* of the antecedent candidates influence the resolution of personal and d-pronouns? (chapter 3.1)

The order of mention of the antecedent candidates influences the resolution of personal and

d-pronouns in an asymmetric manner; personal pronouns are preferentially resolved towards the first-mentioned, and d-pronouns towards the second-mentioned entity.

2. How does the *information structure* of the antecedent clause influence the resolution of personal and d-pronouns? (chapter 3.2)

Pronoun resolution preferences differ after non-canonical and canonical structures. Information structure has an influence on pronoun resolution in that focus information is an important cue to pronoun resolution.

3. How does the inherent semantic factor *animacy* influence the resolution of personal and d-pronouns? (chapter 3.3)

We found an influence of animacy on pronoun resolution only in German. While this influence does not change the direction of the preferences, it changes the timing of emergence of the effect, in that the personal pronoun is resolved earlier after animate antecedents than inanimate antecedents; thus there is less ambiguity.

4. Do pronoun resolution preferences vary inter-individually across participants or tasks? (chapter 3.4)

At least in the case of the d-pronoun in German there were individual differences across participants to either ultimately chose the first-mentioned or the second mentioned entity. Although this final interpretation differed, the on-line behavior was largely unaffected by these individual preferences; the d-pronoun was always resolved towards the second-mentioned entity. In Dutch, the first-mentioned entity was ultimately chosen when the ambiguity could not be resolved (at least on-line during the 2 seconds after pronoun onset). Although at first sight there were differences between on- and off-line results, the additional analyses showed that there was a correlation between both.

5. Is there cross-linguistic evidence for general resolution preferences for personal and d-pronouns? (chapters 3.1, 3.2, 3.3, 3.4)

In general, the resolution preferences of personal and d-pronouns pattern together in German and in Dutch. While personal pronouns prefer first-mentioned topical antecedents, d-pronouns prefer second-mentioned non-topical antecedents after canonical antecedent structures, when the pronouns are presented after topic-focus antecedent structures, they

show overlapping functions as they are both resolved towards the second-mentioned focused entity. These overlapping functions were observed in both German and Dutch. However, we also found a difference between the two languages: the German d-pronoun seemed to be particularly marked for resolution towards the non-topical entity in contrast to the personal pronoun which seemed to be relatively neutral in this regard. This markedness of the d-pronoun was observed through the relatively early and clear second-mention effect compared to the personal pronoun. The robustness of the on-line effect and its independence of animacy information or individual differences in final interpretation preferences underlines this analysis. In contrast, the German personal pronoun was affected by the animacy of the antecedents; its resolution was facilitated by the animacy of its antecedent. This is taken as support for the relatively unmarkedness and thus for the neutrality of the personal pronoun. It is in this last special feature of markedness of the pronominal forms that German and Dutch differ.

6. How do the factors *order of mention, information structure of the antecedent clause* and *animacy* influence L2 pronoun resolution? (chapters 4.1, 4.2, 4.3)

All factors which had an influence on native pronoun resolution also influenced L2 resolution. L1-L2 differences occurred for the factor order of mention for the Dutch L2 learners of German only, but not for the German learners of Dutch. This may reflect proficiency differences and/or differences in the learner varieties. The factor animacy not only changed the timing of the effects, but it even changed the resolution preferences of the d-pronoun in both learner groups, such that the resolution towards the non-topical second-mentioned entity was facilitated when referring to an animate entity.

7. Do L2 pronoun resolution preferences vary inter-individually across participants due to L2 proficiency? (chapter 4.4)

The factor proficiency accounted well for the variation in the Dutch L2 learner data. It affected the resolution of the d-pronoun: the learners switched from a first-mentioned topical preference to a target-like second-mentioned non-topical preference. Thus, L2 learners at a higher proficiency level discriminated between the functions of personal and d-pronouns.

8. Do L2 learners show the same resolution preferences for personal and d-pronouns as native speakers? (And if not, in what ways do they differ?) (chapter 4)

Yes and no. While the German L2 learners of Dutch showed native-like resolution preferences, the Dutch L2 learners (at least at lower proficiency levels) behaved in a non-native-like fashion.

9. Do Dutch L2 learners of German and German L2 learners of Dutch resolve personal and d-pronouns in their second languages likewise? (chapter 4)

Yes and no. When the personal and d-pronouns had overlapping functions (both pronouns preferred the second-mentioned focused entity), the learner groups patterned alike. However, when these functions needed to be discriminated, such as when the pronouns were presented after canonical antecedent structures, the learner groups showed different resolution behaviors: the Dutch L2 learners of German showed a non-target-like behavior while the German L2 learners of Dutch showed a target-like behavior. The findings point in two directions. First, the data of the Dutch L2 learners of German is best explained in terms of proficiency, since its distribution is not native-like. However, although the German L2 learners of Dutch resolved the pronouns in a native-like way overall, there were subtle differences observed. Specifically, following inanimate antecedents, the learners first had a preference to resolve the d-pronoun towards the first-mentioned entity, as did the Dutch L2 learners of German, but then they switched ultimately towards the target-like second-mentioned preference. Since in German the d-pronoun is marked to refer to a non-topical entity, it may well be that the German L2 learners of Dutch transferred this strong preference from their L1 and benefited from this transfer.

5.3. Future work

In this dissertation several results have been presented that invite further investigation. First, as the investigation of pronoun resolution after canonical and non-canonical antecedent structures has shown, the information status of the antecedent candidates has an influence on pronoun resolution. Therefore, future studies should be aware of the fact that order of mention and grammatical role information cannot be manipulated independently from information structure: pronoun resolution after SVO sentences might be different from pronoun resolution after OVS sentences, because the latter are information structurally marked. In addition, when used with L2 learners, OVS structures may well introduce additional processing costs resulting in more noise in the data. Instead of controlling for the information structure of the antecedent sentence, it would be interesting to study it systematically. The experiments in this thesis showed that contrastive focus is an important cue to pronoun resolution. Contrastive focus can be seen as a way to express

givenness and this might not be as different from topicality as a marker of salience. It would be interesting to find out how informational focus (newness) affects the resolution of personal and d-pronouns. This could for example be operationalized via intonation or conversational disfluencies (Arnold, Altmann, Fagnano, & Tanenhaus, 2004; Arnold, Fagnano, & Tanenhaus, 2003).

Another issue which was explored was the role of animacy. Pronoun resolution was shown to be sensitive to this factor (at least in German). I did however not address the question of how antecedents of different animacy levels would influence the resolution of personal and d-pronouns. It would therefore be highly interesting to construct items of mixed animacy (AI and IA) to explore whether the animacy level of the competitor could have an effect on pronoun resolution. This would also be highly interesting for second language acquisition as the L2 learners of this thesis have shown to be strongly influenced by animacy.

Chapter 3.4 explored the match between on-line and off-line results; in the case of the German d-pronoun, it turned out that while final resolution preferences differed (as measured by the off-line task), the on-line resolution showed an independently robust second-mention effect. Thus, this first of all underlines the importance of using on-line methods such as eye-tracking to study pronoun resolution, since it would not have been possible to detect this preference, only the off-line preferences been elicited. Second, it opens an area for future work; it would be highly interesting to study systematically the correlations between on- and off-line results. Do final preferences become visible during the on-line resolution? Furthermore, the fact that we found differences across tasks calls for more investigation. Especially, with regard to the off-line results on the interpretation of personal and d-pronouns, it would be interesting to use off-line tasks which measure spoken language comprehension as opposed to written comprehension. All the off-line comprehension studies that exist so far have used written off-line tasks but it is arguable more natural to study such phenomena in spoken language form.

With regard to the "spoken" sentence materials, an important question that is raised is what effect intonation might have on the resolution of the two pronominal forms. On the one hand, in the past the function of d-pronouns has been compared to the function of stressed personal pronouns in English; one could ask, how do Germans and Dutch stressed personal pronouns pattern? On the other hand, it is largely assumed among German grammarians that d-pronouns are most frequently stressed. But as Ahrenholz (2007) has shown, most occurrences found in a corpus of spoken German were in fact unaccented. Since the aim was to control for the potential confound of stress in the current set of experiments, both were unaccented. However, in future work it would be interesting to investigate the

influence of stress on the resolution of personal and d-pronouns, above all in L2 learners, as this might be another reliable non-structural cue.

Another interesting topic concerns the resolution preferences for L2 learners at different proficiency levels. We found that L2 learners develop from a non-target-like topic-first strategy in pronoun resolution after canonical antecedent structures towards a more target-like system, where they discriminate between the functions of personal and d-pronouns. But when the two forms have overlapping functions, the learners showed no resolution difficulties. Due to the relatively low number of high proficient learners in the L2 Dutch group no definite conclusions can be drawn on this issue. Future studies could focus on these developmental aspects and more systematically investigate different levels of proficiency in order to find out when and how learners become more target-like in their resolution behavior and whether the developmental pattern observed in the Dutch learners would be borne out by learners of different L1-L2 language pairings. As reported in chapter 4.1, the German L2 learners showed relatively target-like resolution behavior, even though they had learned Dutch for only 12 months. Since in German the d-pronoun is marked for non-topical reference, it is likely that the learners benefited from this L1 feature when processing their L2. To further investigate L1 transfer issues, the resolution of personal and d-pronouns should be studied with different source and target languages, as well as in learners from L1 backgrounds which do or do not have d-pronouns or which differ in the markedness feature of such pronouns.

Last but not least, while this thesis shows that personal and d-pronouns can have overlapping functions as well as distinct functions, depending on the linguistic context, it would be interesting to get a clearer picture on how d-pronouns differ from demonstratives. German would be a particularly suitable language for addressing this question empirically, as there is a formal distinction between the d-pronoun *der* and the most closely related demonstrative form *dieser*. This has been studied by Ahrenholz (2007) in production, but has not yet been systematically addressed in comprehension, and it would help us at better understanding, whether and how to classify d-pronouns, which sometimes resemble personal and sometimes demonstrative pronouns.

References

Ahrenholz, B. (2007). *Verweise mit Demonstrativa Im Gesprochenen Deutsch. Grammatik, Zweitspracherwerb und Deutsch als Fremdsprache.* Berlin and New York: Walter De Gruyter Inc.

Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition, 93*, B79–B87.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247-264.

Ariel, M. (1990). *Accessing NP antecedents.* London: Routledge.

Ariel, M. (1994). Interpreting anaphoric expressions: a cognitive versus a pragmatic approach. *Journal of Linguistics, 30*, 3-42.

Ariel, M. (2001). Accessibility theory: an overview. In T. Sanders, J. Schilperoord & W. Spooren (Eds.), *Text representation, linguistic and psycholinguistic aspects* (pp. 29-87). Amsterdam/ Philadelphia: John Benjamins.

Arnold, J. E. (1998). *Reference Form and Discourse Patterns.* Stanford University.

Arnold, J. E. (1999). *Marking salience: The similarity of topic and focus.*Unpublished manuscript, University of Pennsylvania.

Arnold, J. E. (2001). The effects of thematic roles on pronoun use and frequency of reference. *Discourse Processes, 31*(2), 137-162.

Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes, 23*(4), 495-527.

Arnold, J. E., Altmann, R., Fagnano, M., & Tanenhaus, M. K. (2004). The Old and Thee, uh, New. *Psychological Science*, 578-582.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The immediate use of gender information: eyetracking evidence of the time-course of pronoun resolution. *Cognition, 76*, B13-B26.

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research, 32*(1), 25-36.

Audring, J. (2006). Pronominal Gender in Spoken Dutch. *Journal of Germanic Linguistics, 18*(2), 85-116.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

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

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457-474.

Bates, D. M., & Sarkar, D. (2007). lme4: Linear mixed-effects models using S4 classes (version 0.999375-27) [software application]: Retrieved from http://www.r-project.org.

Bini, M. (1993). La adquisición del italiano: mas allá de las propiedades sintácticas del parámetro pro-drop [the acquisition of Italian: beyond the syntactic properties of the pro-drop parameter]. In L. Liceras (Ed.), *La Linguistica y el Analisis de los*

Sistemas no Natives: Doverhouse.

Bittner, D. (2007). *Influence of animacy and grammatical role on production and comprehension of intersentential pronouns in German L1-acquisition1*. Paper presented at the Conference on Intersentential Pronominal Reference in Child and Adult Language.

Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1.18).

Bosch, P., Katz, G., & Umbach, C. (2007a). The Non-subject Bias of German Demonstrative Pronouns. In M. Schwarz-Friesel, M. Consten & M. Knees (Eds.), *Anaphors in Text: Cognitive, formal and applied approaches to anaphoric reference* (pp. 145-164). Amsterdam: Benjamins.

Bosch, P., Rozario, T., & Zhao, Y. (2003). *Demonstrative Pronouns and Personal Pronouns. German der vs. er.* Paper presented at the EACL2003, Budapest.

Bosch, P., & Umbach, C. (2007b). *Reference determination for demonstrative pronouns.* Paper presented at the Conference on Intersentential Pronominal Reference in Child and Adult Language.

Bouma, G., & Hopp, H. (2007). *Coreference Preferences for Personal Pronouns in German.* Paper presented at the Conference on Intersentential Pronominal Reference in Child and Adult Language.

Broeder, P. (1991). *Talking about people: a multiple case study on adult language acquisition* (Vol. 1). Amsterdam/Lisse: Swets & Zeitlinger.

Byron, D., & Allen, J. (1998). *Resolving demonstrative pronouns in the TRAINS93 corpus.* Paper presented at the Second Colloquium on Discourse Anaphora and Anaphor Resolution (DAARC-2).

Carminati, M. N. (2002). *The Processing of Italian Subject Pronouns.* University of Massachusetts Amherst.

Carreiras, M., Gernsbacher, M. A., & Villa, V. (1995). The advantage of first mention in Spanish. *Psychonomic Bulletin & Review, 2*, 124–129.

Chafe, W. L. (1974). Language and consciousness. *Language*, 111-133.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25-55). New York: Academic Press.

Clahsen, H., & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics, 27*, 3-42.

Clahsen, H., & Felser, C. (2006b). Continuity and shallow structures in language processing. *Applied Psycholinguistics, 27*, 107-126.

Clifton, C., Jr., Traxler, M. J., Mohamed, M. T., Williams, R. S., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language, 49*, 317-334.

Comrie, B. (1994). *Coreference: Between Grammar and Discourse.* Paper presented at the Kansai Linguistic Society, Osaka.

Comrie, B. (1997). *Pragmatic Binding: Demonstratives as anaphors in Dutch.* Paper presented at the The 23rd Annual Meeting of the Berkeley Linguistics Society, Berkeley.

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

Cowles, H. W. (2003). *Processing Information Structure: Evidence from Comprehension and Production.* UC San Diego.

