Processing Predicate NPs:

An EEG study

Ann Hermansson

Supervisors: Dr. Eva Klingvall, Lund University
Dr. Fredrik Heinat, Linnæus University

Centre for Languages and Literature, Lund University
MA in Language and Linguistics, English Linguistics
SPVR01 Language and Linguistics: Degree project – Master's (Two Years) Thesis, 30 credits
Spring 2023

Abstract

Theories on discourse accessibility and givenness suggest that nominal discourse entities vary with regard to how likely they are to act as antecedents to anaphoric expressions. Many factors are thought to affect the accessibility of a nominal entity in the mental model, including its morphosyntactic form and grammatical role. Moreover, previous research has shown that reflexive pronouns can be illusorily licensed, if they are preceded by a nominal that matches them in terms of morphosyntactic or semantic features (e.g. number and gender). While some nominals that are unavailable as antecedents by syntactic criteria (cf. Binding Theory) have been studied with regard to discourse accessibility and pronoun resolution, as well as to what extent they give rise to agreement attraction effects, non-referential predicate nominals have not yet been studied in a pronoun-resolution paradigm. The current study made use of an EEG experiment in order to investigate to what extent predicate noun phrases are considered potential antecedents during online pronoun resolution. This was indicated by whether they gave rise to agreement-attraction effects. The results showed that a predicate nominal antecedent yielded a significant P600 response, and also a late sustained anterior negativity (L-AN) response against a baseline, referential nominal condition. The P600 and L-AN responses were interpreted as indications of the predicate nominals being realised as poor antecedents to the anaphoric pronoun. However, gender congruency between the predicate noun phrase and the anaphoric pronoun attenuated the P600 amplitude. The amplitude attenuation was interpreted as the predicate nominal being falsely perceived as an antecedent to the anaphoric pronoun, as indicated by the presence of agreement attraction effects. The results therefore suggest that a predicate noun phrase is considered a potential antecedent to an anaphoric pronoun during the early stages of pronoun resolution.

Acknowledgements

It is not only appropriate, but also a pleasure, to acknowledge a few special people who have been significantly helpful during the writing of this thesis.

First and foremost, I want to direct a big thank you to all my participants, who not only lent me their time and scalps, but also showed such great interest in the project that they told their friends to participate as well.

I gratefully acknowledge the Humanities Lab, for providing the necessary equipment and room to conduct the experiment. Thank you, also, to *Språkmastern*, for providing funding for the participation enumeration, and Jordan Zlatev, program coordinator, for orchestrating it all. Thank you to Henrik Garde, whose MatLab-scripts I had the opportunity to use with great success.

To my EEG-peers: thank you Huayuan Cui and Katja Kopaeva for providing company and support in and outside the lab, and to Jule Nabrotzky, who despite the *kladd* diligently helped me put on the electrode cap on several participants.

I also want to direct a special thank you to my sweet friends, who cheered me on the loudest when I needed it most; although I sometimes felt like studying with you did more ill than good in terms of productivity, I do not think that I would have made it through without your support, or the endless amount of coffee we consumed together. Thank you: Sonja Holmer, Ravn Kirkegaard and Jinhee Kwon.

The greatest thank you of all goes my supervisors, Eva Klingvall and Fredrik Heinat. Thank you for your time, insightful comments, and instruction, as well as coffee, laughs, and conversation. Most of all, thank you for believing in me from the very beginning.

Finally, this thesis would not exist without Mats Johansson, whose consistent encouragement from my very first semester at Lund University is undoubtedly what made me pursue this degree in the first place.

Table of Contents

List of Figures

List of Tables

Abbreviations

DEM	Demonstrative
DEF	Definite
EEG	Electroencephalography
EMG	Electromyography
EOG	Electrooculography
ERP	Event-Related Potential
GLMM	Generalized Linear Mixed Model
IC	Implicit Causality
INDEF	Indefinite
kΩ	Kiloohm
LMM	Linear Mixed Model
LAN	Left Anterior Negativity
L-(L)AN	Late-(Left) Anterior Negativity
ms	Millisecond
N400	A negative ERP component peaking 400 ms post stimulus onset
NP	Noun Phrase
P600	A positive ERP component peaking 600 ms post stimulus onset
POSS	Possessive
PR	Present Tense
ROI	Region Of Interest
SG	Singular
μV	Microvolt
VP	Verb Phrase

1. Introduction

On the topic of effective and efficient communication, philosopher Paul Grice formulated four maxims that speakers follow in order to convey meaning to their interlocutors. These maxims do not only provide diagnostics for clear communication in terms of *what* should be said, but also of *how* it should be said; the maxims of Quantity and Manner, for example, call for a conversational contribution that is sufficiently informative for the purposes of the conversation. That is, the contribution should be relevant and clear – it should be informative without being overly elaborate, while retaining enough information to avoid ambiguity (Grice, 1991, pp. 26; 27).

While following these maxims might seem difficult, there are in fact many linguistic tools and strategies for providing clear and relevant information to one's interlocutor in line with the Gricean principles. For example, pronominal[1] forms are, *inter alia*, effective tools that are commonly used to refer to entities that are salient from either the situational context or the preceding discourse (Ariel, 1990; Huddleston & Pullum, 2002, p. 425). Assuming that the speaker adheres to the Gricean maxims of effective communication by providing informative and unambiguous information, pronoun resolution (i.e. successfully linking a pronoun with an obvious antecedent) should be unproblematic for the interlocutor. Some potential antecedents are more accessible for anaphoric reference than others, for different reasons. Studies have shown that while some Noun Phrases (NPs) are more likely to license an anaphoric pronoun, others are virtually impossible antecedents in reference resolution. The aim of this study is to investigate one of these supposedly impossible reference antecedents, namely predicate NPs.

The accessibility of an NP, and consequentially its possibility to license an anaphoric pronoun, has been proposed to be dependent on several syntactic, semantic, and pragmatic factors. Wittenberg & Trotzke (2021), for example, investigate how the degree of semantic incorporation of the NP in question affects co-reference establishment. Semantic incorporation regards the extent to which the noun is semantically inseparable from the predicate. In English, semantic incorporation occurs with verbs that are "semantically bleached", such as light-verb constructions (van Geenhoven, 1998; Wittenberg & Trotzke, 2021, p. 89):

[1] from Latin: *pro* 'instead of', *nomen* 'name'

(1) Alexa *gave a hug* to Mohit.

A high degree of semantic incorporation has sometimes been supposed to negatively affect an NP's referential accessibility. Nonetheless, Wittenberg & Trotzke found that the morphosyntactic form of semantically incorporated nouns can modulate discourse prominence. Specifically, semantically incorporated nouns appearing alongside a determiner (e.g. *give a hug* in (1)) appear to be more salient than bare nouns in similar constructions (e.g. *go to church*) (Wittenberg & Trotzke, 2021, p. 94).

English predicate NPs carry more or less the entirety of the predicate content, and are in that regard similar to semantically incorporated nouns per the definition of Wittenberg & Trotzke (2021). The copular which introduces a predicate NP merely acts as a bearer of grammatical information, such as tense and number:

(2) I am *a teacher*.
PR.1.SG

Indeed, predicate NPs are "characteristically non-referential" as they do not introduce a new entity to the discourse; rather, they only provide further information about one already established (Huddleston & Pullum, 2002, p. 254). Their non-referentiality, or unavailability for anaphoric reference, could supposedly be linked to their high degree of semantic incorporation. However, their morphosyntactic form is identical to that of indefinite referential NPs (cf. *I saw* a teacher), which in turn are unproblematic in terms of discourse accessibility. With respect to what previous research has found about 'superficial' morphosyntactic factors (e.g. the presence of a determiner) affecting discourse salience, the question remains as to what extent predicate NPs are regarded possible antecedents during online reference resolution.

The effects of semantic incorporation and morphosyntax on discourse salience have been studied using offline measures (e.g. sentence-continuations and acceptability-ratings, see Wittenberg & Trotzke, 2021), but online measures are hitherto underrepresented in this field. Pronoun-resolution studies can say a great deal about how interlocutors store and retrieve discourse entities in their *mental model* for language comprehension purposes and linguistic problem solving. Moreover, brain-imaging techniques such as electroencephalography (EEG) can provide excellent insight into the various neurocognitive processes underlying linguistic processing in real-time. Whether a discourse entity gives rise to the presence or absence of specific Event-Related Potential (ERP) components (that is, fluctuations in electric potentials

as response to linguistic stimuli) can thus be used as a diagnostic of its discourse salience (Luck, 2014).

Predicate NPs, notwithstanding their theoretical inability to establish co-reference with an anaphor, have not been studied as to how they are stored in the mental models as *potential* antecedents for anaphora. Therefore, this thesis both aims to fill this gap in research using an ERP experimental paradigm and thereby provide further insight into how semantic and morphosyntactic information interact and is stored, processed and accessed, online. The study is guided by the following research question(s):

RQ: To what extent are indefinite predicate NPs stored in the mental model as possible antecedents for anaphoric pronouns?

a. Does their morphosyntactic form and/or predicative meaning affect online pronoun resolution, as reflected by ERP patterns?

The thesis is structured as follows. Chapter 2 provides the relevant theoretical background, defines key concepts and outlines previous research. Chapter 3 presents the chosen methodology and how the investigation was carried out. The results are then presented in Chapter 4 and discussed in Chapter 5. The study and its main findings are concluded in Chapter 6.

2. Background

2.1 Mental Model Theory: Overview and Anaphoric Reference

Mental Model theory (Garnham, 2001; Johnson-Laird, 1980, 1981; van Dijk & Kintsch, 1983) builds on the concept of an "internal model of the world" that people construct during speech- and text comprehension (Johnson-Laird, 1980, p. 73). This model is thought to be built on the comprehender's[2] mental representations of the places, events, and situations introduced in the immediate linguistic environment, which can be employed for comprehension purposes (Garnham, 2001, p. 5; 19). The representations, in turn, may be based on real-world knowledge –on entities and situations that people can anchor in real-life events– but they can also be constructed largely on imaginary and fictional situations. Just like both listening and reading involve an incremental unfolding of incoming linguistic input, the construction of the mental model (sometimes: situation model) is thought to be incremental. New information is integrated with old as new context becomes available (Garnham, 2001, pp. 22-23; Haviland & Clark, 1974, pp. 512-513). In other words, the mental model and the representations therein are continuously updated, as the model represents a discourse-based 'possible world' in which information about entities, states and events in the current discourse is stored.

How previously stored information is activated and retrieved becomes particularly relevant during anaphoric referential processing (Johnson-Laird, 1981, p. 141). Once a mental representation has been formed, the model allows the comprehender to keep track of and store discourse entities for later retrieval, and the model is therefore thought to depend on working memory capacity to some extent (Garnham, 2001, p. 23; Kintsch & van Dijk, 1978, p. 368). If, however, the mental model does *not* contain a representation to which, for example, an anaphor can be mapped, anaphor resolution becomes difficult. In such cases, the comprehender might need to rely on a bridging inference, i.e. the process where a possible referent to the anaphor must be inferred or deduced, as no appropriate referent has been explicitly introduced. Bridging

[2] As general language comprehension is discussed here (not specifically reading- or listening based), 'comprehender' is used to broadly refer to a person computing meaning from incoming linguistic material (cf. *reader* for reading comprehension, *interlocutor* or *listener* for conversation-based or listening comprehension)

inferences may allow a comprehender to establish a referent outside the current discourse or linguistic domain, but such processes are correlated with increased processing load (Haviland & Clark, 1974, p. 518; Kintsch & van Dijk, 1978, p. 368; Johnson-Laird, 1981, p. 358; Garnham, 2001, p. 23). In contrast, a referent that is unambiguous and readily available for the resolution of an anaphoric pronoun is clearly preferred from an efficiency-of-processing viewpoint. The context in which an anaphor is introduced should thus preferably provide enough information (be it linguistic or situational) to make the referent of the anaphoric expression clear, aiding anaphor resolution (Garnham, 2001, p. 34). The different aspects of what makes an antecedent accessible, salient or 'given' enough to aid pronoun resolution are discussed in the following sections.

2.2 Anaphoric Reference and Reference Resolution Strategies

In theoretical linguistics, the term 'anaphora' has come to receive (at least) three distinct interpretations. Firstly, it can be used as an umbrella term for any instances in which a linguistic element depends on another for its interpretation; secondly, it can refer to a linguistic element which specifically depends on a previously occurring entity (an antecedent) for its interpretation (in contrast to cataphora; a *cataphor* receives its interpretation from a later-occurring referent); thirdly, it can denote a specific subset of nominal referential expressions, usually reflexives and reciprocals (e.g. *herself*, *each other*) (Chomsky, 1981; Huang, 2015, p. 674). In this thesis, I will use the terms 'anaphora' and 'anaphor' in alignment with a combination of the first and second senses; 'anaphora' refers to the overall theory on co-referentiality and linguistic dependency, and although 'anaphor' can refer to either a typical anaphor or a cataphor, I only discuss anaphoric backward-looking dependencies. Moreover, unless specified, anaphora is discussed in relation to nominal dependencies, rather than verbal. The following subsections provide a theoretical overview of anaphoric reference and introduce some prominent theories on anaphoric reference resolution.

2.2.1 Anaphoric Reference: an Overview

There exists a plethora of theories, studies and experiments on the notions of anaphora and pronominal reference. While some linguistic scholars have provided extensive accounts of the

(morpho-)syntactic rules of anaphora and co-reference (e.g. Chomsky's (1981) *Lectures on Government and Binding*), others have focused more on semantic (e.g. Reinhart, 1983) and pragmatic (e.g. Huang, 2000) factors that affect how anaphoric expressions receive their interpretation. With regard to pronominal reference, Chomsky's (1981) Binding theory[3] claims that a pronoun must be "free in its governing category" (Principle B; Chomsky, 1981, p. 188); this usually means that a pronoun cannot be coreferential with an antecedent in the immediate clause or phrase that contains it:

(3) Daphne$_i$ heard that [the woman$_j$ loves her$_{i/*j}$][4]

As the Binding Theory framework merely formulates that pronouns are dependent on either a preceding or succeeding nominal referent outside the 'local' linguistic domain, pronominal forms cannot depend solely on syntax for their interpretation. Instead, pronominal anaphoric reference is usually considered "a discourse phenomenon", or a psychological phenomenon (Garnham, 2001, p. 55). This means that the co-reference between an anaphor and a possible antecedent must be established either by linguistic means provided in a preceding clause somewhere earlier in the discourse, or outside the linguistic context altogether (via e.g. bridging inferences or situational cues). Usually, however, there is an interplay of both morphosyntactic, semantic and pragmatic/situational factors when pronominal reference is resolved. Garnham (2001) discusses some of these factors in light of Accessibility theory (Ariel, 1990), which is outlined in the next subsection.

2.2.2 Accessibility Theory and the Givenness Hierarchy

Ariel (1990) links the syntactic factors that govern anaphor licensing to context-related and pragmatic factors. Introducing Accessibility Theory, she states that the form of an anaphoric expression is linked to the ease with which a coreferential antecedent can be retrieved from the mental model (i.e. how salient the antecedent is), and suggests that anaphoric expressions can be labelled accordingly as to the degree of Accessibility they mark (Ariel, 1990, p. 16). In other words, a dynamic relationship between the antecedent and anaphor in terms of degree of

[3] The principles of Binding theory have received some criticism in light of both cross-linguistic evidence (see Huang, 2000; 2015) and evidence in English (see e.g. Kaiser et al., 2009). Nevertheless, it provides a good-enough account on the general principles underlying pronoun licensing in (at least) English, wherefore it is referenced here.

[4] Subscripts mark potential co-referential dependencies between nominal entities; the asterisk marks illicit co-reference; square brackets mark the subordinate clause containing the pronoun.

Accessibility is assumed; the form of the referential expression will match the Accessibility of the intended antecedent (Ariel, 1990, p. 17; 69). Accessibility, in turn, is dependent on several linguistic and contextual factors, including 'Distance', 'Competition', 'Saliency' and 'Unity' (from Ariel, 1990, pp. 28-29):

(4) *Factors affecting the Accessibility of an antecedent*

a. Distance: The distance between the antecedent and the anaphor (relevant to subsequent mentions only).
b. Competition: The number of competitors on the role of antecedent.
c. Saliency: The antecedent being a salient referent, mainly whether it is a topic or a non-topic.
d. Unity: The antecedent being within vs. without the same frame/world/point of view/segment or paragraph as the anaphor.

Building on these factors, antecedents can be considered highly accessible to the comprehender if they have been recently mentioned (4a), have no competing potential antecedents (4b), constitute the discourse/sentence topic (4c) and appear within the same sentence/paragraph as the anaphor (4d), and they will then be referenced using markers that signal the corresponding degree of Accessibility. The hierarchical order in which different referential expressions can be organised as to the degree of Accessibility they mark is then proposed as follows:

LOW ACCESSIBILITY

Full name + Modifier
Full name
Long definite description
Short definite description
Last name
First name
Distal demonstrative (+ Modifier)
Proximal demonstrative (+ Modifier)
Stressed pronouns + Gesture
Stressed pronouns
Unstressed pronouns
Zeros

HIGH ACCESSIBILITY

Figure 1. The Accessibility Hierarchy (Ariel, 1991, p. 449)

The 'emptier' the anaphoric/referential expression is, the more it signals that the antecedent is highly accessible in the discourse, i.e. easy to single out as the evident antecedent. For example, a stressed pronoun contains more information than an unstressed one (as it contains a prosodic marker for salience), and an unstressed pronoun, in turn, contains more information than a zero-form. Although not highest in the hierarchy, unstressed personal pronouns are still regarded markers of high accessibility – that is, the information the pronoun provides (in English: information about number, person and possibly gender) should be enough for the comprehender to identify and retrieve the co-referential antecedent with little effort.

Gundel, Hedberg and Zacharski (1993) suppose a hierarchical relationship between the form of a referring expression and the degree of givenness it marks:

in focus >	**activated** >	**familiar** >	**uniquely identifiable** >	**referential** >	**type identifiable**
it	*HE*[5], *this, that, this* N[6]	*that* N	*the* N	indefinite *this* N	*a* N

Figure 2. The Givenness Hierarchy (Gundel et al., 1993, pp. 275, 284)

Similar to that of Ariel (1990), their hierarchy aims to describe different referential expressions' "cognitive statuses" –i.e. the "information about location in memory and attention state" that the expression marks (Gundel et al., 1993, p. 274). The hierarchy is considered implicational, meaning that the cognitive status of a referential expression leftmost in the scale implies all cognitive statuses to the right of it. A referent that is cognitively 'activated' should supposedly also be 'familiar' and 'uniquely identifiable', and so on (Gundel et al., 1993, p. 276). The referent should thus be available in the comprehender's short-term memory after having been retrieved by either linguistic or extralinguistic cues prior to mention (Gundel et al., 1993, p. 278). A referent 'in focus', highest in the hierarchy, is likely to be a highly salient and accessible discourse entity, often the current utterance topic or a topic higher up in the discourse (Gundel et al., 1993, p. 279).