Cowles, H. W., Walenski, M., & Kluender, R. (2007). Linguistic and cognitive prominence in anaphor resolution: Topic, contrastive focus and pronouns. *Topoi, 26*, 3-18.

Crawley, R., Stevenson, R., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research, 4*, 245–264.

De Swart, P., Lamers, M., & Lestrade, S. (2008). Animacy, argument structure, and argument encoding. *Lingua, 118*, 131-140.

Diessel, H. (1999). *Demonstratives. Form, Function, and Grammaticalization.* Amsterdam: John Benjamins.

Dietrich, R., & Van Nice, K. Y. (in press). Belebtheit, Agentivität und inkrementelle Satzproduktion. In C. Habel & T. Pechmann (Eds.), *Sprachproduktion.* Wiesbaden: Westdeutscher Verlag.

Drosdowski, G., Eisenberg, P., Gelhaus, H., Henne, H., Sitta, H., & Wellmann, H. (Eds.). (1995) Duden. Mannheim, Leipzig, Wien, Zürich: Dudenverlag.

Dussias, P. (2003). Syntactic ambiguity resolution in L2 learners: Some effects of bilinguality on L1and L2 processing strategies, *Studies in Second Language Acquisition* (Vol. 25, pp. 529-557).

Felser, C., Roberts, L., Gross, R., & Marinis, T. (2003). The processing of ambiguous sentences by first and second language learners of English, *Applied Psycholinguistics* (Vol. 24, pp. 453-489).

Frederiksen, J. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes, 4*, 323–347.

Frenck-Mestre, C. (1997). Examining second language reading: An on-line look. In A. Sorace, C. Heycock & R. Shillcock (Eds.), *Language Acquisition: Knowledge Representation and Processing - Generative Approaches to Language Acquisition* (pp. 474-478). Edinburgh: HCRC.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. Herredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217-236). Amsterdam: Elsevier.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly J. of Experimental Psychology, 50*(A), 119-148.

Fukumura, K., & Van Gompel, R. P. G. (accepted). The effect of animacy on the choice of referring expression. *Language and Cognitive Processes.*

Gass, S. (1987). The resolution of conflicts among competing systems: a bidirectional perspective. *Applied Psycholinguistics, 8*(4), 329-350.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition, 32*, 99-156.

Gernsbacher, M. A. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Erlbaum.

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language, 27*, 699–717.

Gernsbacher, M. A., Hargreaves, D. J., & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of

clause recency. *Journal of Memory and Language, 28*, 735-755.

Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study.* Philadelphia/ Amsterdam: John Benjamins.

Givón, T. (1984). Universals of Discourse Structure and Second Language Acquisition. In W. E. Rutherford (Ed.), *Language Universals and Second Language Acquisition* (pp. 109-136). Amsterdam: John Benjamins.

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

Gordon, P. C., Hendrick, R., & Foster, K. L. (2000). Language comprehension and probe-list memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 766-775.

Grober, E. H., Beardsley, W., & Caramazza, A. (1978). Parallel function strategy in pronoun assignment. *Cognition, 6*, 117-133.

Gundel, J. K. (1980). *Zero NP-anaphora in Russion: a case of topic-prominence.* Paper presented at the the 16th Annual Meeting of the Chicago Linguistic Society.

Gundel, J. K. (2003). *Information Structure and Referential Givenness/Newness: How Much Belongs in the Grammar?* Paper presented at the HPSG03 Conference, Michigan State University, East Lansing.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language and Cognitive Processes, 69*, 274-307.

Gürel, A. (2003). *Is the overt pronoun constraint universal? Evidence from L2 Turkish.* Paper presented at the The 6th Generative Approaches to Second Language Acquisition, Somerville, MA.

Gürel, A. (2004). Selectivity in L2-induced attrition: A psycholinguistic account. *Journal of Neurolinguistics, 17*, 53-78.

Hegarty, M., Gundel, J. K., & Borthen, K. (2002). Information structure and the accessibility of clausally introduced referents. *Theoretical Linguistics*, 1-24.

Helbig, G., & Buscha, J. (2001). *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht.* Leipzig, München: Langenscheidt.

Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguistis. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision and action* (pp. 1-58). Hove: Psychology Press.

Hopp, H. (2007). *Ultimate Attainment at the Interfaces in Second Language Acquisition: Grammar and Processing* (Vol. 65). Groningen: Groningen Dissertations in Linguistics.

Igoa, J. M. (1996). The relationship between conceptualization and formulation processes in sentence production: some evidence from Spanish. In M. Carreiras, J. E. García-Albea & N. Sebastián-Gallés (Eds.), *Language Processing in Spanish* (pp. 305-351). Mahwah, NJ: Lawrence Erlbaum Associates.

Jackson, C. N. (2008). Processing strategies and the comprehension of sentence-level input by L2 learners of German. *System, 36*, 388-406.

Jackson, C. N., & Roberts, L. (accepted). Animacy affects the processing of subject-object ambiguities in the L2: Evidence from self-paced reading with German L2 learners of Dutch. *Applied Psycholinguistics*.

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

Järvikivi, J., Van Gompel, R. P. G., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution: Contrasting the first-mention and subject-preference accounts. *Psychological Science, 16*(4), 260-264.

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

Juffs, A. (1998). Main verb vs. reduced relative clause ambiguity resolution in second language sentence processing. *Language Learning, 48*, 107-147.

Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in wh-extraction. *Studies in Second Language Acquisition, 17*, 483-516.

Kaiser, E., & Trueswell, J. (2004). *The referential properties of Dutch pronouns and demonstratives: Is salience enough?* Paper presented at the Sinn und Bedeutung 8, FB Sprachwissenschaft. Konstanz, Germany.

Kaiser, E., & Trueswell, J. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes, 23*(5), 709-748.

Kameyama, M. (1999). Stressed and Unstressed Pronouns: Complementary Preferences. In P. Bosch & R. v. d. Sandt (Eds.), *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge: University Press.

Kempe, V., & MacWhinney, B. (1998). The acquisition of case markings by adult learners of Russian and German. *Studies in Second Language Acquisition, 20*(4), 543-587.

Kilborn, K. (1989). Sentence processing in a second language: the timing of transfer. *Language and Speech, 32*(1), 1-23.

Kim, J.-H., & Montrul, S. (2003). *Binding interpretations in Korean heritage speakers.* Paper presented at the the 28th Boston University Conference on Language Development, Boston.

Klein, W. (2008). The topic situation. In B. Ahrenholz, U. Bredel, W. Klein, M. Rost-Roth & R. Skiba (Eds.), *Empirische Forschung und Theoriebildung. Beiträge aus Soziolinguistik, Gesprochene-Sprache- und Zweitspracherwerbsforschung. Festschrift für Norbert Dittmar zum 65. Geburtstag* (pp. 287-305). Frankfurt am Main: Peter Lang.

Klein, W., & Perdue, C. (1997). The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research, 13*(4), 301-347.

Klein, W., & Rieck, B.-O. (1982). Der Erwerb der Personalpronomina im ungesteurten Spracherwerb. *Zeitschrift für Literaturwissenschaft und Linguistik, 22*(86), 67-92.

Krifka, M., & Féry, C. (2008). Information Structure. Notional distinctions, ways of expression. In P. v. Sterkenburg (Ed.), *Unity and diversity of languages* (pp. 123-136). Amsterdam: John Benjamins.

Lakoff, G. (1971). Presupposition and Relative Well-Formedness. In D. Steinberg & L. Jakobovits (Eds.), *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*. Cambridge: Cambridge University Press.

Lambrecht, K. (1994). *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge UK: Cambridge

University Press.
Lamers, M., & De Hoop, H. (2005). Animacy Information in Human Sentence Processing: An Incremental Optimization of Interpretation Approach *Constraint Solving and Language Processing* (Vol. 3438, pp. 158-171). Berlin / Heidelberg: Springer.
Levinson, S. C. (1987). Pragmatics and the grammar of anaphora: a partial pragmatic reduction of binding and control phenomena. *Journal of Linguistics, 23*, 379-434.
Levinson, S. C. (1991). Pragmatic reduction of the binding conditions revisited. *Journal of Linguistics, 27*, 107-161.
LoCoco, V. (1987). Learner comprehension of oral and written sentences in German and Spanish: the importance of word order. *In B. VanPatten, T. R. Dvorak & J. Lee (Eds.), Foreign Language Learning: A Research Perspective* (pp. 119–129). Newbury House: Cambridge, MA.
Luján, M. (1985). *Binding Properties of overt pronouns in null pronominal languages.* Paper presented at the Chicago Linguistic Society, University of Chicago.
Luján, M. (1986). *Stress and binding of pronouns.* Paper presented at the Chicago Linguistic Society, University of Chicago.
Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language, 47*, 50-68.
Mak, W. M., Vonk, W., & Schriefers, H. (2006). Processing relative clauses in Dutch: When rocks crush hikers. *Journal of Memory and Language, 54*, 466-490.
Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27*, 53–78.
Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information processing time with and without saccades. *Perception and Psychophysics, 53*(4), 372-380.
McDonald, J. (1987). Sentence interpretation in bilingual speakers of English and Dutch. *Applied Psycholinguistics, 8*(4), 379-412.
Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition, 7*, 125-142.
Musan, R. (2010). *Informationsstruktur* (Vol. 9). Heidelberg: Universitätverlag Winter.
Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition, 24*, 501-528.
Prat-Sala, M., & Branigan, H. P. (2000). Discourse Constraints on Syntactic Processing in Language Production: A Cross-linguistic Study in English and Spanish. *Journal of Memory and Language, 42*, 168-182.
Reinhart, T. (1982). Pragmatics and linguistics: an analysis of sentence topics. *Philosophica, 27*, 53-94.
Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood squares; Looking at things that aren't there anymore. *Cognition, 76*, 269-295.
Roberts, L., Gullberg, M., & Indefrey, P. (2008). Subject pronoun resolution in second language discourse: an eye-tracking study with Turkish and German learners of Dutch. *Studies in Second Language Acquisition, 30*(3), 333-357.
Roberts, L., Järvikivi, J., Ellert, M., & Schumacher, P. (in prog.). Resolving the pronouns er and der in L2 German discourse: A visual-world eye-tracking study with Finnish L2 learners. *Ms. MPI for Psycholinguistics.*

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics, 1*, 75–116.

Sasaki, Y. (1994). Paths of processing strategy transfers in learning Japanese and English as foreign languages. *Studies in Second Language Acquisition, 16*(1), 43-71.

Schreiber, M. (1999). *Textgrammatik. Gesprochene Sprache. Sprachvergleich. Proformen im gesprochenen Französischen und Deutschen.* Frankfurt/M.: Lang.

Schumacher, P., Roberts, L., Ellert, M., & Järvikivi, J. (in prog.). Resolving the German subject pronouns er and der in auditory sentence comprehension: A visual world eye-tracking study. *Ms. MPI for Psycholinguistics.*

Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior, 13*, 272–281.

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research, 23*(197–229).

Sorace, A. (2000). *Differential Effects of Attrition in the L1 Syntax of Near-native L2 speakers.* Paper presented at the 24th Annual Boston University Conference on Language Development, Boston.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian, *Second Language Research* (Vol. 22, pp. 339-368).

Stevenson, R. (1996). Mental Models, Propositions, and the Comprehension of Pronouns. In J. Oakhill & A. Garnham (Eds.), *Mental Models in Cognitive Science. Essays in Honor of Phil Johnson-Laird.* Erlbaum: Psychology Press.

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

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632-1634.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language, 47*, 69-90.

Trueswell, J., Tanenhaus, M. K., & Garnsey, S. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285-318.

Tsimpli, I. M., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: a study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism, 8*, 257-277.

Van Kampen, J., & Pinto, M. (2007). *Germanic and Romance discourse devices in acquisition.* Paper presented at the LOT, Utrecht.

Van Nice, K. Y., & Dietrich, R. (2003). Task-sensitivity of animacy effects: Evidence from German picture descriptions. *Linguistics, 41*(5), 825-849.

Ward, G., & Birner, B. J. (2001). Discourse and Information Structure. In D. Schiffrin, D. Tannen & H. Hamilton (Eds.), *Handbook of Discourse Analysis* (pp. 119-137). Oxford: Basil Blackwell.

Webber, B. L. (1991). Structure and ostension in the interpretation of discourse deixis.

Language and Cognitive Processes, 6(2), 107-135.

Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophsysiology, 36*, 559-570.

Weinrich, H. (1993). *Textgrammatik der deutschen Sprache*. Mannheim: Duden.

Williams, J., Möbius, P., & Kim, C. (2001). Native and non-native processing of English wh-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics, 22*, 509-540.

Wilson, F. (2009). *Processing at the syntax-discourse interface in second language acquisition.* University of Edinburgh, Edinburgh.

Zifonun, G., Hoffmann, L., Strecker, B., & al. (1997). *Grammatik der deutschen Sprache*. Berlin, New York: de Gruyter.

Appendix

Appendix A: Experimental items

1 G Der Arzt ist freundlicher als der Koch. Der ist heute Morgen recht früh aufgestanden. Nur die Bäckerin war noch früher wach als er.

D De arts is vriendelijker dan de kok. Hij is vanochtend redelijk vroeg opgestaan. Alleen de bakker was nog eerder wakker dan hij.

E The doctor is friendlier than the cook. He got up quite early this morning. Only the baker (fem) was awake even earlier than he.

2 G Der Polizist ist schlauer als der Gärtner. Der ist nun schon zum zehnten Mal in eine neue Wohnung gezogen. Die letzte Wohnung war von Mäusen befallen gewesen.

D De politieagent is slimmer dan de tuinman. Die is nu al voor de tiende keer naar een nieuwe woning verhuisd. De laatste woning werd bewoond door muizen

E The policeman is cleverer than the gardener. He has moved to a new apartment for the tenth time now. The last apartment was infested with mice.

3 G Der Matrose ist netter als der Kapitän. Der will dieses Wochenende in den Zoo fahren. Der Zoo ist bekannt, viele Papageien zu haben.

D De matroos is aardiger dan de kapitein. Hij wil dit weekend naar de dierentuin rijden. De dierentuin staat er bekend om dat het veel papagaaien heeft.

E The sailor is nicer than the captain. He wants to drive to the zoo this weekend. The zoo is known to have many parrots.

4 G Der Lehrer ist glücklicher als der Musiker. Der wollte gestern Abend ins Kino gehen. Leider waren die Plätze schon ausverkauft, sodass er stattdessen ins Schwimmbad ging.

D De leraar is gelukkiger dan de musicus. Die wilde gisteravond naar de bioscoop gaan. Helaas waren de plaatsen uitverkocht, waarop hij besloot om naar het zwembad te gaan.