2.2.3 Grammatical Role Parallelism and Implicit Causality

While different referential expressions can signal to the comprehender how easy an antecedent can be retrieved from the mental model, the linguistic factors that actually influence the saliency or accessibility of an antecedent have hitherto only been briefly mentioned. Von Heusinger & Schumacher (2019) mention some of the "prominence-lending cues" that affect the saliency of discourse entities that compete for the role of antecedent in pronominal reference (p. 119). For example, grammatical functions have been found to affect discourse prominence: subjects are usually more prominent than objects, topics are usually more prominent than non-topics and Agents[7] are usually more prominent than Patients (von Heusinger & Schumacher, 2019, p. 119). Moreover, research has found that in cases where two

[5] Stressed personal pronoun (Gundel et al., 1993, p. 278)
[6] Definite demonstrative determiner (as contrasted to indefinite *this*) (Gundel et al., 1993, p. 276)
[7] As in thematic role.

entities compete for the role of antecedent, 'grammatical role parallelism' is likely to affect co-reference establishment. That is, an ambiguous pronoun in a co-/subordinated clause is often assigned an antecedent that shares its grammatical role, even in languages with relatively free word-order (Hall & Yoshida, 2020, p. 301; Sauermann & Gagarina, 2017, p. 7; Smyth, 1994, p. 219).

However, grammatical role parallelism does not always apply to online pronoun resolution. For example, Caramazza et al. (1977) found that the establishment of co-reference between competing antecedents can be dependent on implicit causality, a semantic propensity of (certain) verbs. Consider the following examples (from Caramazza et al., 1977, p. 601):

(5) Jane hit Mary because she had stolen a tennis racket.
(6) Jane angered Mary because she had stolen a tennis racket.

Although the pronouns in both sentences are syntactically ambiguous in that two discourse entities match the pronoun in terms of person (third), number (singular) and gender (female), Caramazza et al. observed that the pronoun in (5) is more likely to be coreferential with *Mary* than *Jane*, whereas *Jane* is the likely antecedent in (6). In (5), comprehenders are likely to imagine an explanation for what Mary might have done to provoke a hit from Jane, rather than assume that Jane hit Mary due to ill-will. That is, comprehenders are likely to predict that the continuation of the sentence is about the causer of the event, and the verb is thus said to carry *implicit causality* (henceforth: IC). Although the pronoun is the grammatical subject in the subordinate clause, the IC of the verb makes the object in the matrix clause more salient in (5); hence, the object is more accessible for subsequent anaphoric mention, despite the mismatch in grammatical function/role (Caramazza et al., 1977). These tendencies have been replicated in ERP studies; for example, Van Berkum et al. (2007) found that a pronoun that was incongruous with the preferred antecedent of a verb with IC rendered a short-lived P600 effect, which the authors took to indicate processing difficulties, prediction error and/or the failure of resolving "proactive referential commitment" (Van Berkum et al., 2007, p. 167).

2.2.4 Establishing Co-reference: Morphological and Semantic Cues

Although grammatical role parallelism and IC both contribute to the resolution of a potentially ambiguous pronoun, there are undoubtedly other linguistic cues that aid pronoun resolution. For example, many languages make use of morphological markings, by which antecedents can

easily be disambiguated. In a language like Spanish, most nouns carry inherent information about gender (either masculine or feminine), which is reflected in morphological features on the noun and any possible determiner or modifier that precedes or follows it (Green, 1988, p. 93-95; Carreiras et al., 1996, p. 643; 647). Some animate/human nouns are variable in terms of inherent gender, but a preceding article can disambiguate whether the referent is female or male (e.g. *el*/*la cliente* 'the.M/the.F customer'). Any anaphoric expression must then agree with its antecedent in terms of number and gender:

(7)	*Este*	*es*	*un*	*libro.*	*Es*	*el*
	DEM.3.SG.M	be.3.SG.PR	INDEF.SG.M	book.M	be.3.SG.PR	DEF.SG.M
	mío.					
	POSS.SG.M					

'This is a book. It is mine.'

The possessive pronoun *mío* and the definite article *el* correspond morphologically to the gender and number of the noun they are coreferential with, namely *libro*.

In English, nouns can be morphologically marked for semantic gender (e.g. feminine marking suffix *-ess*, as in *countess*, *waitress*, *lioness*), but most nouns carry no such overt marking. Notwithstanding the lack of gender-marking morphology, nouns can still be 'definitionally' gendered (Osterhout et al., 1997) – that is, the meaning of the word (typically that which would be listed in a lexicon entry) encodes explicit information about gender and it can be regarded lexically unambiguous in that regard (Saeed, 2016, p. 56; 58; 259). This is true for most English kinship terms (e.g. *aunt, father, sister*…) as well as some nouns denoting people/individuals. Such semantically gendered nouns are thought to carry a lexical 'gender' component (cf. Componential analysis; Saeed, 2016, p. 259). However, most English role-nouns (e.g. *nurse, doctor, secretary*) are neither morphologically nor semantically encoded for gender. Nevertheless, they have been found to often be 'stereotypically' marked for gender, a phenomenon that is further outlined in the next subsection.

2.2.5 Stereotypical Gender Effects on Reference Resolution

Many scholars have specifically focused on how societal norms and stereotypes regarding gender roles and representations affect the construction of a mental model, and how they affect anaphor resolution abilities (Carreiras et al., 1996; Duffy & Keir, 2004; Garnham et al., 2012;

Kennison & Trofe, 2003; Kreiner et al. 2008; Oakhill et al., 2005; Osterhout et al., 1997; Reali et al., 2015). Many English nouns, especially (but not exclusively) role nouns and nouns denoting human individuals, carry no semantic gender component or morphological gender marking. Instead, gender representations of such nouns are dependent on some sort of inference from probability or stereotype. In fact, gender stereotypes have been found to be present in children's mental representations of specific roles/professions (see Wilbourn & Kee, 2010) and employed instantly during language processing in order to make predictions about the upcoming discourse (Carreiras et al., 1996, p. 657).

Interestingly, stereotype information appears to be activated even in languages where gender is overtly (i.e. morphologically) marked. In a series of self-paced-reading experiments, Carreiras et al. (1996) found that both speakers of Spanish (a language with overt gender marking on all nouns, as mentioned above) and English (a language with usually no overt gender marking on role nouns) made use of gender stereotypes during online language processing. For Spanish-speakers, marking a noun for a gender counterintuitive to its stereotypical gender (e.g. *la carpintero* 'the [female] carpenter' instead of *el carpintero* 'the [male] carpenter') resulted in increased processing load (as reflected in increased reading times of the sentence containing the noun), compared to when the morphological marker and stereotypical gender matched (Carreiras et al., 1996, pp. 651; 657-658). A subsequent anaphoric pronoun, however, was processed with ease; the gender-marking definite article, providing updated information about the actual gender of the referent, thereby seemed to facilitate pronoun resolution in Spanish. For the English-speaking participants, however, an increase in processing load was not visible until the mismatched personal pronoun was encountered (e.g. *the carpenter … she*), as that was when the gender-mismatch first became apparent (Carreiras et al., 1996, p. 652). In sum, stereotypical gender bias influenced language comprehension for both Spanish and English speakers, albeit visible on different stages in the discourse.

The mismatch effect of gender stereotypicality on pronoun resolution and the immediate activation of stereotypical gender information found in Carreiras et al. (1996) have been found in other psycholinguistic experimental paradigms (e.g. word-based self-paced reading: Kennison & Trofe, 2003; lexical decision task: Oakhill et al., 2005; eye-tracking: Reali et al., 2015). Duffy and Keir (2004) also found a facilitative effect from context on pronoun resolution in an eye-tracking experiment. A context-providing clause in which the gender of a

role-noun was specified eliminated any mismatch-effects between an antecedent and the anaphoric pronoun. The authors took this to indicate that stereotype-based gender information is instantaneously employed during language comprehension, but that this stereotype can be overridden by context in which gender is specified (Duffy & Keir, 2004, p. 557). Consequently, the eliminated mismatch effect would indicate that the new information had been incorporated in the updated mental model. Similar conclusions were drawn by Kreiner et al. (2008) in an eye-tracking study. They found that introducing a gendered pronoun in a cataphoric reference construction (i.e. establishing the gender of the referent before it is encountered) led to increased processing load for nouns that have gender encoded lexically (either semantically or morphologically), but *not* for stereotypically gendered nouns (Kreiner et al., 2008, pp. 256; 257-258). To summarize, perceived or inferred gender of a referent (i.e. stereotypical gender) has been found to affect pronoun resolution abilities to the same extent as morphologically and/or semantically encoded gender, but context can inhibit mismatch effects.

2.2.6 Structurally Unavailable Referents: Agreement Attraction and Intrusion Effects

The previous subsection established that the mental representations of human referents usually contain information about the referent's gender, stereotypically or linguistically encoded. Moreover, it was also established how a referent's stereotypical gender is seemingly incorporated in the mental model immediately when a role-noun is processed and becomes readily available for later employment, for example when agreement with a subsequent anaphor needs to be established. These two observations become relevant for the discussion of the phenomenon known as Agreement Attraction.