E The teacher is luckier than the musician. He wanted to go to the movies yesterday night. Unfortunately, the seats were already sold out, so that he went to the swimming pool instead.

5 G Der Feuerwehrmann ist dümmer als der Boxer. Der will nächste Woche mit einer Reisegruppe in den Urlaub fahren. Das Meer ist zu dieser Jahreszeit jedoch noch

ziemlich kalt.

D De brandweerman is dommer dan de bokser. Hij wil volgende week met een reisgroep op vakantie gaan. De zee is in dit jaargetijde echter nog behoorlijk koud.

E The fireman is sillier than the boxer. He wants to go on vacation with a tour group next week. The sea is however still pretty cold at this season of the year.

6 G Der Clown ist frecher als der Indianer. Der will für nächste Woche eine große Feier organisieren. Dort soll eine Klavierspielerin auftreten.

D De clown is brutaler dan de indiaan. Die wil voor volgende week een groot feest organiseren. Daar zal een pianospeelster spelen.

E The clown is cheekier than the Indian. He wants to organize a big party for next week. A pianist (fem) is supposed to perform there.

7 G Der König ist bekannter als der Zauberer. Der hat seit längerer Zeit starken Husten. Vielleicht wird der Husten durch eine Katzenallergie ausgelöst.

D De koning is bekender dan de tovenaar. Hij moet al een lange tijd veel hoesten. Misschien wordt de hoest veroorzaakt door een kattenallergie.

E The king is more famous than the magician. He has got a strong cough for a long while now. Maybe the cough is caused by a cat allergy.

8 G Der Bäcker ist pünktlicher als der Professor. Der hat vor einiger Zeit ein Motorrad gekauft. Mit dem Motorrad kommt man leichter durch den Stau.

D De bakker is nauwkeuriger dan de professor. Die had een tijdje geleden een motor gekocht. Met de motor kom je makkelijker door de file heen.

E The baker is more punctual than the professor. He has bought a motorcycle some time ago. With the motorcycle one can get easier through the traffic jam.

9 G Der Bauer ist reicher als der Torwart. Der hat letztes Jahr keinen einzigen Tag frei gehabt. Dieses Jahr ist ein längerer Aufenthalt in einem naheliegenden Schloss geplant.

D De boer is rijker dan de keeper. Hij had afgelopen jaar geen een dag vrij. Dit jaar is er een langer verblijf in een naastgelegen slot gepland.

E The farmer is richer than the goalkeeper. He has not had one day off last year. This year, a long stay in a closely located castle is planned.

10 G Der Jäger ist lustiger als der Tennisspieler. Der will heute Abend mit einer Kollegin Pizza essen gehen. Sie kennt ein gutes italienisches Restaurant.

D De jager is grappiger dan de tennisspeler. Die wil vanavond met een collega pizza gaan eten. Zij kent een goed Italiaans restaurant.

E The hunter is funnier than the tennis player. He wants to go out for a pizza with a colleague (fem) tonight. She knows a good Italian restaurant.

11 G Der Postbote ist intelligenter als der Fischer. Der hat vor einiger Zeit im Lotto gewonnen. Mit dem Geld konnten die Schulden für das Haus abbezahlt werden.

D De postbode is intelligenter dan de visser. Hij had een tijdje geleden de lotto gewonnen. Met het geld konden de schulden voor het huis worden afbetaald.

E The mailman is more intelligent than the fisherman. He has won the lottery some time ago. With the money, the debts on the house could be payed off.

12 G Der Vater ist schneller als der Sohn. Der will nächste Woche nach München auf das Oktoberfest fahren. Mit dem Auto wird die Fahrt ca. vier Stunden dauern.

D De vader is sneller dan de zoon. Die wil volgende week naar het Oktoberfest in München rijden. Met de auto zal de rit ca. 4 uur duren.

E The father is faster than the son. He wants to go to Munich to the Oktoberfest next week. The drive will take about four hours by car.

13 G Der Apfel ist süßer als der Pfirsich. Der wurde heute Morgen ganz frisch gepflückt. Bei Sonne kann sich der Geschmack am besten entfalten.

D De appel is zoeter dan de perzik. Hij wordt iedere morgen heel vers geplukt. Door de zon kan de smaak zich het beste ontwikkelen.

E The apple is sweeter than the peach. It was picked very freshly this morning. The taste can develop best under sunshine.

14 G Der Schrank ist schwerer als der Tisch. Der stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.

D De kast is zwaarder dan de tafel. Die is afkomstig uit een meubelwinkel in België. De sofa zal volgende week geleverd worden.

E The cupboard is heavier than the table. It originates from a furniture store in Belgium. The sofa is supposed to be delivered next week.

15 G Der Computer ist teurer als der Fernseher. Der wurde letzte Woche im Internet versteigert. Auch eine Gitarre hatte der Käufer so erwerben können.

D De computer is duurder dan de televisie. Hij werd afgelopen week geveild op internet. De koper heeft ook een gitaar weten te bemachtigen.

E The computer is more expensive than the television. It was sold at an auction on

the internet last week. The customer was able to purchase also a guitar this way.

16 G Der Kuchen ist leckerer als der Keks. Der liegt auf dem Teller im Küchenregal. Dort werden die Kinder nicht nachschauen.

D De taart is lekkerder dan de koek. Die ligt op het bord op het keukenaanrecht. Daar zullen ze kinderen niet kijken.

E The cake is more delicious than the cookie. It lies on the plate in the kitchen rack. The children won't look there.

17 G Der Mantel ist moderner als der Rock. Der ist vor zwei Wochen per Post geliefert worden. Die Briefe kamen zusammen mit dem Paket an.

D De mantel is moderner dan de rok. Hij is twee weken geleden per post geleverd. De brieven kwamen samen met het pakket aan.

E The coat is more modern than the skirt. It was delivered via mail two weeks ago. The letters arrived together with the parcel.

18 G Der Stuhl ist bequemer als der Sessel. Der gilt mittlerweile schon als Sammlerstück. Bei der letzten Auktion wurde sogar der Sattel Napoleons zu einem geringeren Preis verkauft.

D De stoel is comfortabeler dan de leunstoel. Die wordt intussen beschouwd als verzamelstuk. Bij de laatste veiling is zelfs Napoleons zetel voor een lagere prijs verkocht.

E The chair is more comfortable than the lazy chair. It counts meanwhile already as a collectible. At the last auction, even the saddle of Napoleon was sold at a lower price

19 G Der Wecker ist lauter als der Staubsauger. Der wurde letzte Woche repariert. Jetzt muss nur noch die Lampe repariert werden.

D De wekker is luider dan de stofzuiger. Hij is afgelopen week gerepareerd. Nu moet alleen nog de lamp gerepareerd worden.

E The alarm clock is louder than the vacuum cleaner. It was repaired last week. Now, only the lamp still needs to be repaired.

20 G Der Brief ist langweiliger als der Roman. Der liegt auf dem Tisch im Wohnzimmer neben der Palme. Die Palme hatte die Nachbarin aus dem Urlaub mitgebracht.

D De brief is saaier dan de roman. Die ligt op de tafel in de woonkamer naast de palmboom. De palmboom heeft de buurvrouw van vakantie meegebracht.

E *The letter is more boring than the novel. It lies on the table in the living room next to the palm. The neighbor (fem) had brought the palm from vacation.*

21 G Der Laubbaum ist älter als der Tannenbaum. Der soll nächste Woche von Köln nach Wien transportiert werden. Dazu wird ein großes Auto benötigt.

D De loofboom is ouder dan de dennenboom. Hij zal volgende week vanuit Keulen naar Wenen getransporteerd worden. Daarvoor is een grote auto nodig.

E *The broadleaf tree is older than the fir tree. It is supposed to be transported from Cologne to Vienna next week. A bigger car is needed for this.*

22 G Der Koffer ist leichter als der Rucksack. Der steht bereits fertig gepackt im Flur. Heute abend will die Krankenschwester verreisen.

D De koffer is lichter dan de rugzak. Die staat al ingepakt op de gang. Vanavond wil de verpleegster op reis gaan.

E *The suitcase is lighter than the back pack. It stands already readily packed in the hallway. The nurse is going on a journey tonight.*

23 G Der Kugelschreiber ist genauer als der Bleistift. Der wird von Anne gerne für einfache Skizzen benutzt. Zuletzt hatte Anne an dem Entwurf eines Kleides gearbeitet.

E *The ball pen is more precise than the pencil. It is fondly used for easy sketches by Anne. Lastly, Anne had worked on the draft of a dress.*

D De balpen is preciezer dan de viltstift. Hij wordt door Anne graag voor simpele schetsen gebruikt. Laatst had Anne een ontwerp voor een jurk gemaakt.

E *The ball pen is more precise than the felt tip. It is fondly used for easy sketches by Anne. Lastly, Anne had worked on the draft of a dress.*

24 G Der Saft ist frischer als der Käse. Der liegt im Einkaufskorb neben dem Gemüse. Das Gemüse soll morgen zubereitet werden.

D De jus is verser dan de kaas. Die ligt in de winkelmand naast de groenten. De groenten zullen morgen klaargemaakt worden.

E *The juice is fresher than the cheese. It lies in the shopping basket next to the vegetables. The vegetables shall be prepared tomorrow.*

Table 0.1: Experimental items of chapter 3.1 (same items in non-canonical order in chapter 3.2)

Appendix B: Forced-Choice Questionnaire

German Questionnaire, Version 1

Name:
Alter:
Geschlecht: w / m
Muttersprache/n:

Version 1 Offline-Fragebogen

Aufgabe: Die folgenden Kurzgeschichten enthalten jeweils immer ein fettgedrucktes Wort. Bitte umkreise den Satzteil, auf den sich das fettgedruckte Wort bezieht.

z.B. (Martha und Arne) sind zusammen ins Kino gegangen.
Leider kamen **sie** zu spät.

1	Die Biene ist fleißig und arbeitet jeden Tag an den Honigwaben. Abends kehrt **sie** erschöpft zum Bienennest zurück.	hf31
2	Zwei Papageien teilen sich einen großen Käfig und doch streiten sie sich manchmal. Dann rupfen **sie** sich gegenseitig Federn aus, um sich zu ärgern. Doch eigentlich mögen sie einander gern.	qr33
3	Die Tänzerin tanzt für ihr Leben gern. Oft trägt sie dann ein weißes Kleid und Ballettschuhe. Manchmal tanzt **sie** auch vor vielen Zuschauern.	ak30
4	Die Zauberin hat ihren Zauberstab immer bei sich. Und bei Vollmond füllt **sie** einen großen Kessel mit Zaubertrank.	fl25
5	Der Apfel ist süßer als der Pfirsich. **Er** wurde heute Morgen ganz frisch gepflückt. Bei Sonne kann sich der Geschmack am besten entfalten.	ui13
6	Der Kugelschreiber ist genauer als der Bleistift. **Er** wird von Anne gerne für einfache Skizzen benutzt. Zuletzt hatte Anne an dem Entwurf eines Kleides gearbeitet.	gi23
7	Die Pilotin konnte ihr Flugzeug besonders gut steuern. Ihr Chef war stolz auf sie. **Er** ehrte sie daher mit einem Titel.	af35
8	Der Jäger ist lustiger als der Tennisspieler. **Der** will heute Abend mit einer Kollegin Pizza essen gehen. Sie kennt ein gutes italienisches Restaurant.	ea10
9	Der Dackel ist kleiner als der Wolf. Die Jägerin nimmt den Dackel gerne mit	sj36

	in den Wald, denn **er** hat eine besonders gute Nase.	
10	Der Vater ist schneller als der Sohn. **Der** will nächste Woche nach München auf das Oktoberfest fahren. Mit dem Auto wird die Fahrt ca. vier Stunden dauern.	la12
11	Die Königin ist sehr mächtig, trägt eine Krone und besitzt einen goldenen Thron. **Sie** regiert das Land mit viel Geschick.	bj29
12	Die Bäuerin lebte mit ihrer Familie auf einem großen Bauernhof. Dort versorgte **sie** die Kühe und die Schafe.	ku26
13	Das Schwimmbad ist angenehmer als das Meer. Anne will heute nach der Arbeit schwimmen gehen. Dazu hat **sie** ihre Schwimmsachen mit ins Büro genommen.	fo47
14	Der Arzt ist freundlicher als der Koch. **Er** ist heute Morgen recht früh aufgestanden. Nur die Bäckerin war noch früher wach als er.	la01
15	Der Postbote ist intelligenter als der Fischer. **Er** hat vor einiger Zeit im Lotto gewonnen. Mit dem Geld konnten die Schulden für das Haus abbezahlt werden.	ha11
16	Das Kartenspiel ist interessanter als das Geschenk. Klaus ist ein talentierter Kartenspieler. Letztes Jahr hat **er** bei der Meisterschaft den ersten Platz gemacht und eine Krone gewonnen.	ld54
17	Der Schrank ist schwerer als der Tisch. **Der** stammt aus einem Möbelgeschäft in Belgien. Das Sofa soll nächste Woche geliefert werden.	fi14
18	Der Matrose ist netter als der Kapitän. **Er** will dieses Wochenende in den Zoo fahren. Der Zoo ist bekannt, viele Papageien zu haben.	pa03
19	Die Polizistin fährt meistens den Streifenwagen. Aber letztens mußte **sie** eine Haustür aufbrechen, um jemanden zu retten.	dp49
20	Die Bärenmutter versorgt ihre Jungen gut. Manchmal gibt **sie** ihnen leckere Honigwaben. Dafür muß sie aber erst die Bienen verscheuchen.	rm43
21	Der Koffer ist leichter als der Rucksack. **Der** steht bereits fertig gepackt im Flur. Heute abend will die Krankenschwester verreisen.	ti22
22	Die Prinzessin hatte ein wunderschönes blaues Kleid. **Sie** wohnte in einem großen Schloß.	ik34
23	Der Polizist ist schlauer als der Gärtner. **Der** ist nun schon zum zehnten Mal in eine neue Wohnung gezogen. Die letzte Wohnung war von Mäusen befallen gewesen.	ta02
24	Die Schauspielerin ist jung und berühmt und steht oft vor der Kamera. Manchmal wird **sie** von Journalisten interviewt. Aber das mag sie nicht besonders.	ul27