In two eye-tracking experiments, Dillon et al. (2013) investigated how the morphosyntactic features of illicit antecedents impacted the resolving of two kinds of linguistic dependencies. Such an impact, commonly referred to as agreement attraction, had previously been observed in studies on both language production (e.g. Bock & Miller, 1991) and comprehension (e.g. Wagers et al., 2009), in that "having a syntactically illicit, but feature-matched NP [led] to processing facilitation" (Dillon et al., 2013, p. 87). Agreement attraction can be thought of as a morphosyntactic 'spill-over effect' that an intrusive NP has on a subsequent linguistic dependency formation, licensing it erroneously (Dillon et al., 2013, p. 87). In Dillon et al., the 'illicit antecedents' were intrusive NPs that dis-/agreed in number with either a reflexive pronoun (*himself* vs. *themselves* in (8)) or with the predicate verb (*was* vs.

were in (9)) (from Dillon et al., 2013, p. 89; italics added to mark intrusive NPs and reflexive pronoun/predicate verb):

(8) a. The new executive who oversaw *the middle managers* apparently doubted *himself* on most major decisions.

b. The new executive who oversaw *the middle managers* apparently doubted *themselves* on most major decisions.

(9) a. The new executive who oversaw *the middle managers* apparently *was* dishonest about the company's profits.

b. The new executive who oversaw *the middle managers* apparently *were* dishonest about the company's profits.

Interestingly, results only showed agreement attraction effects on illicit subject-verb agreement (Dillon et al., 2013, p. 88; 98). Similar results of no intrusion effects on reflexive licensing were found by Xiang et al. (2009) in an ERP experiment (further outlined in section 2.4.1), whose results appear in stark contrast to previous self-paced-reading and eye-tracking experiments on reflexive binding that have shown clear attraction effects of gender-agreement between an intrusive NP and a reflexive pronoun (e.g. Badecker & Straub, 2002; Kennison, 2003).

Scholars like Kennison (2003) and Badecker and Straub (2002) have put forth evidence from behavioural studies that suggests that even antecedents that are structurally unavailable[8] may give rise to agreement-attraction effects during pronoun resolution. While Kennison concluded that "all entities, both structurally available and structurally unavailable, are included in the initial set of potential antecedents" (Kennison, 2003, p. 351), Badecker & Straub claim that this 'initial set' consists "only [of] highly salient or focused entities in the local discourse context" (Badecker & Straub, 2002, p. 764). Note, however, that their 'structurally unavailable' entities were actual referential NPs. Similarly, the intrusive NP in Dillon et al. (2013) (cf. (8) and (9) above) indeed carries real reference to an actual discourse entity. The current study, however, contrasts referential NPs with non-referential (predicate) equivalents, seeking to investigate how and to what extent predicative NPs are stored in the mental model as potential antecedents to a subsequent anaphor. If predicate NPs also are included in the 'initial set of antecedents' during online pronoun resolution, they should give

[8] Usually with regard to the principles of Binding Theory (Chomsky, 1981)

rise to illusory pronoun licensing or agreement attraction effects, along the reasoning of Kennison, and Badecker and Straub. The conflicting results of intrusion/agreement-attraction effects found in some, but not all, linguistic dependency formation processes, naturally call for more research in this field. Moreover, the discourse salience of predicate NPs, as reflected by their possibility to affect pronoun licensing, has (to the author's knowledge) hitherto not been researched in a pronoun-resolution paradigm. Aiming to fill this gap in linguistic research, the current study can thus give novel insight into how nominal predicates affect subsequent linguistic dependency formations.

2.3 Predication and copular verbs

As has been established, the mental model has the capability to store and aid retrieval of different discourse entities for subsequent reference. According to van Heusinger and Schumacher (2019), a discourse entity is "assessed relative to members of the same type" (p. 118), meaning that it competes with equivalents for the role of antecedent. For nominal anaphora, these entities are commonly NPs, either simple (e.g. proper nouns/names) or complex (e.g. noun and modifier), but not all noun phrases introduce discourse referents. In this section, the English verb phrase and some predication strategies are introduced, with a specific focus on the contrast between nominal complements of referential and predicative status.

2.3.1 Verbal Predication

Typically, the verb is considered the centre, or head, of a clause and "is the syntactically most important element within it: properties of the verb determine what other kinds of elements are required or permitted" (Huddleston & Pullum, 2002, p. 50). A verb is classified with regard to the minimal number of arguments it requires in order to form a proposition (Adger, 2003, p. 78). A verb requiring only one argument is referred to as a one-place predicate and is normally regarded intransitive in that the only argument is usually also the grammatical subject. A predicate verb in transitive constructions, however, takes one or two additional internal arguments, situated within the verb phrase domain – commonly direct and/or indirect objects. (Adger, 2003, p. 78). The semantic relationship between a predicate and its argument(s) is

realised via Theta (θ) roles (also: Thematic roles). The predicator (the head of the predicate construction) is said to be "bearer of a theta-grid", meaning that it has theta-roles to assign the arguments it requires (Roy, 2013, p. 5). With this brief overview of basic verbal predication in English, let us turn to the case of non-verbal predication, with a specific focus on nominal predicates (Roy, 2013, p. 7).

2.3.2 Nominal Predication

Superficially, there appears to be little structural difference between a transitive verb (*met* in (10a)) and a copular verb (*is* in (10b)). Consider and compare the two structures in (10):

(10) a. Sarah met a nurse.

b. Sarah is a nurse.

Both sentences unfold incrementally in the pattern of NP [VP [NP]] and both verbs at least appear to take identical complements as internal arguments (*a nurse*). However, the nominal complement in a copular construction does not introduce a new referent to the discourse; in (10b), *a nurse* merely assigns a property to the already-introduced subject (Huddleston & Pullum, 2002, p. 53, 217). In copular constructions, the nominal complement is commonly considered the actual predicator of the sentence (Roy, 2013, pp. 8-9). Although Huddleston & Pullum (2002) suppose that the copular verb to some extent is the predicator in copular constructions, they still seem to align with Roy (2013), who, building on scholars such as Bowers (1993) and Adger & Ramchand (2003)[9], claims that "the copula itself [does not play] any role in mediating the relationship between the predicate and its subject" (Roy, 2013, p. 8). The theta-role assigning predicator is thus not the verb *be* but rather the predicative complement[10]. Hence, the similarity between (10a) and (10b) is only superficial; the deep syntactic structure of nominal predicates is, in fact, distinct from that of verbal predicates (Roy, 2013, p. 12). An outline of the syntactic structure of copular verbs is outside the scope of this paper; however, this study aligns with both Roy (2013) and Huddleston & Pullum (2002) in the view that a copular verb introduces a predicative complement that assigns a property or

[9] Proponents of the so-called PredP-hypothesis (see Roy, 2013, for an overview).
[10] See Roy, 2013, p. 12.

state to the subject[11] whereas complements of transitive verbs introduce new, or reference old, discourse entities.

Although the term 'copular' is broadly used for any verb whose complement links predicatively to a predicand[12], *be* is sometimes regarded the only 'true' copula in English. Being more or less semantically empty, the verb does not bear any theta-grid or inherent argument structure, and is traditionally regarded as merely an occupant of the verbal position in the clause, bearing inflection for tense (Butler, 2003, p. 269; Huddleston & Pullum, 2002, p. 218; Roy, 2013, pp. 11-12). Other structures that feature predicative complements are preferably labelled 'complex-intransitive'[13] according to Huddleston and Pullum (2002, p. 53), but the terms 'semi-copular' or 'pseudo-copular' appear in other scholarly material (e.g. Butler, 2003, p. 269). This study does not focus solely on structures featuring the copular *be*, but rather any verbal construction for which a nominal complement is considered (subject) predicative (e.g. *become*, *work as*, etc.; Huddleston & Pullum, 2002, pp. 271-272). For that reason, 'copular' will here be used in the wider sense, simply denoting a verb that takes a predicative complement. In cases where a distinction between copular verbs with or without semantic content is needed, *semi-copular* will refer to verbs that are not necessarily semantically empty.

2.3.3 Noun Incorporation

There are some constructions that seemingly fall somewhere between transitive and copular constructions. Noun incorporation, for example, is a phenomenon (especially) common in some agglutinative/polysynthetic languages, in which nominal elements are morphologically incorporated into larger predicate constructions together with the predicate verb. Although such incorporation occurs through morphosyntactic operations, the phenomenon is referred to as 'semantic incorporation', in that the incorporated NP "contribute[s] to the bulk of the predicate meaning" (Wittenberg & Trotzke, 2021, p. 88). In West Greenlandic, noun incorporation is a productive predication strategy, in which the incorporated noun forms a complex predicate construction together with the incorporating verb (van Geenhoven, 2002, p. 760, 762).

[11] Or object, in object-predicative constructions (Huddleston & Pullum, 2002, p. 217)
[12] The entity which is predicated (Aarts, 2014).
[13] Although not transitive, the copular construction in still more complex than a typical intransitive clause (hence 'complex-intransitive').

Wittenberg and Trotzke (2021) drew a parallel between West Greenlandic semantic incorporation and nouns appearing in English light verb constructions. Building on theories of discourse salience and pronominal reference (e.g. Ariel, 1990 & Gundel et al., 1993), they set out to investigate the discourse prominence of nouns in light verb constructions, such as *a kick* in (11), as compared to other nominal complements (from Wittenberg & Trotzke, 2021, p. 91):

(11) The teenager gave *a kick* to his rival.

(12) The teenager gave *a note* to his rival.

As previous studies had found that the semantic status of a referent affects the pronominal form with which it is retrieved (e.g. Scholten & Aguilar-Guevara, 2010), Wittenberg & Trotzke predicted that nouns in light-verb constructions would consistently be referenced with markers of low(er) accessibility; in this case, they hypothesized that in choosing between a pronoun and repeating the whole NP, participants would be inclined to repeat the whole NP (Wittenberg & Trotzke, 2021, p. 91, 96). In the end, they found that semantic incorporation did not modulate the frequency with which nouns in light-verb constructions were referenced with a pronoun. Instead, they noted that the morphosyntactic form of the antecedent (i.e., whether it was realised as a full determiner phrase or not) was a more prominent factor affecting anaphor form, than the degree of semantic incorporation (Wittenberg & Trotzke, 2021, pp. 93-94).

Wittenberg and Trotzke's (2021) predictions on the degraded discourse saliency of nouns incorporated in complex predicate constructions can be extended to that of nominal predicates. Nominal predicates are much like light-verb constructions, in that the complex predicate includes a nominal element which is virtually inseparable from the semantic interpretation of the predicate expression. Moreover, the verb is not only "semantically bleached", as for light-verb constructions (Wittenberg & Trotzke, 2021, p. 89); instead, the verb is more or less semantically empty, and it is the nominal complement which constitutes the better part of the predicative meaning. Previous studies on the discourse prominence and cognitive status of nominals in complex predicates have predominantly employed offline behavioural measures (e.g. sentence-completion: Scholten & Aguilar-Guevara, 2010, Wittenberg & Trotzke, 2021), leaving online measures unexplored and underrepresented in this particular field. However, since the late 20th century, online methods such as the electroencephalography (EEG) technique has been used for psycholinguistic research in general and pronoun-resolution studies in particular. It could thus serve as an appropriate method for expanding the view on how the predicate status of nouns in nominal predicate

constructions affects how they are stored in the mental model, and their ability to support anaphoric reference resolution measures. The next section introduces the technical aspects of this methodology and discusses how it previously has been applied in studies on discourse prominence and referential processing.

2.4 Electroencephalography, Event-Related Potentials and Referential Processing

In linguistic research, the EEG technique has been used to investigate different cognitive processes involved in language processing. By attaching electrodes to the scalp, the researcher can detect and record different shifts in electric potential that occur as a response to manipulated linguistic experimental stimuli at a certain electrode site. After averaging many experimental trials and subtracting activation deviations from a baseline pattern, the modulation of neural activity in response to linguistic stimuli can thus be argued to be indicative of certain linguistic, cognitive processes. Researchers have thereby been able to identify a number of Event-Related Potentials (ERPs) related to specific linguistic processing processes (Kemmerer, 2015, p. 61). As EEG data can be recorded continuously and shifts in electrical potentials can be registered with millisecond precision, ERP responses can thus give insight into language comprehension strategies in real time (Luck, 2014, p. 12; Kemmerer, 2015, pp. 60-61).

2.4.1 The P600 Component

A classic, but complex, ERP component that has been linked to different kinds of syntactic computation processes is the P600 component, reflected as a late positive deflection between 500-1000 ms after stimulus onset (usually peaking around 600 ms after stimulus onset) over the mid/posterior parts of the scalp (Swaab et al., 2011, p. 419). Although the component has historically been viewed as a response diagnostic of syntactic violation detection/processing

(e.g. Hagoort et al., 1993[14]; Osterhout & Mobley, 1995, p. 748), it has also been found for syntactically complex or ambiguous constructions (such as the processing of long-distance dependencies or garden-path sentences), and the component is therefore thought to be indicative of syntactic/structural processing in general (Gouvea et al., 2010, p. 155, 172; Kaan et al., 2000; Kaan & Swaab, 2003; Osterhout & Holcomb, 1992; Swaab et al., 2011, p. 419, 420). For example, Gosselke Berthelsen et al. (2020) found that the presence of a linguistic component triggering morphosyntactic processing (in their case, grammatical tone) rendered a P600 response against a morphosyntactically empty control condition (Gosselke Berthelsen et al., 2020, pp. 8). Moreover, the component has also been found for *semantically* anomalous constructions (e.g. violations of typical semantic thematic roles), further challenging the view on the P600 as an exclusively syntax-related response (for a review and discussion of linguistic factors modulating the so-called 'semantic P600', see Bornkessel-Schlesewsky & Schlesewsky, 2008; Kuperberg et al., 2003, p. 127; van Herten et al., 2005, p. 252). Drawing from such results, some scholars have thus come to view the P600 as a neurophysiological response to 'integration difficulties', reflecting a general updating of the mental model or serving as an indication of a difficulty of integrating an entity into the current discourse for various reasons, not exclusively syntax-related (Bornkessel-Schlesewsky & Schlesewsky, 2008; Burkhardt, 2007, p. 1853; 1854; Heinat & Klingvall, 2020, pp. 2, 5; Steinhauer et al., 2010, p. 1536). Indeed, Delogu et al. (2019), among others, have found supporting evidence for this view, in that discourse entities that were syntactically correct but difficult to fit into the current discourse (by being implausible in the semantic context) gave rise to a P600-like response, as compared to a baseline and a control condition (Delogu et al., 2019, p. 6; 8; 10).

Gouvea et al. (2010) discuss how the P600 response is likely to not only vary with regard to amplitude, but also latency and duration, depending on what type of dependency formation is computed. Building on the results of both their own and previous studies, they suggest that "the P600 amplitude and duration directly reflect structure-building (and dismantling) operations, whereas the retrieval processes that are needed to initiate structure building are reflected only in the onset latency of the P600" (Gouvea et al., 2010, pp. 175, 183-184). In the context of pronoun resolution computations, this suggests that the more difficult

[14] However, they discuss the component under another name, namely the Syntactic Positive Shift (SPS).

the antecedent is to retrieve and employ for pronoun resolution, the more delayed should the P600 onset be (Gouvea et al., 2010, p. 173).

Xiang et al. (2009) used EEG to study and compare intrusion effects in two different grammatical dependency computations, namely NPI[15]-licensing and referential (reflexive) resolution. In their study, unlicensed NPIs gave rise to a prominent P600 effect, but intrusive licensors attenuated the component amplitude (Xiang et al., 2009, p. 48, 51). In contrast, the ERP responses to illicitly licensed reflexives gave rise to the same P600-like response found for the control incongruent condition (Xiang et al., 2009, p. 50, 51). This absence of intrusion effects for the resolution of a reflexive pronoun suggests that the readers detected and processed the syntactic dependency fast upon encountering the reflexive pronoun, thereby detecting the antecedent as inappropriate despite the match in gender and number features (Xiang et al., 2009, pp. 51-52).

The current study makes use of the previous observations that the P600 component indicates difficulties in resolving syntactic dependencies (e.g. Gouvea et al., 2010), can serve as a diagnostic of parsing difficulty and reflects an update/restructuring of the mental model (Burkhardt, 2007; Delogu et al., 2019) and can be modulated by intrusive licensors (Xiang et al., 2009, pp. 45-46). Additionally, it operates against the backdrop of previous EEG-based research on pronoun resolution, which, *inter alia*, has found that antecedent-anaphor gender mismatches (both definitional and stereotypical) give rise to P600 effects at the onset of an anaphoric pronoun, reflecting pronoun resolution failure (Nieuwland, 2014, p. 2; Osterhout et al., 1997, p. 276, 278; Osterhout & Mobley, 1995, p. 753, 759; Van Berkum, 2004, as cited by Barkley et al., 2015, p. 146; Van Berkum et al., 2004, as cited by Van Berkum et al., 2007, p. 162). Moreover, the study builds on observations related to another ERP component found for morphosyntactic agreement computations, namely the Left Anterior Negativity (LAN) component.

2.4.2 The Left Anterior Negativity (LAN) Component

The LAN is a negative-deflecting component visible at (as the name implies) left anterior regions, typically around 400 ms post stimulus onset (Barkley et al., 2015, p. 144; Friederici et al., 1993, pp. 189-190; García-Sierra et al., 2021, p. 4; Kluender & Kutas, 1993; Neville et al.,

[15] Negative Polarity Item

1991; Swaab et al., 2011, p. 427). It is often observed alongside a P600 response in a so-called biphasic pattern, and is generally suggested to be "indicative of processing difficulty, whether this is caused by an ungrammaticality, simply a grammatically unexpected continuation, or storage or retrieval from working memory in complex constructions" (Kaan & Swaab, 2003, p. 633, 626). Much like the P600, however, the LAN is not only indicative of morphosyntactic violation detection, but rather morphosyntactic processing in general (again, see Gosselke Berthelsen et al., 2020). For example, Barkley et al. (2015) explicitly link the component to retrieval processes, and suggest that it reflects the backward-looking process in which previously parsed linguistic material is searched for the appropriate antecedent to an anaphoric element (the 'back-association hypothesis'; Barkley et al., 2015, p. 145; see also Kluender & Kutas, 1993). The potential presence of a LAN effect would then indicate that a backward-dependency computation is attempted, and the amplitude of such an effect would reflect the ease with which it is computed (Barkley et al., 2015; García-Sierra et al., 2021, pp. 4, 11-12). In referential processing, the LAN effect has also been found to be modulated by violations of antecedent-anaphor gender congruency, including congruencies building on the stereotypical gender of the antecedent (King & Kutas, 1998, as cited in Federmeier et al., 2003, p. 155). The presence of anterior negativities (not necessarily left-distributed) is also commonly linked to working memory load and structural/morphosyntactic prediction (Coulson et al., 1998, p. 51; Federmeier et al., 2003, pp. 153-154, 156; Gouvea et al., 2010, p. 179; Kluender & Kutas, 1993, pp. 205-206).