25	Der Computer ist teurer als der Fernseher. **Er** wurde letzte Woche im Internet versteigert. Auch eine Gitarre hatte der Käufer so erwerben können.	mi15
26	Die Ärztin steht jeden Morgen zeitig auf und zieht sich ihren weißen Kittel an. **Sie** behandelt ihre Patienten mit viel Geduld. Doch abends ist sie oft erschöpft.	jh52
27	Der Saft ist frischer als der Käse. **Der** liegt im Einkaufskorb neben dem Gemüse. Das Gemüse soll morgen zubereitet werden.	ii24
28	Das Futter ist nahrhafter als das Brötchen. Trotzdem will die Ente lieber das Brötchen fressen. Passanten hatten **es** in den Teich geworfen.	pl44
29	Das Schaf ist kuscheliger als das Schwein. Eier kann **es** leider nicht legen.	mp32
30	Die Köchin liebte frisches Gemüse und konnte **es** sehr gut zubereiten. Gerade rührte sie in einem großen Kochtopf und schmeckte es ab. Ihre Gäste waren später begeistert.	kt28
31	Der Löwe ist stärker als der Tiger. Die Königin benutzt **ihn** deshalb gerne als Symbol.	ds45
32	Die Bäckerin backte die besten Brötchen der Stadt und hatte sich damit sehr beliebt gemacht. Jeden Morgen holte sie die frischen Brötchen aus ihrem großen Ofen. **Diese** dufteten in der ganzen Strasse.	au53
33	Die Hasen haben zu Ostern besonders viel zu tun. **Sie** färben Eier und legen sie in kleine Nester.	uh42
34	Der König ist bekannter als der Zauberer. **Er** hat seit längerer Zeit starken Husten. Vielleicht wird der Husten durch eine Katzenallergie ausgelöst.	qa07
35	Der Bauer ist reicher als der Torwart. **Er** hat letztes Jahr keinen einzigen Tag frei gehabt. Dieses Jahr ist ein längerer Aufenthalt in einem naheliegenden Schloss geplant.	ja09
36	Der Brief ist langweiliger als der Roman. **Der** liegt auf dem Tisch im Wohnzimmer neben der Palme. Die Palme hatte die Nachbarin aus dem Urlaub mitgebracht.	oi20
37	Die Schlange ist hinterhältiger als die Katze. Die Tänzerin behängt sich gerne mit der Schlange während **sie** auftritt.	nd40
38	Die Schule, in der die Lehrerin arbeitet, ist klein und beschaulich und liegt in einem Bergdorf. Die Lehrerin ist bei fast allen Schülern beliebt. **Sie** kann ihnen viel beibringen.	nj37
39	Im Dschungel leben viele Tiere. Darunter auch ein Elefant und ein Tiger. Überraschenderweise verstehen **sie** sich gut.	hm57
40	Der Feuerwehrmann ist dümmer als der Boxer. **Er** will nächste Woche mit einer Reisegruppe in den Urlaub fahren. Das Meer ist zu dieser Jahreszeit	ra05

	jedoch noch ziemlich kalt.	
41	Die alte Frau ging zu ihrem Häuschen. Es sah sehr schön aus und die Eingangstüre war neu. Sie hatte **sie** gerade erst gekauft.	iq44
42	Der Laubbaum ist älter als der Tannenbaum. **Er** soll nächste Woche von Köln nach Wien transportiert werden. Dazu wird ein großes Auto benötigt.	si21
43	Der Stuhl ist bequemer als der Sessel. **Der** gilt mittlerweile schon als Sammlerstück. Bei der letzten Auktion wurde sogar der Sattel Napoleons zu einem geringeren Preis verkauft.	oi18
44	Die alte Gärtnerin hatte ihr ganzes Leben den Pflanzen gewidmet. Jeden Tag nahm **sie** Hake und Spaten in die Hand und pflegte die Gärten. Mehr brauchte sie nicht, sie war zufrieden.	nt27
45	Die Maus ist flinker als die Katze. Ganz schnell ist **sie** in einer der Mülltonnen am Straßenrand verschwunden.	eu36
46	Die Reporterin sitzt die meiste Zeit an ihrem Schreibtisch, aber manchmal interviewt **sie** auch berühmte Personen vor der Kamera. Deshalb hat sie auch viel Spaß an ihrem Beruf.	gf54
47	Der Wecker ist lauter als der Staubsauger. **Er** wurde letzte Woche repariert. Jetzt muss nur noch die Lampe repariert werden.	si19
48	Die Klavierspielerin ist sehr talentiert. **Sie** spielt die Noten rauf und runter. Und ihr Klavier ist ihr liebstes Stück geworden.	bq33
49	Die Sportlerin hatte gerade ihre Topform erreicht. Doch dann brach **sie** ihr Bein während des Trainings. Ihr Trainer war darüber sehr enttäuscht.	qg29
50	Der Mantel ist moderner als der Rock. **Er** ist vor zwei Wochen per Post geliefert worden. Die Briefe kamen zusammen mit dem Paket an.	ki17
51	Die Spieler der Fussballmannschaft tragen blaue Trikots und treffen das Tor immer. **Sie** werden sehr erfolgreich sein.	rb46
52	Die Jägerin beobachtet gern die Rehe im Wald. Ihr Gewehr hat sie auch oft dabei, aber sie traut sich nicht, **es** einzusetzen.	ge52
53	Das Motorrad ist schneller als das Fahrrad. Trotzdem fährt Boris lieber mit dem Fahrrad in die Berge. So kann **er** sich sportlich betätigen.	ef48
54	Die Zwerge tragen alle große Mützen. Wenn die Prinzessin **sie** besucht, sind alle ganz aufgeregt.	mr31
55	Das Auto ist gefährlicher als das Flugzeug. Trotzdem hat die Tänzerin Angst vor dem Fliegen. Deshalb nimmt **sie** beruhigende Medikamente vor dem Abflug.	tk39
56	Der Lehrer ist glücklicher als der Musiker. **Der** wollte gestern Abend ins Kino gehen. Leider waren die Plätze schon ausverkauft, sodass er stattdessen ins	da04

	Schwimmbad ging.	
57	Der Bäcker ist pünktlicher als der Professor. **Der** hat vor einiger Zeit ein Motorrad gekauft. Mit dem Motorrad kommt man leichter durch den Stau.	pa08
58	Der Clown ist frecher als der Indianer. **Der** will für nächste Woche eine große Feier organisieren. Dort soll eine Klavierspielerin auftreten.	oa06
59	Der Kuchen ist leckerer als der Keks. **Der** liegt auf dem Teller im Küchenregal. Dort werden die Kinder nicht nachschauen.	bi16
60	Das Kind hatte einen guten Freund. Im Sommer gingen **sie** oft zusammen Eis essen. Sie lachten viel und hatten Spaß.	jl43

Table 0.2: German off-line questionnaire of chapter 3.1, Version 1

Appendix C: Sound files parameters of German and Dutch recordings

		Dutch Recordings				German Recordings			
Item	**Pronoun**	**Duration of wav1 (ms)**	**Sentence**	**Amount of syllables**	**Syllables per sec**	**Duration of wav1 (ms)**	**Sentence**	**Amount of syllables**	**Syllables per sec**
1	d-pronoun	2600	De arts is vriendelijker dan de kok.	10	3.85	2210	Der Arzt ist freundlicher als der Koch.	9	4.07
2	d-pronoun	2928	De politieagent is slimmer dan de tuinman.	13	4.44	2557	Der Polizist ist schlauer als der Gärtner.	11	4.30
3	d-pronoun	2777	De matroos is aardiger dan de kapitein.	12	4.32	2587	Der Matrose ist netter als der Kapitän.	12	4.64
4	d-pronoun	3060	De leraar is gelukkiger dan de musicus.	13	4.25	2439	Der Lehrer ist glücklicher als der Musiker.	12	4.92
5	d-pronoun	2493	De brandweerman is dommer dan de bokser.	11	4.41	2391	Der Feuerwehrmann ist dümmer als der Boxer.	12	5.02
6	d-pronoun	2674	De clown is brutaler dan de indiaan.	11	4.11	2210	Der Clown ist frecher als der Indianer.	10	4.52
7	d-pronoun	2673	De koning is bekender dan de tovenaar.	12	4.49	2308	Der König ist bekannter als der Zauberer.	12	5.20
8	d-pronoun	2866	De bakker is nauwkeuriger dan de professor.	13	4.54	2521	Der Bäcker ist pünktlicher als der Professor.	12	4.76
9	d-pronoun	1877	De boer is rijker dan de keeper.	9	4.79	2341	Der Bauer ist reicher als der Torwart.	10	4.27
10	d-pronoun	2925	De jager is grappiger dan de tennisspeler.	13	4.44	2423	Der Jäger ist lustiger als der Tennisspieler.	13	5.37

11	d-pronoun	2806	De postbode is intelligenter dan de visser.	14	4.99	2748	Der Postbote ist intelligenter als der Fischer.	14	5.09
12	d-pronoun	2201	De vader is sneller dan de zoon.	9	4.09	2225	Der Vater ist schneller als der Sohn.	9	4.04
13	d-pronoun	2002	De appel is zoeter dan de perzik.	10	5.00	2255	Der Apfel ist süßer als der Pfirsich.	10	4.43
14	d-pronoun	2198	De kast is zwaarder dan de tafel.	9	4.09	2187	Der Schrank ist schwerer als der Tisch.	8	3.66
15	d-pronoun	2950	De computer is duurder dan de televisie.	13	4.41	2507	Der Computer ist teurer als der Fernseher.	12	4.79
16	d-pronoun	2034	De taart is lekkerder dan de koek.	9	4.42	2439	Der Kuchen ist leckerer als der Keks.	10	4.10
17	d-pronoun	2218	De mantel is moderner dan de rok.	10	4.51	2053	Der Mantel ist moderner als der Rock.	10	4.87
18	d-pronoun	3089	De stoel is comfortabeler dan de leunstoel.	11	3.56	2221	Der Stuhl ist bequemer als der Sessel.	10	4.50
19	d-pronoun	2469	De wekker is luider dan de stofzuiger.	11	4.46	2389	Der Wecker ist lauter als der Staubsauger.	11	4.60
20	d-pronoun	2278	De brief is saaier dan de roman.	9	3.95	2591	Der Brief ist langweiliger als der Roman.	11	4.25
21	d-pronoun	2392	De loofboom is ouder dan de dennenboom.	11	4.60	2691	Der Laubbaum ist älter als der Tannenbaum.	11	4.09
22	d-pronoun	2075	De koffer is lichter dan de rugzak.	10	4.82	2221	Der Koffer ist leichter als der Rucksack.	10	4.50
23	d-pronoun	3063	De balpen is preciezer dan de viltstift.	11	3.59	3096	Der Kugelschreiber ist genauer als der Bleistift.	13	4.20

24	d-pronoun	2236	De jus is verser dan de kaas.	8	3.58	2218	Der Saft ist frischer als der Käse.	9	4.06
1	personal pronoun	2221	De arts is vriendelijker dan de kok.	10	4.50	2194	Der Arzt ist freundlicher als der Koch.	9	4.10
2	personal pronoun	3066	De politieagent is slimmer dan de tuinman.	13	4.24	2388	Der Polizist ist schlauer als der Gärtner.	11	4.61
3	personal pronoun	2667	De matroos is aardiger dan de kapitein.	12	4.50	2504	Der Matrose ist netter als der Kapitän.	12	4.79
4	personal pronoun	2616	De leraar is gelukkiger dan de musicus.	13	4.97	2391	Der Lehrer ist glücklicher als der Musiker.	12	5.02
5	personal pronoun	2589	De brandweerman is dommer dan de bokser.	11	4.25	2685	Der Feuerwehrmann ist dümmer als der Boxer.	12	4.47
6	personal pronoun	2609	De clown is brutaler dan de indiaan.	11	4.22	2325	Der Clown ist frecher als der Indianer.	10	4.30
7	personal pronoun	2451	De koning is bekender dan de tovenaar.	12	4.90	2179	Der König ist bekannter als der Zauberer.	12	5.51
8	personal pronoun	2578	De bakker is nauwkeuriger dan de professor.	13	5.04	2356	Der Bäcker ist pünktlicher als der Professor.	12	5.09
9	personal pronoun	1893	De boer is rijker dan de keeper.	9	4.75	2325	Der Bauer ist reicher als der Torwart.	10	4.30
10	personal pronoun	2652	De jager is grappiger dan de tennisspeler.	13	4.90	2537	Der Jäger ist lustiger als der Tennisspieler.	13	5.12
11	personal pronoun	2687	De postbode is intelligenter dan de visser.	14	5.21	2798	Der Postbote ist intelligenter als der Fischer.	14	5.00
12	personal pronoun	1830	De vader is sneller dan de zoon.	9	4.92	2081	Der Vater ist schneller als der Sohn.	9	4.32

13	personal pronoun	2379	De appel is zoeter dan de perzik.	10	4.20	2204	Der Apfel ist süßer als der Pfirsich.	10	4.54
14	personal pronoun	2213	De kast is zwaarder dan de tafel.	9	4.07	2120	Der Schrank ist schwerer als der Tisch.	8	3.77
15	personal pronoun	2894	De computer is duurder dan de televisie.	13	4.49	2692	Der Computer ist teurer als der Fernseher.	12	4.46
16	personal pronoun	2303	De taart is lekkerder dan de koek.	9	3.91	2322	Der Kuchen ist leckerer als der Keks.	10	4.31
17	personal pronoun	2291	De mantel is moderner dan de rok.	10	4.36	2272	Der Mantel ist moderner als der Rock.	10	4.40
18	personal pronoun	2681	De stoel is comfortabeler dan de leunstoel.	11	4.10	2152	Der Stuhl ist bequemer als der Sessel.	10	4.65
19	personal pronoun	2429	De wekker is luider dan de stofzuiger.	11	4.53	2320	Der Wecker ist lauter als der Staubsauger.	11	4.74
20	personal pronoun	2242	De brief is saaier dan de roman.	9	4.01	2458	Der Brief ist langweiliger als der Roman.	11	4.48
21	personal pronoun	2263	De loofboom is ouder dan de dennenboom.	11	4.86	2573	Der Laubbaum ist älter als der Tannenbaum.	11	4.28
22	personal pronoun	2146	De koffer is lichter dan de rugzak.	10	4.66	2389	Der Koffer ist leichter als der Rucksack.	10	4.19
23	personal pronoun	2757	De balpen is preciezer dan de viltstift.	11	3.99	2860	Der Kugelschreiber ist genauer als der Bleistift.	13	4.55
24	personal pronoun	2203	De jus is verser dan de kaas.	8	3.63	2287	Der Saft ist frischer als der Käse.	9	3.94
M		**2490.5**		**10.9167**	**4.39**	**2400.8333**		**10.875**	**4.52**
SD		**485.1634014**		**2.233366**	**0.74**	**401.2016312**		**2.10491**	**0.76**

Table 0.3: Overview of sound file parameters of the German and Dutch recordings for the experiments in chapter 3.1 and 4.1.

Appendix D: Loglikelihood Tests – Fit of the models

Chapter 3.1

L1 Germans

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	4191.675	4222.87755	-2089.8375				
		Interaction	7	4192.75961	4229.16259	-2089.37981	0.91538208	1	0.33869042	
2	200-400	Simple	6	6824.89237	6856.09492	-3406.44619				
		Interaction	7	6826.42095	6862.82392	-3406.21048	0.47142309	1	0.49233321	
3	400-600	Simple	6	7211.1191	7242.32165	-3599.55955				
		Interaction	7	7211.24232	7247.64529	-3598.62116	1.87678617	1	0.17069987	
4	600-800	Simple	6	7300.94029	7332.14284	-3644.47015				
		Interaction	7	7298.80473	7335.2077	-3642.40237	4.13556211	1	0.04199112	*
5	800-1000	Simple	6	7398.81413	7430.01668	-3693.40707				
		Interaction	7	7392.92426	7429.32723	-3689.46213	7.88987642	1	0.00497123	**
6	1000-1200	Simple	6	7460.97653	7492.17908	-3724.48826				
		Interaction	7	7444.09746	7480.50044	-3715.04873	18.8790648	1	1.39E-05	***
7	1200-1400	Simple	6	7478.07708	7509.27963	-3733.03854				
		Interaction	7	7468.34127	7504.74425	-3727.17064	11.7358061	1	0.00061309	***
8	1400-1600	Simple	6	7500.66253	7531.86508	-3744.33126				
		Interaction	7	7485.93895	7522.34192	-3735.96947	16.7235789	1	4.32E-05	***
9	1600-1800	Simple	6	7463.52408	7494.72663	-3725.76204				
		Interaction	7	7456.14897	7492.55194	-3721.07449	9.37510529	1	0.00219952	**
10	1800-2000	Simple	6	7479.2118	7510.41435	-3733.6059				
		Interaction	7	7472.13075	7508.53372	-3729.06537	9.08104676	1	0.00258272	**

Table 0.4: Loglikelihood test results evaluating the goodness-of-fit of the interaction model over the simple model. **Note:** † p<.1; * p<.05; ** p<.01; *** p<.001.

L1 Dutch

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	4357.74234	4388.95383	-2172.87117				
		Interaction	7	4358.97656	4395.38997	-2172.48828	0.76578054	1	0.38152555	
2	200-400	Simple	6	6872.57688	6903.78838	-3430.28844				
		Interaction	7	6873.58218	6909.99559	-3429.79109	0.99470788	1	0.31859444	
3	400-600	Simple	6	7224.13173	7255.34323	-3606.06587				
		Interaction	7	7223.4033	7259.81672	-3604.70165	2.72843385	1	0.09857591	†
4	600-800	Simple	6	7401.34398	7432.55548	-3694.67199				
		Interaction	7	7397.6386	7434.05202	-3691.8193	5.70537724	1	0.01691302	*
5	800-1000	Simple	6	7449.47791	7480.68941	-3718.73895				
		Interaction	7	7444.78278	7481.19619	-3715.39139	6.69513197	1	0.00966765	**
6	1000-1200	Simple	6	7513.84411	7545.0556	-3750.92205				
		Interaction	7	7506.08632	7542.49973	-3746.04316	9.75778988	1	0.00178565	**
7	1200-1400	Simple	6	7504.71255	7535.92405	-3746.35628				
		Interaction	7	7488.90998	7525.32339	-3737.45499	17.802578	1	0.0000245	***
8	1400-1600	Simple	6	7470.74614	7501.95764	-3729.37307				
		Interaction	7	7460.97647	7497.38989	-3723.48824	11.7696689	1	0.00060204	***
9	1600-1800	Simple	6	7443.29999	7474.51149	-3715.64999				
		Interaction	7	7437.75971	7474.17313	-3711.87986	7.54027549	1	0.00603348	**
10	1800-2000	Simple	6	7465.69695	7496.90844	-3726.84847				
		Interaction	7	7448.9192	7485.33262	-3717.4596	18.7777425	1	0.0000147	***

Table 0.5: *Loglikelihood test results evaluating the goodness-of-fit of the interaction model over the simple model.* **Note:** † p<.1; * p<.05; ** p<.01; *** p<.001.

Chapter 3.2

L1 Germans

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	5935.25239	5967.04306	-2961.6262				
		Interaction	7	5937.24446	5974.33358	-2961.62223	0.00793086	1	0.92903792	
2	200-400	Simple	6	7533.78083	7565.5715	-3760.89041				
		Interaction	7	7535.74629	7572.8354	-3760.87314	0.03454155	1	0.8525596	
3	400-600	Simple	6	7994.73493	8026.5256	-3991.36746				
		Interaction	7	7996.40817	8033.49729	-3991.20409	0.32675638	1	0.56757532	
4	600-800	Simple	6	8138.1118	8169.90247	-4063.0559				
		Interaction	7	8139.21608	8176.30519	-4062.60804	0.89572432	1	0.34393077	
5	800-1000	Simple	6	8240.91022	8272.7009	-4114.45511				
		Interaction	7	8240.81903	8277.90814	-4113.40951	2.09119841	1	0.14814982	
6	1000-1200	Simple	6	8213.35747	8245.14814	-4100.67873				
		Interaction	7	8214.25439	8251.3435	-4100.12719	1.10307864	1	0.29359146	
7	1200-1400	Simple	6	8231.29154	8263.08221	-4109.64577				
		Interaction	7	8233.06256	8270.15167	-4109.53128	0.22898774	1	0.63227542	
8	1400-1600	Simple	6	8235.23164	8267.02231	-4111.61582				
		Interaction	7	8237.22092	8274.31004	-4111.61046	0.01071904	1	0.91754012	
9	1600-1800	Simple	6	8241.7311	8273.52177	-4114.86555				
		Interaction	7	8243.67114	8280.76025	-4114.83557	0.05995796	1	0.80656239	
10	1800-2000	Simple	6	8225.90427	8257.69494	-4106.95213				
		Interaction	7	8227.45707	8264.54618	-4106.72853	0.44719978	1	0.50366774	

Table 0.6: Loglikelihood test results evaluating the goodness-of-fit of the interaction model over the simple model. **Note:** † p<.1; * p<.05; ** p<.01; *** p<.001.

L1 Dutch

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	3136.20298	3166.46718	-1562.10149				
		Interaction	7	3137.67528	3172.98351	-1561.83764	0.52769905	1	0.46757584	
2	200-400	Simple	6	5995.70037	6025.96457	-2991.85019				
		Interaction	7	5997.29165	6032.59988	-2991.64582	0.4087227	1	0.52261853	
3	400-600	Simple	6	6201.30649	6231.57069	-3094.65325				
		Interaction	7	6203.29919	6238.60742	-3094.6496	0.00729984	1	0.93191231	
4	600-800	Simple	6	6253.28629	6283.55049	-3120.64314				
		Interaction	7	6254.63389	6289.94212	-3120.31694	0.65240248	1	0.41925504	
5	800-1000	Simple	6	6287.23901	6317.50321	-3137.61951				
		Interaction	7	6289.16899	6324.47722	-3137.58449	0.07002398	1	0.79130187	
6	1000-1200	Simple	6	6275.2153	6305.47949	-3131.60765				
		Interaction	7	6276.94549	6312.25372	-3131.47274	0.26980994	1	0.60345929	
7	1200-1400	Simple	6	6282.80923	6313.07343	-3135.40462				
		Interaction	7	6283.84886	6319.15709	-3134.92443	0.96036879	1	0.32709398	
8	1400-1600	Simple	6	6279.70331	6309.96751	-3133.85166				
		Interaction	7	6281.49167	6316.7999	-3133.74584	0.2116405	1	0.64548462	
9	1600-1800	Simple	6	6259.00668	6289.27088	-3123.50334				
		Interaction	7	6260.85955	6296.16778	-3123.42977	0.14713448	1	0.70128893	
10	1800-2000	Simple	6	6284.10562	6314.36982	-3136.05281				
		Interaction	7	6286.01912	6321.32736	-3136.00956	0.08649379	1	0.76868281	

Table 0.7: Loglikelihood test results evaluating the goodness-of-fit of the interaction model over the simple model. **Note:** † p<.1; * p<.05; ** p<.01; *** p<.001.

Chapter 3.3

L1 Germans

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	4191.675	4222.87755	-2089.8375				
		Animacy	11	4199.53262	4256.7373	-2088.76631	2.14237455	5	0.82911472	
2	200-400	Simple	6	6824.89237	6856.09492	-3406.44619				
		Animacy	11	6832.45609	6889.66076	-3405.22805	2.43628235	5	0.786059	
3	400-600	Simple	6	7211.1191	7242.32165	-3599.55955				
		Animacy	11	7217.68578	7274.89045	-3597.84289	3.43332571	5	0.63349943	
4	600-800	Interaction	7	7298.80473	7335.2077	-3642.40237				
		Animacy	11	7298.52242	7355.72709	-3638.26121	8.28231337	4	0.08176724	†
5	800-1000	Interaction	7	7392.92426	7429.32723	-3689.46213				
		Animacy	11	7396.63639	7453.84107	-3687.3182	4.2878653	4	0.36844708	
6	1000-1200	Interaction	7	7444.09746	7480.50044	-3715.04873				
		Animacy	11	7448.63905	7505.84372	-3713.31952	3.45841317	4	0.48422982	
7	1200-1400	Interaction	7	7468.34127	7504.74425	-3727.17064				
		Animacy	11	7467.04144	7524.24611	-3722.52072	9.29983528	4	0.05402671	†
8	1400-1600	Interaction	7	7485.93895	7522.34192	-3735.96947				
		Animacy	11	7482.05968	7539.26435	-3730.02984	11.8792667	4	0.01827202	*
9	1600-1800	Interaction	7	7456.14897	7492.55194	-3721.07449				
		Animacy	11	7460.7689	7517.97357	-3719.38445	3.3800722	4	0.49634626	
10	1800-2000	Interaction	7	7472.13075	7508.53372	-3729.06537				
		Animacy	11	7477.3763	7534.58097	-3727.68815	2.75444973	4	0.59972079	

Table 0.8: Loglikelihood test results evaluating the goodness-of-fit of the animacy model over the interaction model or the simple model.
Note: † p<.1; * p<.05; ** p<.01; *** p<.001.

L1 Dutch

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	4357.74234	4388.95383	-2172.87117				
		Animacy	11	4363.39598	4420.61706	-2170.69799	4.3463555	5	0.5006998	
2	200-400	Simple	6	6872.57688	6903.78838	-3430.28844				
		Animacy	11	6875.1033	6932.32438	-3426.55165	7.47357842	5	0.18773378	
3	400-600	Simple	6	7224.13173	7255.34323	-3606.06587				
		Animacy	11	7220.8492	7278.07028	-3599.4246	13.282538	5	0.02087015	*
4	600-800	Interaction	7	7397.6386	7434.05202	-3691.8193				
		Animacy	11	7399.87027	7457.09135	-3688.93513	5.76833897	4	0.21712972	
5	800-1000	Interaction	7	7444.78278	7481.19619	-3715.39139				
		Animacy	11	7448.93474	7506.15582	-3713.46737	3.84803263	4	0.42696288	
6	1000-1200	Interaction	7	7506.08632	7542.49973	-3746.04316				
		Animacy	11	7511.06709	7568.28817	-3744.53355	3.01922605	4	0.55461314	
7	1200-1400	Interaction	7	7488.90998	7525.32339	-3737.45499				
		Animacy	11	7493.71361	7550.93469	-3735.8568	3.19636711	4	0.52551792	
8	1400-1600	Interaction	7	7460.97647	7497.38989	-3723.48824				
		Animacy	11	7466.03556	7523.25664	-3722.01778	2.94091288	4	0.56776169	
9	1600-1800	Interaction	7	7437.75971	7474.17313	-3711.87986				
		Animacy	11	7440.89358	7498.11466	-3709.44679	4.86612961	4	0.30131127	
10	1800-2000	Interaction	7	7448.9192	7485.33262	-3717.4596				
		Animacy	11	7448.23444	7505.45552	-3713.11722	8.68476476	4	0.06948039	†

Table 0.9: Loglikelihood test results evaluating the goodness-of-fit of the animacy model over the interaction model or the simple model.
Note: † p<.1; * p<.05; ** p<.01; *** p<.001.

Chapter 3.4

L1 Germans

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	4191.675	4222.87755	-2089.8375				
		Off-line	11	4196.27275	4253.47743	-2087.13638	5.40224478	5	0.36878419	
2	200-400	Simple	6	6824.89237	6856.09492	-3406.44619				
		Off-line	11	6828.65636	6885.86104	-3403.32818	6.23600926	5	0.28392688	
3	400-600	Interaction	6	7211.1191	7242.32165	-3599.55955				
		Simple	11	7216.77763	7273.9823	-3597.38881	4.34147581	5	0.50136928	
4	600-800	Interaction	7	7298.80473	7335.2077	-3642.40237				
		Off-line	11	7304.2183	7361.42297	-3641.10915	2.58642942	4	0.62922897	
5	800-1000	Interaction	7	7392.92426	7429.32723	-3689.46213				
		Off-line	11	7398.56471	7455.76938	-3688.28235	2.35955017	4	0.6699492	
6	1000-1200	Interaction	7	7444.09746	7480.50044	-3715.04873				
		Off-line	11	7449.02842	7506.2331	-3713.51421	3.06903911	4	0.54633909	
7	1200-1400	Interaction	7	7468.34127	7504.74425	-3727.17064				
		Off-line	11	7473.06552	7530.27019	-3725.53276	3.27575776	4	0.51278238	
8	1400-1600	Interaction	7	7485.93895	7522.34192	-3735.96947				
		Off-line	11	7491.74909	7548.95377	-3734.87455	2.18985355	4	0.7008873	
9	1600-1800	Interaction	7	7456.14897	7492.55194	-3721.07449				
		Off-line	11	7460.92764	7518.13232	-3719.46382	3.22132608	4	0.52149332	
10	1800-2000	Interaction	7	7472.13075	7508.53372	-3729.06537				
		Off-line	11	7473.04707	7530.25174	-3725.52353	7.08368031	4	0.13153163	

Table 0.10: Loglikelihood test results evaluating the goodness-of-fit of the off-line model over the interaction model or the simple model.
Note: † p<.1; * p<.05; ** p<.01; *** p<.001.