However, previous research reports the biphasic LAN/P600 pattern with varying frequency. Specifically, it appears as if the P600 does not always appear alongside any LAN effect (see e.g. Coulson et al., 1998, p. 28; Federmeier et al., 2003, p. 155; Swaab et al., 2011, p. 428; Tanner & van Hell, 2014). The sometimes-absent LAN has thus made scholars question whether the effect could be an artefact of EEG referencing and/or averaging processes, subject-variability (in that only some participants show a biphasic LAN-P600 pattern for morphosyntactic violation processing), or component overlap (Caffarra et al., 2019, pp. 9-10; Molinaro et al., 2011; see also Tanner, 2015, for a commentary on these issues). Nonetheless, the LAN *has* also been reported in isolation without any subsequent P600. For example, Ciaccio et al. (2023) found the 'classic' biphasic LAN/P600 response to a morphosyntactic violation when EEG data was averaged over all participants. However, they also found that participants could be grouped based on whether their individual ERP patterns showed

negativity- (i.e. LAN[16]) or positivity-dominance (i.e. P600). The authors thus found evidence for the two components appearing in isolation, and argued for a dissociation between what types of linguistic processing they reflect (Ciaccio et al., 2023, pp. 12-13). At least to some extent, the LAN thus seems to index the above-mentioned processes of morphosyntactic processing, working memory load and/or backward-dependency computations (see e.g. Tanner, 2015, pp. 154-155).

2.5 The Current Study

The theoretical background and previous research on pronoun resolution can be summarised as follows. First, comprehenders construct a mental model during language comprehension, in which mental representations of discourse entities and referents are created and stored, dependent on working memory capacity (Garnham, 2001; Johnson-Laird, 1980, 1981; van Dijk & Kintsch, 1983). Second, entities that are encountered during discourse processing can be anaphorically referenced, and an anaphoric expression usually takes a form that is in alignment with the discourse accessibility, saliency or givenness of the entity it references (Ariel, 1990; Gundel et al., 1993). Third, although English role nouns are usually not coded for gender by morphological means, information about definitional and stereotypical gender of role nouns alike has been found to be activated and used for anaphoric resolution purposes to similar extents (Carreiras et al., 1996; Duffy & Keir, 2004; Garnham et al., 2012; Kennison & Trofe, 2003; Kreiner et al. 2008; Oakhill et al., 2005; Osterhout et al., 1997; Reali et al., 2015). Fourth, discourse entities that are syntactically unavailable for subsequent anaphoric mention can nonetheless illusorily license an anaphor (Badecker & Straub, 2002; Kennison, 2002), but not always and not necessarily in all dependency-formations (Xiang, 2009; Dillon et al., 2013). Fifth, online dependency computations have successfully been studied using EEG, for which the P600 and LAN components have been found indicative of different cognitive processes and operations involved in anaphor resolution (Barkley et al., 2015; Burkhardt, 2007; Delogu et al., 2019; García-Sierra et al., 2021; Gouvea et al., 2010; Kaan & Swaab, 2003; Nieuwland, 2014; Osterhout et al., 1997; Osterhout & Mobley, 1995; Van Berkum, 2004).

[16] Or N400.

Finally, the morphosyntactic form of semantically incorporated nouns (specifically whether they appear alongside a determiner) has been found to modulate their discourse salience (Wittenberg & Trotzke, 2021). Although predicate NPs are non-referential (Huddleston & Pullum, 2002; Roy, 2013), their morphosyntactic form is identical to indefinite referential NPs. Consequently, they could be hypothesized to give rise to agreement attraction effects, if the comprehender parses them as NPs with which co-reference with a subsequent anaphor could possibly be established (cf. Kennison, 2003). Based on the theoretical background outlined above, this study sets out to test the following hypotheses:

Hypotheses

H1. Predicate NPs are accessible as potential antecedents for anaphora in the mental model, due to the morphosyntactic form that they share with referential NPs.

H2. Predicate NPs are not accessible as potential antecedents for anaphora in the mental model, due to their predicative function.

The current study makes use of a semantic gender mismatch and pronoun resolution paradigm, explained in detail in Chapter 3. The test implications of the above two hypotheses, however, can be explained with reference to an example of an experimental item:

Experimental condition	**Sentence**
REF-MATCH (baseline)	a. Henry met **a nurse** at the hospital and **she** laughed a lot.
REF-MISMATCH	b. Henry met **a surgeon** at the hospital and **she** laughed a lot.
PRED-MATCH	c. Henry was **a nurse** at the hospital and **she** laughed a lot.
PRED-MISMATCH	d. Henry was **a surgeon** at the hospital and **she** laughed a lot.

Table 1. An experimental item in all conditions. PRED = Predicate complement NP; REF = Referential complement NP; MIS-/MATCH = gender feature mis-/match between the complement NP and anaphoric pronoun.

If a predicate NP is considered a potential antecedent during pronoun resolution (H1), we can expect:

(13) *Test implications for H1*

i. similar differences in LAN/P600 amplitude between REF-MISMATCH relative to REF-MATCH, as for PRED-MISMATCH relative to PRED-MATCH; specifically, lower component amplitudes for PRED-MATCH and REF-MATCH than their MISMATCH equivalents.

Overall, H1 predicts that a predicate NP will behave similarly to a referential one during online pronoun resolution. If we assume that gender congruency will affect the ease of association between the anaphor (*she* in Table 1) and the intrusive NP (*a nurse*/*a surgeon*) for a referential NP, the same effect should thus be assumed for a predicate NP. (13i) therefore reflects that an intrusive predicate NP that matches a subsequent anaphoric pronoun in terms of gender will give rise to agreement attraction effects and illusorily license the anaphor. This is realised as significant differences in EEG amplitude between the PRED-MATCH and MISMATCH conditions. Based on previous research, (13i) also reflects that the pronoun in the REF-MISMATCH condition should be more difficult to integrate and associate with the referential NP than the MATCH version, given the gender incongruency. The graph in Figure 3 can illustrate the expected ERP amplitude contrast between conditions, as predicted by H1. It shows how the difference between MATCH and MISMATCH in the REF condition should be similar to that of MATCH and MISMATCH in the PRED condition:

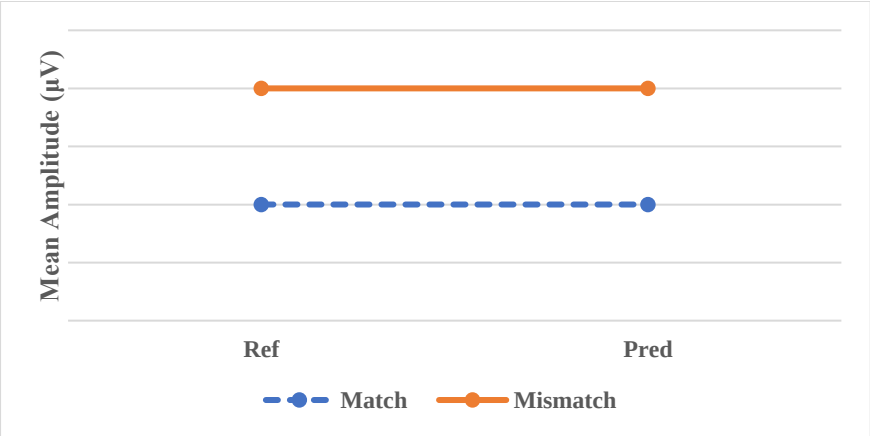

Figure 3. The expected effects of NP-TYPE and NP-PRONOUN GENDER MATCH on mean amplitude (H1)

In sum, H1 expects a main effect of gender congruency between the intrusive NP and the pronoun within both the predicate NP and referential NP conditions, but does not assume any significant effects on the EEG patterns based on NP-Type.

If a predicate NP is not considered a potential antecedent to an anaphoric pronoun due to its predicative function (H2), the predicate NP should quickly be realised as an impossible antecedent, regardless of any potential gender-feature match between the NP and the subsequent pronoun. The implications of H2 can thus be phrased as follows:

(14) *Test implications for H2*

i. significant differences in P600 and LAN amplitudes between the PRED-MATCH, PRED-MISMATCH and REF-MISMATCH conditions relative to REF-MATCH; specifically, higher component amplitudes for both PRED conditions and the REF-MISMATCH condition than the REF-MATCH condition.

ii. no significant difference in LAN/P600 amplitude between the PRED conditions and REF-MISMATCH.

iii. no significant difference in LAN/P600 amplitude between PRED-MATCH and PRED-MISMATCH.