L1 Dutch

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Interaction	6	4357.74234	4388.95383	-2172.87117				
		Off-line	11	4360.62011	4417.84119	-2169.31006	7.12222121	5	0.21170811	
2	200-400	Interaction	6	6872.57688	6903.78838	-3430.28844				
		Off-line	11	6880.65893	6937.88001	-3429.32946	1.91795376	5	0.86037727	
3	400-600	Interaction	6	7224.13173	7255.34323	-3606.06587				
		Off-line	11	7230.69585	7287.91693	-3604.34793	3.43588023	5	0.63311119	
4	600-800	Interaction	7	7397.6386	7434.05202	-3691.8193				
		Off-line	11	7403.02175	7460.24283	-3690.51088	2.61685089	4	0.62384098	
5	800-1000	Interaction	7	7444.78278	7481.19619	-3715.39139				
		Off-line	11	7447.22081	7504.44189	-3712.61041	5.56196402	4	0.23433625	
6	1000-1200	Interaction	7	7506.08632	7542.49973	-3746.04316				
		Off-line	11	7506.87095	7564.09203	-3742.43548	7.21536384	4	0.12493558	
7	1200-1400	Interaction	7	7488.90998	7525.32339	-3737.45499				
		Off-line	11	7487.87171	7545.09279	-3732.93585	9.03826835	4	0.06015004	†
8	1400-1600	Interaction	7	7460.97647	7497.38989	-3723.48824				
		Off-line	11	7464.14063	7521.36171	-3721.07031	4.83584414	4	0.30455936	
9	1600-1800	Interaction	7	7437.75971	7474.17313	-3711.87986				
		Off-line	11	7442.43586	7499.65694	-3710.21793	3.32385023	4	0.50516229	
10	1800-2000	Interaction	7	7448.9192	7485.33262	-3717.4596				
		Off-line	11	7450.02236	7507.24343	-3714.01118	6.89684777	4	0.1414408	

Table 0.11: Loglikelihood test results evaluating the goodness-of-fit of the animacy model over the interaction model or the simple model.
Note: † p<.1; * p<.05; ** p<.01; *** p<.001.

Chapter 4.1

L2 Dutch learners of German

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	3961.46769	3992.83803	-1974.73385				
		Interaction	7	3963.04167	3999.64039	-1974.52084	0.42602	1	0.51394837	
2	200-400	Simple	6	7075.7434	7107.11373	-3531.8717				
		Interaction	7	7076.75503	7113.35375	-3531.37751	0.98836931	1	0.32014126	
3	400-600	Simple	6	7505.2118	7536.58213	-3746.6059				
		Interaction	7	7505.55115	7542.14987	-3745.77557	1.66065335	1	0.19751513	
4	600-800	Simple	6	7613.19782	7644.56815	-3800.59891				
		Interaction	7	7612.11459	7648.71331	-3799.0573	3.08322458	1	0.07910355	†
5	800-1000	Simple	6	7690.12945	7721.49979	-3839.06473				
		Interaction	7	7690.56657	7727.16528	-3838.28328	1.56288964	1	0.21124262	
6	1000-1200	Simple	6	7731.98848	7763.35881	-3859.99424				
		Interaction	7	7730.61381	7767.21253	-3858.3069	3.37466795	1	0.06620592	†
7	1200-1400	Simple	6	7770.60319	7801.97352	-3879.3016				
		Interaction	7	7767.69743	7804.29615	-3876.84871	4.90576212	1	0.02676724	*
8	1400-1600	Simple	6	7749.93656	7781.30689	-3868.96828				
		Interaction	7	7745.8541	7782.45282	-3865.92705	6.08246187	1	0.01365304	*
9	1600-1800	Simple	6	7718.2904	7749.66073	-3853.1452				
		Interaction	7	7715.863	7752.46172	-3850.9315	4.42739875	1	0.03536637	*
10	1800-2000	Simple	6	7677.61	7708.98033	-3832.805				
		Interaction	7	7677.36983	7713.96855	-3831.68492	2.24016514	1	0.13446662	

Table 0.12: Loglikelihood test results evaluating the goodness-of-fit of the interaction model over the simple model. **Note:** † p<.1; * p<.05; ** p<.01; *** p<.001.

L2 German learners of Dutch

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	4788.4441	4820.44224	-2388.22205				
		Interaction	7	4790.14284	4827.474	-2388.07142	0.30126889	1	0.58308803	
2	200-400	Simple	6	7816.72642	7848.72456	-3902.36321				
		Interaction	7	7818.48159	7855.81276	-3902.2408	0.24483053	1	0.62073882	
3	400-600	Simple	6	8056.87967	8088.8778	-4022.43983				
		Interaction	7	8056.37261	8093.70377	-4021.1863	2.50706173	1	0.11333707	
4	600-800	Simple	6	8206.02844	8238.02658	-4097.01422				
		Interaction	7	8203.27381	8240.60497	-4094.63691	4.75463217	1	0.02921954	*
5	800-1000	Simple	6	8364.0637	8396.06184	-4176.03185				
		Interaction	7	8359.90701	8397.23817	-4172.9535	6.15669719	1	0.01309154	*
6	1000-1200	Simple	6	8443.34034	8475.33847	-4215.67017				
		Interaction	7	8437.09566	8474.42682	-4211.54783	8.24467481	1	0.00408717	**
7	1200-1400	Simple	6	8453.03322	8485.03136	-4220.51661				
		Interaction	7	8439.7676	8477.09876	-4212.8838	15.2656256	1	0.0000934	***
8	1400-1600	Simple	6	8497.44944	8529.44758	-4242.72472				
		Interaction	7	8479.99932	8517.33048	-4232.99966	19.4501283	1	0.0000103	***
9	1600-1800	Simple	6	8464.41336	8496.41149	-4226.20668				
		Interaction	7	8446.63303	8483.96419	-4216.31652	19.7803224	1	0.00000869	***
10	1800-2000	Simple	6	8459.34194	8491.34008	-4223.67097				
		Interaction	7	8434.26022	8471.59139	-4210.13011	27.0817129	1	1.95E-07	***

Table 0.13: *Loglikelihood test results evaluating the goodness-of-fit of the interaction model over the simple model.* **Note:** † p<.1; * p<.05; ** p<.01; *** p<.001.

Chapter 4.2

L2 Dutch learners of German

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	3608.11339	3638.55301	-1798.05669				
		Interaction	7	3610.07829	3645.59118	-1798.03915	0.03509574	1	0.85139519	
2	200-400	Simple	6	6179.74165	6210.18127	-3083.87083				
		Interaction	7	6181.38596	6216.89885	-3083.69298	0.35568799	1	0.55091082	
3	400-600	Simple	6	6466.27131	6496.71093	-3227.13566				
		Interaction	7	6466.20919	6501.72208	-3226.1046	2.06212134	1	0.15099983	
4	600-800	Simple	6	6487.41757	6517.85719	-3237.70878				
		Interaction	7	6489.16685	6524.67974	-3237.58343	0.25071425	1	0.6165726	
5	800-1000	Simple	6	6566.63377	6597.07339	-3277.31689				
		Interaction	7	6564.13118	6599.64407	-3275.06559	4.50258692	1	0.03384362	*
6	1000-1200	Simple	6	6610.68507	6641.12469	-3299.34254				
		Interaction	7	6607.99558	6643.50847	-3296.99779	4.68949059	1	0.03034764	*
7	1200-1400	Simple	6	6628.78153	6659.22115	-3308.39076				
		Interaction	7	6629.14166	6664.65454	-3307.57083	1.63987167	1	0.20034308	
8	1400-1600	Simple	6	6601.8418	6632.28142	-3294.9209				
		Interaction	7	6603.81168	6639.32457	-3294.90584	0.03011701	1	0.862225	
9	1600-1800	Simple	6	6609.55005	6639.98967	-3298.77503				
		Interaction	7	6609.19756	6644.71045	-3297.59878	2.35249046	1	0.12508303	
10	1800-2000	Simple	6	6625.53656	6655.97618	-3306.76828				
		Interaction	7	6627.60356	6663.11644	-3306.80178	0	1	1	

Table 0.14: Loglikelihood test results evaluating the goodness-of-fit of the interaction model over the simple model. **Note:** † p<.1; * p<.05; ** p<.01; *** p<.001.

L2 German learners of Dutch

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	3884.58949	3914.61878	-1936.29474				
		Interaction	7	3884.73106	3919.76523	-1935.36553	1.85843034	1	0.17280607	
2	200-400	Simple	6	5897.01976	5927.04905	-2942.50988				
		Interaction	7	5898.60897	5933.64314	-2942.30449	0.41078658	1	0.52157054	
3	400-600	Simple	6	5998.82868	6028.85797	-2993.41434				
		Interaction	7	6000.75399	6035.78816	-2993.37699	0.07469344	1	0.78462184	
4	600-800	Simple	6	6108.97966	6139.00895	-3048.48983				
		Interaction	7	6110.8985	6145.93267	-3048.44925	0.08116635	1	0.77572299	
5	800-1000	Simple	6	6099.44764	6129.47694	-3043.72382				
		Interaction	7	6100.84759	6135.88176	-3043.42379	0.60005477	1	0.43855713	
6	1000-1200	Simple	6	6165.80389	6195.83318	-3076.90195				
		Interaction	7	6165.71331	6200.74749	-3075.85666	2.09057976	1	0.14820982	
7	1200-1400	Simple	6	6172.69359	6202.72289	-3080.3468				
		Interaction	7	6170.72793	6205.7621	-3078.36396	3.96566726	1	0.04643711	*
8	1400-1600	Simple	6	6142.78976	6172.81905	-3065.39488				
		Interaction	7	6142.55933	6177.5935	-3064.27966	2.23042654	1	0.13531649	
9	1600-1800	Simple	6	6096.3211	6126.35039	-3042.16055				
		Interaction	7	6097.59155	6132.62572	-3041.79578	0.72954896	1	0.39302943	
10	1800-2000	Simple	6	6120.23009	6150.25938	-3054.11505				
		Interaction	7	6122.04008	6157.07426	-3054.02004	0.19001042	1	0.66290795	

Table 0.15: Loglikelihood test results evaluating the goodness-of-fit of the interaction model over the simple model. **Note:** † p<.1; * p<.05; ** p<.01; *** p<.001.

Chapter 4.3
L2 Dutch learners of German

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	3961.46769	3992.83803	-1974.73385				
		Animacy	11	3967.39467	4024.90695	-1972.69734	4.07302176	5	0.53895127	
2	200-400	Simple	6	7075.7434	7107.11373	-3531.8717				
		Animacy	11	7081.57487	7139.08714	-3529.78743	4.16852721	5	0.52541626	
3	400-600	Simple	6	7505.2118	7536.58213	-3746.6059				
		Animacy	11	7509.49859	7567.01086	-3743.74929	5.71321286	5	0.33513204	
4	600-800	Interaction	7	7612.11459	7648.71331	-3799.0573				
		Animacy	11	7614.02429	7671.53656	-3796.01214	6.09030394	4	0.19250506	
5	800-1000	Simple	6	7690.12945	7721.49979	-3839.06473				
		Animacy	11	7694.49194	7752.00422	-3836.24597	5.63751148	5	0.34310271	
6	1000-1200	Interaction	7	7730.61381	7767.21253	-3858.3069				
		Animacy	11	7729.89997	7787.41225	-3853.94999	8.71383584	4	0.06866408	†
7	1200-1400	Interaction	7	7767.69743	7804.29615	-3876.84871				
		Animacy	11	7769.99385	7827.50612	-3873.99692	5.70358189	4	0.22240556	
8	1400-1600	Interaction	7	7745.8541	7782.45282	-3865.92705				
		Animacy	11	7744.696	7802.20827	-3861.348	9.15810131	4	0.05726693	†
9	1600-1800	Interaction	7	7715.863	7752.46172	-3850.9315				
		Animacy	11	7715.52379	7773.03606	-3846.7619	8.33920854	4	0.07991369	†
10	1800-2000	Simple	6	7677.61	7708.98033	-3832.805				
		Animacy	11	7680.66761	7738.17989	-3829.33381	6.94238662	5	0.22496047	

Table 0.16: Loglikelihood test results evaluating the goodness-of-fit of the animacy model over the interaction model or the simple model.
Note: † p<.1; * p<.05; ** p<.01; *** p<.001.

L2 German learners of Dutch

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	4788.4441	4820.44224	-2388.22205				
		Animacy	11	4796.01475	4854.678	-2387.00737	2.42935817	5	0.78709424	
2	200-400	Simple	6	7816.72642	7848.72456	-3902.36321				
		Animacy	11	7821.03511	7879.69836	-3899.51755	5.69131928	5	0.33742262	
3	400-600	Simple	6	8056.87967	8088.8778	-4022.43983				
		Animacy	11	8056.57915	8115.24241	-4017.28958	10.300513	5	0.0671546	†
4	600-800	Interaction	7	8203.27381	8240.60497	-4094.63691				
		Animacy	11	8193.94082	8252.60408	-4085.97041	17.3329868	4	0.00166517	**
5	800-1000	Interaction	7	8359.90701	8397.23817	-4172.9535				
		Animacy	11	8335.71896	8394.38221	-4156.85948	32.1880508	4	0.00000175	***
6	1000-1200	Interaction	7	8437.09566	8474.42682	-4211.54783				
		Animacy	11	8429.20632	8487.86958	-4203.60316	15.8893368	4	0.00317131	**
7	1200-1400	Interaction	7	8439.7676	8477.09876	-4212.8838				
		Animacy	11	8443.44093	8502.10418	-4210.72046	4.32666843	4	0.36359757	
8	1400-1600	Interaction	7	8479.99932	8517.33048	-4232.99966				
		Animacy	11	8483.03855	8541.7018	-4230.51928	4.96076506	4	0.29134699	
9	1600-1800	Interaction	7	8446.63303	8483.96419	-4216.31652				
		Animacy	11	8449.14599	8507.80924	-4213.57299	5.48704739	4	0.24087037	
10	1800-2000	Interaction	7	8434.26022	8471.59139	-4210.13011				
		Animacy	11	8438.81206	8497.47531	-4208.40603	3.44816646	4	0.48580339	

Table 0.17: Loglikelihood test results evaluating the goodness-of-fit of the animacy model over the interaction model or the simple model.
Note: † p<.1; * p<.05; ** p<.01; *** p<.001.