(14i) and (14ii) reflect that the pronouns in both PRED conditions should be more difficult to integrate into the discourse than the pronoun in the REF-MATCH condition. They also specify that there should be no significant difference in component amplitudes for the PRED conditions and the REF-MISMATCH condition, as they all contain illicit antecedents. (14iii) further specifies that a predicate NP should fail to give rise to any illusory pronoun licensing, regardless of a potential gender match between the predicate NP and the subsequent anaphor. The expected ERP amplitude contrast between conditions can thus be exemplified as in the following graph:

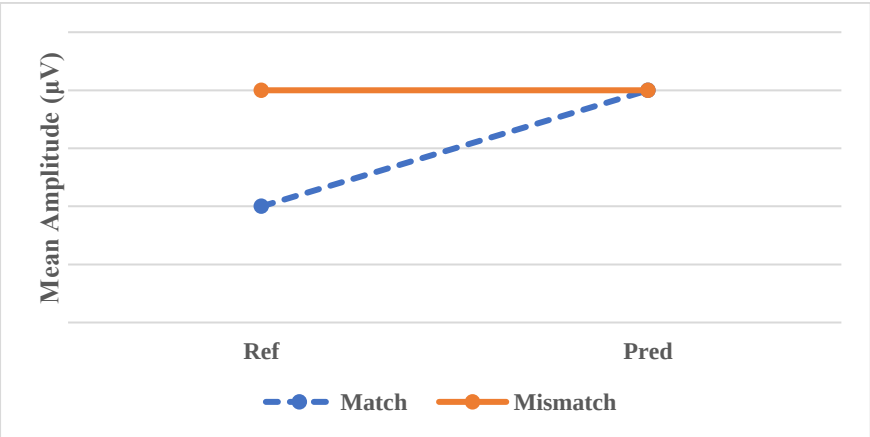

Figure 4. The expected effects of NP-TYPE and NP-PRONOUN GENDER MATCH on mean amplitude (H2)

In sum, H2 predicts that there should be an interaction of factors and their effects on ERP amplitudes; in other words, the hypothesis predicts that the gender mismatch factor should only affect ERP amplitudes in the condition with a referential NP, but not in the predicate-NP condition.

3. Materials and Method

3.1 Participants

Twenty-five self-reported native speakers of English[17] (F=15, N-B[18]=1), aged between 20 and 30 (*M*=24.36, *SD*=2.89), volunteered to participate in the study. Five participants were bilingual but had grown up speaking English in an English-speaking environment from an early age[19]. All participants reported being right-handed, having no diagnosed neurodevelopmental divergences or reading/writing difficulties, and had normal/corrected-to-normal vision. The participants were recruited through posters, social media, acquaintances, and word-of-mouth. Before the experiment, participants received general information about the purpose of the study, detailed information about the data collection procedure and how the collected data would be stored and used. They then gave informed consent by signing a consent form and were ensured that they could end their participation at any time without penalty. As compensation for their participation, the participants received a cinema ticket voucher, worth approximately 150SEK.

3.2 Materials

The experimental materials consisted of 160 items, manipulated with respect to two factors with two levels each: NP-TYPE (REF vs. PRED) and NP-PRONOUN GENDER MATCH (MATCH vs. MISMATCH). All items thus appeared in four conditions and were distributed equally across four lists in a 2x2 Latin Square design. All items were constructed similarly in terms of structure: the first clause was introduced by a proper noun in subject position (M = 80, F = 80) (*Henry* in Table 1); the clause continued with a verb phrase taking either a referential or predicate NP as complement (*saw/was a nurse/a surgeon* in Table 1); the first clause was then

[17] Varieties spoken: North American (*n*=11), Australian (*n*=3), British (*n*=6), Indian (*n*=1), New Zealand (*n*=2), Scottish (*n*=1), South African (*n*=1)
[18] Non-Binary
[19] The bilingual speaker with the latest Age of Acquisition: 11

coordinated with a second clause with coordinating conjunctions *and* (n = 82) or *but* (n =78); lastly, the coordinated clause featured a personal pronoun (either *he* or *she*) in subject position. The nominal complement and personal pronoun were always separated by at least two words, and the personal pronoun was always followed by at least three words. These measures were taken to control for possible spill-over effects from the gendered noun, as well as to avoid possible wrap-up effects that might arise from measuring ERP responses from the last word of the sentence. The below item (Table 1, repeated from p. 24) is a representative example of a stimulus item in all four conditions, also demonstrating the above explained sentence structure:

Condition	**Experimental item**
REF-MATCH (baseline)	a. Henry met **a nurse** at the hospital and **she** laughed a lot.
REF-MISMATCH	b. Henry met **a surgeon** at the hospital and **she** laughed a lot.
PRED-MATCH	c. Henry was **a nurse** at the hospital and **she** laughed a lot.
PRED-MISMATCH	d. Henry was **a surgeon** at the hospital and **she** laughed a lot.

Table 1. An experimental item in all conditions.

Note that the NP-PRONOUN GENDER MATCH factor was manipulated with respect to whether there was semantic gender congruency between the referential/predicate NP and the subsequent personal pronoun, and *not* congruency between the proper noun subject of the first clause and the subsequent pronoun. The global gender mismatch between the proper noun subject of the first clause and the pronominal subject of the second clause was held constant in all conditions.

Nouns can be considered stereotypically gendered if they carry either a strong male or a strong female connotation according to a majority of speakers. Although the current study could not collect such information due to time constraints, the nouns used in the material were all taken from studies in which stereotypes were either researched or controlled for using normative data (Carreiras et al., 1996; Duffy & Keir, 2004; Kennison & Trofe, 2003; Kreiner et al., 2008; Oakhill et al., 2005; Osterhout et al., 1997; Wilbourn & Kee, 2010). 160 nouns, either stereotypically (n = 80, M =40, F =40) or definitionally (n = 80, M=40, F=40) gendered, were chosen to constitute the referential/predicate NPs in this study. These nouns were taken

from lists appearing in previous studies on stereotypically gendered (Carreiras et al., 1996; Duffy & Keir, 2004; Garnham et al., 2012; Kennison & Trofe, 2003; Kreiner et al. 2008; Oakhill et al., 2005; Osterhout et al., 1997; Reali et al., 2015; Wilbourn & Kee, 2010) or definitionally gendered (Osterhout et al., 1997) nouns.

For a small number of the definitionally gendered nouns that were taken from Osterhout et al. (1997), the corresponding male/female counterparts were added (for *congressman*: *congresswoman*; for *stepmother*: *stepfather*; for *baroness*: *baron*; for *nun*: *monk*; for *hostess*: *host*; for *lady*: *lord*; for *patriarch*: *matriarch;* for *actress*: *actor*). Nouns that are derived with suffixes *-man*[20]*, -woman* and *-ess* were considered definitionally gendered, as well as such nouns for which a dictionary definition contained explicit information about gender (Oxford University Press, 2023a; 2023b). On the basis of this, an additional eleven nouns were selected for this study (*cameraman, camerawoman, clergyman*, *clergywoman*, *count*, *countess*, *headmaster*, *headmistress, newsman*, *newswoman*, *seamstress*).

160 unique transitive verbs and complex verbal constructions[21] were used to construct readings in which the complement NP was rendered referential (*saw* in Table 1). It should be mentioned that the Implicit Causality (IC) of these verbs was not explicitly accounted for. Nonetheless, causal adverbial clauses (introduced by e.g. *because*) were avoided altogether and clauses were, as mentioned, joined via coordination processes only. Moreover, coordination has previously been shown to be able to reverse the IC of a verb (Erlich, 1980, as cited in Garnham & Oakhill, 1985, p. 388). Nonetheless, some limitations of not controlling for IC are discussed in section 5.3. The predicate readings of a complement NP were constructed with the help of 28 unique (semi-)copulas[22], which appeared a total of 4-7 times throughout the material (*was* in Table 1). Lists of the verbs and nouns used in the study, along with a list of all experimental items in all conditions, can be found in the Appendices.

As previously mentioned, the stimuli sentences were constructed as to give rise to a global gender mismatch between the proper name subject in the first clause (*Henry* in Table 1) and the pronominal subject in the second clause (*she* in Table 1), rendering the proper name subject unavailable as an antecedent. This global gender mismatch was held constant in all

[20] Although the suffix *-man* can also be used gender neutrally (Oxford University Press, 2023b, para III)
[21] As in verbal phrases consisting of more than one word, such as *had lunch with*, *yelled at*, etc.
[22] Including verbs taking prepositional complements (such as *as* and *like*) (Huddleston & Pullum, 2002, p. 255).

experimental items and in all conditions. The gender mismatch between the proper name and pronoun would thus trigger the participant to search through the prior discourse in search of an antecedent that matches the pronoun in terms of gender and number (Huang, 2000, pp. 8-9), in alignment with the theories on prominence lending cues and grammatical role parallelism (Hall & Yoshida, 2020, p. 301; Sauermann & Gagarina, 2017, p. 7; Smyth, 1994, p. 219; von Heusinger & Schumacher, 2019, p. 119) and the back-association hypothesis (Barkley et al., 2015, p. 145). The intervening NP, either referential or predicative, was therefore manipulated with regard to the NP-PRONOUN GENDER MATCH factor. Thus, the two NP types could be contrasted by comparing to what extent they give rise to agreement-attraction effects, reflecting whether they are considered potential antecedents to the anaphoric pronoun or not. All gendered nouns were distributed so that they appeared in all conditions (REF/PRED, MATCH/MISMATCH). The main critical word and time-locking point at which differences in ERP responses between conditions were predicted was the personal pronoun in the second clause.