Chapter 4.4

L2 Dutch learners of German

Time window	in ms	Model	Df	AIC	BIC	logLik	Chisq	Chi.Df	Pr..Chisq	
1	0-200	Simple	6	3961.46769	3992.83803	-1974.73385				
		Proficiency	11	3970.64336	4028.15563	-1974.32168	0.8243355	5	0.97545525	
2	200-400	Simple	6	7075.7434	7107.11373	-3531.8717				
		Proficiency	11	7083.32339	7140.83566	-3530.6617	2.42000697	5	0.78849101	
3	400-600	Simple	6	7505.2118	7536.58213	-3746.6059				
		Proficiency	11	7512.61112	7570.12339	-3745.30556	2.60067818	5	0.76126222	
4	600-800	Interaction	7	7612.11459	7648.71331	-3799.0573				
		Proficiency	11	7614.5403	7672.05258	-3796.27015	5.57428634	4	0.23327641	
5	800-1000	Simple	6	7690.12945	7721.49979	-3839.06473				
		Proficiency	11	7690.47669	7747.98896	-3834.23835	9.65276466	5	0.08569354	†
6	1000-1200	Interaction	7	7730.61381	7767.21253	-3858.3069				
		Proficiency	11	7733.21431	7790.72658	-3855.60715	5.3994996	4	0.2487058	
7	1200-1400	Interaction	7	7767.69743	7804.29615	-3876.84871				
		Proficiency	11	7771.63802	7829.15029	-3874.81901	4.05940877	4	0.39802545	
8	1400-1600	Interaction	7	7745.8541	7782.45282	-3865.92705				
		Proficiency	11	7750.3476	7807.85987	-3864.1738	3.5065041	4	0.47689007	
9	1600-1800	Interaction	7	7715.863	7752.46172	-3850.9315				
		Proficiency	11	7713.59804	7771.11031	-3845.79902	10.2649588	4	0.03619334	*
10	1800-2000	Simple	6	7677.61	7708.98033	-3832.805				
		Proficiency	11	7670.33915	7727.85142	-3824.16957	17.2708519	5	0.00401376	**

Table 0.18: Loglikelihood test results evaluating the goodness-of-fit of the proficiency model over the interaction model or the simple model.
Note: † $p<.1$; * $p<.05$; ** $p<.01$; *** $p<.001$.

Appendix E: European Reference Frame

The Levels of the Common European Framework as measured by the German Placement Test by the Goethe Institute[67]

The German Placement Test categorized the participants into Basic Learners (levels A1 and A2), Independent Users (levels B1 and B2), and Proficient Users (levels C1 and C2) according to the following proficiency scores.

Proficiency Category	Levels	GPT Scores
Basic User	A1	1-5
	A2	6-10
Independent User	B1	11-15
	B2	16-20
Proficient User	C1	21-27
	C2	28-30

Table 0.19: German Placement Test Scores and according Levels of the European Reference Frame (Goethe Institute)

Descriptions of the Common European Framework by the Council of Europe[68]

Basic User

A1: Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

A2: Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.

[67] http://www.goethe.de/cgi-bin/einstufungstest/einstufungstest.pl
[68] http://www.coe.int/t/dg4/linguistic/CADRE_EN.asp

Independent User

B1: Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.

B2: Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.

Proficient User

C1: Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.

C2: Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.

Samenvatting in het Nederlands

In coherente discourse wordt vaak meermaals naar dezelfde discourse-entiteit verwezen, iets wat doorgaans bereikt wordt door het gebruik van persoonlijke voornaamwoorden (pronomina) zoals *hij, zij* en *het*. In tegenstelling tot bijvoorbeeld het Engels hebben Duits en Nederlands twee pronominale vormen die gebruikt kunnen worden om naar een mannelijke entiteit in het enkelvoud te verwijzen, zoals in voorbeeld (1) te zien is: het 'normale' pronomen (Duits *er*, Nederlands *hij*) en het zgn. d-pronomen (Duits *der*, Nederlands *die*).

(1) Duits: **Peter$_i$** wollte Tennis spielen. Doch **er$_i$/der$_i$** war krank.
Nederlands: **Peter$_i$** wilde gaan tennissen. Maar **hij$_i$/die$_i$** was ziek.

In dit proefschrift onderzoek ik hoe moedertaalsprekers en tweedetaalleerders van het Duits en het Nederlands normale pronomina en d-pronomina in ambigue contexten interpreteren wanneer meer dan één antecedent voorhanden is. Specifiek werd de invloed van de volgorde, de informatiestructurele status en de animaatheid van de mogelijke antecedenten op de voorkeur voor de interpretatie van beide typen pronomina onderzocht. Ook werden de individuele verschillen onderzocht in de voorkeuren voor de interpretatie van pronomina tussen deelnemers aan de studie, door te kijken naar de verschillen in *on-* en *off-line* interpretatie van pronomina en naar verschillen in de taalvaardigheid van de tweedetaalleerders in hun tweede taal.

Voorgaande psycholinguïstische studies naar de interpretatie van normale en d-pronomina hadden tot doel de vraag te beantwoorden hoe de onderlinge volgorde en de grammaticale rol van twee mogelijke antecedenten de interpretatie van het verwijzende pronomen beïnvloeden. In deze studies werd daarom de interpretatie van pronomina die volgen op een zin met SVO-volgorde (waarin het subject als eerste genoemd wordt en het object als tweede) vergeleken met de interpretatie van pronomina die volgen op een OVS-zin (waarin het object als eerste genoemd wordt en het subject als tweede). Het veranderen van de woordvolgorde in de voorzin beïnvloedt echter ook de informatiestructurele status van de beide mogelijke antecedenten, wat op zijn beurt weer invloed op de interpretatie van de pronomina kan hebben. Om deze reden zijn in de experimenten die in dit proefschrift besproken worden de pronomina gepresenteerd na comparatiefstructuren zoals *"De kast is zwaarder dan de tafel. Hij..."*. Omdat beide antecedenten in zulke zinnen in de nominatief staan, kon de vraag hoe de volgorde van de beide antecedenten in de eerste zin de interpretatie van het pronomen beïnvloedt beter worden beantwoord. Ook kon het effect

van een niet-canonieke volgorde van de antecedenten onderzocht worden met behulp van comparatiefstructuren van het type *"Zwaarder dan de tafel is de kast. Hij..."*. Daarnaast werd gekeken naar de rol van animaatheid van de mogelijke antecedenten, in tegenstelling tot eerdere studies, die uitsluitend naar animate antecedenten gekeken hebben. In de experimenten werden twee verschillende taken gebruikt: een *visual-world eye-tracking*-taak, waarmee *on-line* interpretatievoorkeuren onderzocht kunnen worden, en een vragenlijst met een *forced-choice* begripstaak, waarmee de uiteindelijke interpretatievoorkeuren gemeten konden worden. Het doel van deze opzet was om zowel de inferenties van de deelnemers van moment tot moment te kunnen volgen als ook om de resultaten van de studies te kunnen vergelijken met de resultaten van eerdere studies.

In sectie 3.1 worden de experimenten beschreven waarin de interpretatie van normale en d-pronomina na gewone comparatiefstructuren in het Duits en het Nederlands onderzocht werden. De resultaten van de *eye-tracking*-expermenten lieten geen verschillen zien tussen de interpretatievoorkeuren in het Duits en het Nederlands (het normale pronomen werd bij voorkeur geïnterpreteerd als verwijzend naar het als eerst genoemde nomen, een zgn. *first-mention*-effect, het d-pronomen werd bij voorkeur geïnterpreteerd als tweede genoemde nomen, een *second-mention*-effect). De twee talen tonen echter verschillen in de timing van de effecten: in het Duits vond de interpretatie van het normale pronomen relatief laat plaats in vergelijking met het d-pronomen, terwijl de interpretatie in het Nederlands direct plaatsvond. Dit duidt erop dat het normale pronomen in het Duits een grotere ambiguïteit vertoont dan in het Nederlands, een resultaat dat overeenkomt met het idee dat het d-pronomen in het Duits speciaal voor non-topicaliteit gemarkeerd is, terwijl het normale pronomen wat dit betreft neutraal is. De resultaten laten zien dat dit voor het Nederlands ogenschijnlijk niet geldt.

De resultaten van het *off-line* begripsexperiment verschilden enigszins van de resultaten van de *on-line eye-tracking*-taak. In beide talen was de interpretatie van het normale pronomen als verwijzend naar de eerste entiteit maximaal (*at ceiling*), maar de interpretatie van het d-pronomen als verwijzend naar de tweede entiteit was slechts marginaal significant. Dit resultaat wijkt af van wat in eerdere *off-line* experimenten met invul- en beoordelingtests met standaard volgordes (SVO) is gevonden. Hoewel in sommige experimenten geen voorkeur is gevonden voor normale pronomina en een duidelijke voorkeur voor de als tweede genoemde entiteit voor d-pronomina (voor het Duits: Bosch, et al., 2007a; Wilson, 2009), is in andere experimenten een sterke voorkeur voor de eerstgenoemde entiteit voor normale pronomina en een voorkeur voor de als tweede genoemde entiteit voor d-pronomina gevonden (voor het Nederlands: Kaiser & Trueswell, 2004; voor het Duits: Roberts, et al., in progress). Het feit dat de *off-line* interpretatie van normale en d-

pronomina geen eenduidig resultaat hebben opgeleverd kan het gevolg zijn geweest van kleine verschillen in de materialen en de taken. Ten eerste zou het kunnen zijn dat de *off-line* resultaten in de hier gepresenteerde studie van andere resultaten afwijken aangezien we gebruik hebben gemaakt van comparatiefstructuren (waarbij we bovendien werkwoorden die mogelijk de voorkeur nog beïnvloeden hebben vermeden). Bovendien moesten onze proefpersonen in het *off-line* begripsexperiment, in tegenstelling tot eerdere studies, ook een keuze voor een van de twee genoemde entiteiten maken als ze niet zeker waren van de interpretatie.

In sectie 3.2 wordt het onderzoek besproken dat ingaat op de vraag of de interpretatievoorkeuren voor normale en d-pronomina met een inleidende zin met niet-canonieke structuur afwijken van pronomina na zinnen met een standaard-structuur, oftewel of de informatiestructuur van de inleidende zin invloed heeft op de interpretatie van de pronomina. Deze vraag is vooral belangrijk vanwege eerdere studies, die geprobeerd hebben de invloed van de volgorde en de grammaticale rol van beide mogelijke antecedenten op de interpretatie te bepalen. In de experimenten werden de interpretatievoorkeuren onderzocht van pronomina met voorzinnen met een structuur zoals *"Zwaarder dan de tafel is de kast. Hij..."*. De resultaten laten zien dat veranderingen in de structuur van de eerste zin de interpretatie van beide typen pronomina beïnvloedt. Om precies te zijn: in vergelijking met de experimenten uit sectie 3.1, waarin de inleidende zin een topic-comment-structuur had, was de als tweede genoemde entiteit pragmatisch zeer gemarkeerd wanneer de inleidende zin een topic-focus-structuur had, zoals in de experimenten in sectie 3.2. Het asymmetrische interpretatiepatroon bij standaard-volgorde (normale pronomina verwijzen bij voorkeur naar het als eerste genoemde nomen, d-pronomina bij voorkeur naar het als tweede genoemde nomen) werd dus niet opnieuw gevonden, in het Duits noch in het Nederlands. Zowel normale als d-pronomina worden geïnterpreteerd als verwijzend naar het tweede nomen, een effect dat in de *eye-tracking*-experimenten zeer snel na het presenteren van het pronomen al optreedt. Dit geeft aan dat de informatiestructuur onmiddellijk betekenisonderscheidend werkt en dat focusinformatie een sterke aanwijzing is bij de interpretatie van pronomina. Dit resultaat komt overeen met resultaten van onderzoek naar Engelse pronomina waaruit blijkt dat aan de ene kant topics erg opvallend zijn voor de interpretatie van pronomina, maar dat pronomina ook een voorkeur kunnen hebben voor elementen met contrastieve focus. (Arnold, 1998, 1999, 2001; Cowles, 2003; Cowles, et al., 2007). Focusinformatie heeft invloed op de interpretatie van zowel normale als d-pronomina en leidt tot overlappende interpretatiemogelijkheden, niet alleen wanneer er slechts één mogelijke antecedent is, maar ook als er twee mogelijke antecedenten zijn. Zulke overlappende effecten kunnen niet

verklaard worden door referentietheorieën die de nadruk leggen op de asymmetrische distributionele eigenschappen van beide typen pronomina.

Gezien het feit dat beide pronominatypen in het Duits en het Nederlands zowel naar animate als ook naar niet animate entiteiten kunnen verwijzen, is ook het effect van animaatheid op de interpretatie van pronomina onderzocht (sectie 3.3). In voorgaande studies zijn uitsluitend animate entiteiten gebruikt; het was daarom nooit duidelijk in hoeverre de interpretatievoorkeuren ook naar niet animate entiteiten gegeneraliseerd kunnen worden. Verrassend genoeg bleek dat er in het Duits een animaatheidseffect is bij de interpretatie van pronomina, maar in het Nederlands niet. In beide talen gingen de resultaten in dezelfde richting voor zowel animate als niet animate elementen; het normale pronomen werd geïnterpreteerd als verwijzend naar het als eerste genoemde nomen, het d-pronomen als verwijzend naar het als tweede genoemde nomen. In het Duits werd het normale pronomen echter na animate entiteiten relatief snel geïnterpreteerd (na 600 ms, vergeleken met 1400 ms voor de animate en niet animate entiteiten samen), wat suggereert dat de animaatheidsinformatie de interpretatie vereenvoudigde.

In het laatste deel van hoofdstuk 3 wordt de invloed van individuele verschillen op de interpretatie van pronomina besproken. Hoewel de *off-line* resultaten in sectie 3.1 zoals gezegd overeenstemmen met de *on-line* resultaten was er toch een kwalitatief verschil in de uiteindelijke interpretatie; de voorkeur van het normale pronomen voor het als eerste genoemde nomen was maximaal (*at ceiling*), terwijl de voorkeur van het d-pronomen voor het als tweede genoemde nomen slechts marginaal significant was. Na een heranalyse van de gegevens uit de *off-line* test werden individuele verschillen in de interpretatievoorkeuren voor het d-pronomen gevonden. Het bleek dat sommige proefpersonen bij de d-pronomina een voorkeur hadden voor het als eerste genoemde nomen terwijl anderen een voorkeur hadden voor het als tweede genoemde nomen. De data zijn vervolgens opnieuw geanalyseerd om te zien of zulke individuele verschillen ook een effect hadden op de *on-line* verwerking van de interpretatie. Opmerkelijk genoeg vonden we voor sprekers van het Duits dat er tijdens het verwerken van de uiting een effect was in de richting van het als tweede genoemde nomen (bij 800 ms), zelfs wanneer het d-pronomen uiteindelijk als verwijzend naar het als eerste genoemde nomen geïnterpreteerd werd. Dit effect onderstreept dat het d-pronomen in het Duits speciaal voor non-topicaliteit gemarkeerd is. In het Nederlands weken de resultaten enigszins van de Duitse resultaten af. Wanneer in de *off-line* taak het als tweede genoemde antecedent werd gekozen, was deze voorkeur ook duidelijk in de *on-line* taak te zien. Wanneer het als eerste genoemde nomen werd gekozen, was er geen *on-line* voorkeur; dat betekent dat als de deelnemers niet zeker waren over de coreferentie-relatie, ze de als eerste genoemde topicale entiteit kozen.

In het tweede gedeelte van het proefschrift is onderzocht hoe tweedetaalleerders normale en d-pronomina in hun tweede taal (T2) interpreteren. De Duitstalige experimenten werden uitgevoerd met Nederlandse T2-leerders van het Duits, de Nederlandstalige experimenten met Duitse T2-leerders van het Nederlands. We hebben onderzocht of de interpretatie van pronomina in T2 beïnvloedt wordt door de volgorde waarin de beide mogelijke antecedenten genoemd worden (sectie 4.1). Voor de Nederlandse T2-leerders van het Duits toonden zowel de *off-line* als de *on-line* resultaten aan dat dit niet het geval was: zij hadden een algemene voorkeur voor de interpretatie van beide typen pronomina als verwijzend naar de als eerste genoemde, topicale, entiteit. Hiervoor zijn twee verklaringen mogelijk: de Nederlandse T2-leerders verwarren beide typen pronomina of ze hebben twee vormen voor één functie (het verwijzen naar de topicale antecedent). De Duitse T2-leerders van het Nederlands daarentegen maakten wel een onderscheid tussen de functies van beide pronominale vormen: het normale pronomen werd geïnterpreteerd als verwijzend naar de als eerste genoemde entiteit, het d-pronomen als verwijzend naar de als tweede genoemde entiteit. Zij vertoonden dus met de doeltaal overeenkomende interpretatievoorkeuren. Aangezien de algemene generalisering dat normale pronomina naar topics verwijzen en d-pronomina een voorkeur voor niet-topics hebben voor beide talen geldt, was het niet duidelijk waarom de resultaten Nederlandse T2-leerders van het Duits niet met de interpretatievoorkeuren in de doeltaal overeenkwamen. Verschillende verklaringen zijn hier mogelijk:

- Fonologische overeenkomsten: het normale pronomen en het d-pronomen in het Duits hebben sterk op elkaar lijkende vormen (*er – der*) in tegenstelling tot de Nederlandse vormen (*hij – die*), wat tot verwarring van de beide vormen zou kunnen leiden.
- Input: beide T2-groepen hebben de tweede taal vooral in klassikaal onderwijs geleerd, waardoor het mogelijk is dat het Duitse d-pronomen minder vaak is voorgekomen dan het Nederlandse d-pronomen (aangezien dat ook een normaal aanwijzend voornaamwoord is en het in geschreven taal voorkomt).
- Taalvaardigheid: de Nederlandse T2-leerders van het Duits waren minder taalvaardig in de doeltaal dan de Duitse T2-leerders van het Nederlands.
- Gemarkeerdheid en L1-transfer: aangezien het Duitse d-pronomen speciaal voor non-topicaliteit is gemarkeerd, kan het zijn dat de Duitse T2-leerders deze eigenschap in hun Nederlands overgenomen hebben.

De eerste verklaring werd verworpen op grond van de kwaliteit van de *on-line* effecten en de *off-line* taak. Hoewel beide vormen uiteindelijk als verwijzend naar de als eerste genoemde entiteit werden geïnterpreteerd, was er een verschil in timing: het d-pronomen

werd later geïnterpreteerd dan het normale pronomen, wat suggereert dat het d-pronomen sterker ambigu is. De interactie tussen de volgorde van de antecedenten en het pronomentype was significant, wat er ook op wijst dat de beide pronomina op andere wijze geïnterpreteerd werden, ook al is het resultaat in beide gevallen gelijk. Bij de *off-line* taak kregen de proefpersonen de zinnen in geschreven vorm, zodat ze zagen dat de pronomina verschillende vormen hadden. Om die reden zijn we er niet van uitgegaan dat de beide pronomentypen verward konden worden. In plaats daarvan hebben we gekeken of de Nederlandse T2-leerders twee vormen voor één functie hadden. De tweede verklaring (het ontbreken van het d-pronomen in het taalaanbod van de Nederlandse T2-leerders) leek waarschijnlijker. Roberts et al. (in progress) hebben echter de interpretatievoorkeuren voor normale en d-pronomina onderzocht van Finse T2-leerders van het Duits die ook in een onderwijsomgeving Duits hadden geleerd en hebben hierbij een robuust effect gevonden in de voorkeur voor interpretatie van d-promonima als verwijzend naar het als tweede genoemde nomen. Dat betekent dat ook de tweede verklaring niet waarschijnlijk is. De rol van de vaardigheid in T2 is in een aparte studie onderzocht (sectie 4.4). Uitgaande van leerdersvarianten, is een mogelijke verklaring voor de waargenomen interpretatiepatronen de eventuele invloed van principes in de brontaal op de interpretatie van pronomina in de doeltaal door Duitse T2-leerders van het Nederlands. In het Duits is het d-pronomen speciaal gemarkeerd voor non-topicaliteit. Deze eigenschap kunnen de Duitse T2-leerders van het Nederlands uit hun eerste taal over hebben genomen. Dat zou betekenen dat, hoewel het Duits en het Nederlands typologisch zeer nauw verwant zijn en dezelfde algemene interpretatievoorkeuren voor normale en d-pronomina laten zien, er toch een subtiel verschil is in de gemarkeerdheid van het d-pronomen dat de waargenomen verschillen zou kunnen veroorzaken.

In sectie 4.2 ligt de nadruk op de informatiestructuur van de inleidende zin. In deze sectie worden experimeten besproken waarin de comparatiefstructuur in de niet-canonieke volgorde werd aangeboden. Eerdere studies naar Duits als tweede taal (Roberts, et al., in progress; Wilson, 2009) vonden geen duidelijke interpretatievoorkeuren na inleidende zinnen met een niet-canonieke OVS-structuur; een mogelijke hypothese is dat het verwerken van OVS-structuren het taalverwerkingssysteem sterk belast, wat de daaropvolgende interpretatie van de pronomina beïnvloed kan hebben. Een aanwijzing voor zulke moeilijkheden is het feit dat, nadat Wilson (2009) haar T2-leerders op basis van hun vaardigheid in T2 in twee groepen had opgedeeld, ze een verschil vond in de interpretatievoorkeuren voor d-pronomina (de groep met lagere taalvaardigheid had een voorkeur voor het als eerste genoemde nomen, de groep met hogere taalvaardigheid voor het als tweede genoemde nomen), terwijl er geen verschil was in de interpretatievoorkeuren voor d-promonina na inleidende zinnen met een standaard-SVO-structuur (beide groepen

hadden geen voorkeur). Aangezien in onze studie beide mogelijke antecedenten in de nominatief stonden, zijn we ervan uitgegaan dat de eventuele extra verwerkingsbelasting minimaal zou blijven. Bovendien, gezien het feit dat er cross-linguistische aanwijzingen zijn uit de data van moedertaalsprekers dat focus een sterkere aanwijzing is dan de volgorde waarin de nomina worden genoemd, waren we geïnteresseerd in de vraag of dit ook voor T2-leerders geldt. De resultaten ondersteunen deze hypothese: beide groepen T2-leerders bleken gevoelig voor deze informatiestructurele aanwijzing. De focusinformatie was sterker dan eventuele effecten van de volgorde van de antecedenten en de T2-leerders vertoonden met de doeltaal overeenkomende interpretatievoorkeuren. De *off-line* resultaten bevestigden met de *on-line* resultaten: de als tweede genoemde, gefocusseerde entiteit werd voor beide typen pronomina als antecedent genomen. Over het geheel genomen moet geconcludeerd worden dat de T2-leerders minder moeite hadden met dit soort antecedentenstructuur (waarbij beide typen pronomina overlappende functies hebben) dan bij de standaardzinnen uit de experimenten in deze studie en bij de niet-canonieke OVS-structuren uit eerdere studies, aangezien beide groepen op doeltaalniveau presteerden wanneer geen onderscheid in functie van de beide pronominatypen gemaakt hoefde te worden. Pas wanneer T2-leerders onderscheid moeten maken tussen verschillende functies van de twee pronominatypen lijken de resultaten af te wijken van prestaties door moedertaalsprekers, zoals in de standaard-contexten, waarin normale pronomina bij voorkeur naar topics verwijzen en d-pronomina bij voorkeur naar niet-topics. In zulke contexten weken de interpretatievoorkeuren van de Nederlandse T2-leerders af van de voorkeuren van moedertaalsprekers.

In sectie 4.3 is het effect van animaatheid op de interpretatie van pronomina door T2-leerders onderzocht. In de literatuur over tweedetaalverwerving wordt geopperd dat T2-leerders hoofdzakelijk op niet-structurele lexicaal-semantische en pragmatische aanwijzingen vertrouwen bij het verwerken van hun tweede taal (Clahsen & Felser, 2006a), wat betekent dat animaatheid een belangrijke aanwijzing zou kunnen zijn voor T2-leerders. Voor beide groepen T2-leerders was de interpretatievoorkeur voor normale pronomina voor animate en niet animate elementen gelijk. Alleen in de timing was er een verschil, in die zin dat het normale pronomen sneller geïnterpreteerd werd voor animate elementen. Dit suggereert dat animaatheid het optreden van een *first-mention*-effect vergemakkelijkt, wat het gevolg zou kunnen zijn van het feit dat animate entiteiten vaker topics zijn. Verrassend genoeg werd er voor het d-pronomen een verschil in voorkeur waargenomen in beide groepen T2-leerders. De analyse van de Nederlandse T2-leerders van het Duits liet zien dat de *first-mention*-voorkeur voor het d-pronomen alleen bij niet animate antecedenten optrad en dus niet bij animate antecedenten. De Duitse T2-leerders van het Nederlands vertoonden een *second-mention*-voorkeur voor het d-pronomen na *animate* elementen. Met niet

animate elementen vertoonden ze eerst een *first-mention*-voorkeur die echter later in een *second-mention*-voorkeur veranderde. Hoewel Nederlandse T2-leerders van het Duits voorheen een interpretatiestrategie vertoonden waarbij ongeacht het type pronomen het topic standaard werd gekozen, toonde de animaatheidsanalyse aan dat ze de beide typen pronomina niet op dezelfde manier behandelden: ze vertoonden een eerste aanzet tot het onderscheiden van beide functies, tenminste na animate antecedenten. Verder is het ook zeer interessant dat de Duitse T2-leerders van het Nederlands, die bij het d-pronomen een op de doeltaal gelijkende *second-mention*-voorkeur vertoonden in de niet-gedifferentieerde analyse (sectie 4.1), een vroege *first-mention*-voorkeur hadden bij het d-pronomen, wanneer dat op niet animate antecedenten volgde; pas later wisselde deze voorkeur naar een *second-mention*-voorkeur. Dit resultaat geeft aan dat in beide T2-leerdergroepen de *second-mention*-voorkeur voor d-pronomina versterkt wordt na animate antecedenten. Dit kan het gevolg zijn van het bovengenoemde feit dat bij een verandering van topic het waarschijnlijker is dat de discourse voortgezet wordt met een animate entiteit, aangezien dat met grotere waarschijnlijkheid een volgend topic is. De T2-leerders lijken erg op deze lexicaal-semantische en pragmatische aanwijzingen te vertrouwen bij het interpreteren van pronomina in de tweede taal.

In sectie 4.4 is ook de taalvaardigheid van de T2-leerders onderzocht. Aangezien de Duitse T2-leerders een vrij homogene, op de doeltaal gelijkende distributie vertoonden, was deze analyse alleen zinvol voor de Nederlandse T2-leerders van het Duits, die in het experiment in sectie 4.1 niet op doeltaalniveau presteerden. Deze T2-leerders werden in twee groepen opgesplitst met respectievelijk een lage (niveau B) en een hoge (niveau C) taalvaardigheid. Er kon aangetoond worden dat de factor taalvaardigheid een goede voorspeller was voor de interpretatievoorkeur. Beide groepen vertoonden zowel *on-line* als *off-line* een *first-mention*-voorkeur voor het normale pronomen. Ze verschilden echter in de *on-line*-interpretatie van het d-pronomen: de groep met laag niveau interpreteerde dat als verwijzend naar de als eerste genoemde antecedent en de groep met hoog niveau interpreteerde het als verwijzend naar het als tweede genoemde antecedent. De *off-line* resultaten weerspiegelden de *on-line* resultaten voor de groep met laag niveau; de groep met hoog niveau vertoonde *off-line* geen voorkeur, wat een gevolg kan zijn geweest van het lage aantal items (de groep bestond uit slecht 6 proefpersonen). De taalvaardigheidsanalyse toont echter aan dat naarmate T2-leerders een grotere taalvaardigheid ontwikkelen in de tweede taal, hun interpretatie zich ontwikkelt van een systeem waarin twee pronominale vormen één functie hebben naar een systeem waarin ze twee functies onderscheiden, zoals in de doeltaal.

Curriculum Vitae

Miriam Ellert was born in 1979 in Bonn, Germany. In 1996, she spent one year as a foreign exchange student at the Victor Valley High School in California (USA), and graduated from the Grotenbach-Gymnasium in Gummersbach, Germany in 2000. She then pursued university studies in Strasbourg (France), Madrid (Spain) and Aachen (Germany). She graduated from the Rheinisch-Westfälische Technische Hochschule in Aachen in Psychology in 2005. She then was awarded a scholarship from the Max-Planck-Gesellschaft to prepare her Ph.D. thesis at the Max Planck Institute for Psycholinguistics in Nijmegen (The Netherlands). After that time, she received a grant from the Deutsche Akademische Austausch Dienst which allowed her to spend six months in the English Linguistics Department at the University of Mannheim (Germany). Currently, she has a postdoctoral position in the German Linguistics Department at the University of Göttingen (Germany).

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 midsagittalarticulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological

approach. *Daniëlle van den Brink*

25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*

50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prososdic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous Pronoun Resolution in L1 and L2 German and Dutch. *Miriam Ellert*

Mark Twain – *The Awful German Language*
translated into German by Ana Maria Brock[69]

"Es ist ein rauher Tag. Hört den Regen, wie *er* strömt, und den Hagel, wie *er* prasselt [...]. Ach, das arme Fischweib, *es* steckt im Sumpfe fest [...]. *Es* öffnet den Mund, um Hilfe zu rufen, aber wenn ein Laut aus *ihm* herausgedringt, ach! wird *er* vom Wüten des Sturmes erstickt. Und jetzt hat die Katze einen der Fische erwischt, und *sie* wird gewiß mit *ihm* entkommen [...].

Der Leser kann selbst sehen, daß diese Pronomengeschichte für die ungeübte Zunge eine sehr mißliche Sache ist."

[69] See beginning of this thesis for the original English version.