Apart from the 160 test items, the participants saw an equal number of unrelated filler items, making the total number of sentences on each list 320. The order of appearance of the items was pseudorandomized. In an attempt to account for possible learning-effects, each list appeared in an additional version in which the order of appearance was reversed, resulting in a total of eight lists. The participants then saw one of the eight lists. Additionally, the participants answered 80 yes-or-no comprehension questions, randomly distributed throughout the session. The ratio of questions to items were 1:4, meaning that there were 40 questions for the test sentences and 40 for the filler sentences.

3.3 Procedure

While some participants were familiar with the EEG technique, most had never been part of an EEG experiment or seen the technique and its equipment up close. Therefore, especial consideration was taken to guide the participants through the procedure of the experiment; the purpose of each step in the preparation was made explicit to the participants and they were also made familiar with what the EEG data would look like.

The experimental stimuli were presented on a computer screen in a Rapid Serial Visual Presentation paradigm with the help of PsychoPy (v. 2021.2.3) (Peirce et al., 2019). The participants sat comfortably in a chair in front of the computer screen, at a distance of approximately 1 m. After an instruction slide, the participants practiced on three non-experimental sentences, two of which were followed by a comprehension question. After the practice session, the experimental items and comprehension questions were presented to the participants. Each trial began with a fixation cross (500 ms), after which the experimental sentence was presented one word at a time. Each word was visible for 300 ms and followed by an Inter-Stimulus Interval (ISI) of 200 ms. At the end of each trial, there was a 1000-ms pause followed by either a comprehension question or an indication to begin the next trial (in this case, an ellipsis). The participants were instructed to give an answer to the question using one of two keys on a keyboard, marked with green for 'yes' and red for 'no'. The questions were visible until the participant had provided an answer, after which an ellipsis was presented, indicating that the next trial could begin.

The comprehension questions mainly served as a means of keeping the participants' attention on the content of the stimulus. Keeping participants' attention in this way is a method common in neuro- and psycholinguistic research, by which the participant is kept unaware of the actual purpose of the study (Kaiser, 2013, p. 143). As the comprehension questions served a secondary purpose, no predictions about accuracy ratings were included as test implications to either hypothesis. Nonetheless, the behavioural results are presented in Chapter 4 and briefly discussed in Chapter 5. No participant was discarded because of low accuracy scores (range: 78% –100%, *M*=92.6%, *SD*=5.5%).

The experiment was divided into two blocks of 160 trials, but the participants were encouraged to take breaks in-between sentences, if needed. At the end of the experiment, the participants could see how many questions they had answered correctly. While the electrodes were removed, the participants were asked what their overall experience was, and the purpose of the study could also be disclosed. On average, participants completed the experiment in roughly an hour and spent 2.5-3 hours in the lab.

3.4 EEG data – Recording and Pre-processing

The EEG data was recorded from 30 electrodes mounted on *EASYCAP* recording caps (arranged in the 10-20 system; Luck, 2014, p. 167). EOG data (eye movements: saccades and blinks) was monitored using four electrodes (placed parallel to the lateral canthus of each eye, and below/above the left eye). The data was re-referenced offline to the average of two mastoid electrodes (Luck, 2014, p. 162-164). Impedances were kept as low as possible; this meant <10 kΩ for eye electrodes, <5kΩ for scalp electrodes and <2kΩ for the mastoid electrodes. A frontal, central electrode (AFz) was used as ground. The data was amplified using a *SynAmps*^{2} amplifier and recorded at a sampling rate of 1000 Hz using the *Curry 7* software provided by *Neuroscan*.

The EEG data was filtered offline using the *EEGlab* software (Delorme & Makeig, 2004), passing frequencies above 0.1 Hz and below 70 Hz, and then manually screened for EMG data (Gouvea et al., 2010; Luck, 2014). An Independent Component Analysis (ICA) was then carried out, in which EOG and channel-drift components were identified and rejected. The EEG data from the electrodes in the relevant Regions of Interest (ROIs) was divided into 1400-ms epochs (including a 200-ms pre-stimulus baseline) and then baseline-corrected. The extracted epochs were screened for remaining artefacts using the *ERPLAB Toolbox* (Lopez-Calderon & Luck, 2014) for *MATLAB* (v.9.12.0) (The MathWorks Inc., 2022), in which epochs with voltage amplitudes exceeding ±75µV were discarded. No participant had more than 25% of epochs removed per condition or in total (*M*=4.8%, *SD*=5.5%, maximum: 21%). The mean remaining number of trials per condition and participant was: REF-MATCH: *M*=38.12, *SD*=2.29; REF-MISMATCH: *M*=37.96, *SD*=2.54; PRED-MATCH: *M*=38.6, *SD*=2.59; PRED-MISMATCH: *M*=38.04, *SD*=2.25.

3.5. Statistical analysis

Following Delogu et al. (2019) and Gouvea et al. (2010), the ROIs (based on laterality and anteriority/posteriority) and relevant ERP time windows were predefined as the basis for analysis. The two ROIs were: a left anterior region (electrodes: Fp1, F7, F3, FT7, FC3) and a posterior region (electrodes: TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, O1, Oz, O2). The

left anterior region is relevant for the LAN component, peaking somewhere 300-500ms post critical word onset, and the posterior region is relevant for the P600 component. The P600 component usually peaks around 600 ms after the critical word, but there is variation in how the latency window is defined in the literature. Building on previous studies, the P600 time window for this study was therefore set to 500-1000 ms post stimulus onset (Barkley et al., 2015; Delogu et al, 2019; Gouvea et al., 2010).

For statistical analysis, a Linear Mixed-Effects Model (LMM) analysis was performed in *R* (v.4.2.3) in the R Studio interface, using the *lme4* (v. 1.1-30) and *lmerTest* (v. 3.1-3) packages (Bates et al., 2015; Kuznetsova et al., 2017; R Core Team, 2023). Data from the ROIs and time windows outlined above was treated as different datasets on which the analyses were run. Mixed-effect models aim to account for how pre-defined independent variables are predicted to interact with the dependent variable(s), varying by effects that are non-independent (Winter, 2020, pp. 234-236). In the case of this study, the two independent variables (also: fixed effects) were the two categorical factors described in section 3.2 above (i.e. NP TYPE (REF, PRED) and NP-PRONOUN GENDER MATCH (MATCH, MISMATCH)). The continuous dependent variable was Amplitude (μV) of the LAN and P600 in the 300-500-ms and 500-1000-ms time windows, respectively. The random effects were Participants and Items. Whereas other repeated measures analyses assume a constant effect between independent and dependent variables, a mixed-effects model is not biased in assuming a constant effect. A mixed-effects model can thus provide an account for any possible individual variance between participants and items (Winter, 2020, pp. 234-235).

Following the definitions of the fixed and random effects above, the below example shows the maximal (in terms of complexity) LM model that could be entered into R, in *lme4* syntax:

(15) `lmer(Amplitude ~ NP-Pronoun Gender Match * NP Type + (NP-Pronoun Gender Match + NP Type|Participant) + (NP-Pronoun Gender Match + NP Type|Item)`

The above model thus fits the possible interaction between the independent variables on the dependent variable, with by-participant and by-item varying intercepts and by-participant and by-item varying slopes (Winter, 2020, p. 241; Baayen et al., 2008). However, in situations where model outputs indicated overparameterization, the model was simplified by removing random effect correlations (see Baayen et al., 2008, p. 395). The justification of any such model

simplification was controlled with a Likelihood ratio test using the *anova* function, which allowed for comparisons between LMM models of different complexities and with differing numbers of parameters. If the model comparison indicated that the simplification was justified (in that both fitted models explain the data equally well), the simpler model was chosen. The EEG results presented in Chapter 4 are therefore based on the maximally converging, non-overparameterized LMM models for which any simplification could be justified by a Likelihood ratio test.

Following Delogu et al. (2019), the P600 onset latencies were investigated by examining the mean amplitude of the P600 in smaller, 200-ms windows, varying by a 100-ms onset (i.e. 500-700 ms, 600-800 ms, 700-900 ms, 800-1000 ms). While this does not provide an exact indication of when the P600 onsets, it can say something about when the effect first became significant. The mean amplitude between conditions in the smaller time windows was thus compared, indicating when the effect of either factor first became visible, and when it was most pronounced.

Pairwise comparisons between different conditions were performed in both the main LAN/P600 amplitude and the P600 effect significance onset ('latency') analyses, using follow-up Estimated Marginal Means (EMM) analyses. The analyses were run on the maximally converging LMMs, with the help of the *emmeans* package (v. 1.8.2; Lenth, 2023). An EMM analysis builds on the comparisons calculated in the LMM and provides straight-forward overviews on how the dependent variable contrasts between conditions. To resolve the false-positive problem and account for Type I errors, the *p*-values in the EMM are adjusted for multiple comparison (the Tukey method, four estimates). (16) provides an example of how the EMM models were entered into R, in *emmeans* syntax:

(16) `emmeans(LMmodel, pairwise ~ NP-Pronoun Gender Match + NP Type)`

With respect to the behavioural data, no hypotheses were formulated regarding the accuracy of responses. Nonetheless, a post-hoc GLMM (Generalized Linear Mixed Model) analysis was run on the behavioural data (using the *glmer* function) to investigate whether there were any significant effects of, and interactions between, the independent variables on comprehension question accuracy rates. Response accuracy (a categorical, binary, variable) was treated as the dependent variable, the independent variables were treated as fixed effects,

and items and participants as random effects. The maximally converging model entered into R using *lme4* syntax was:

(17)
```
glmer(Response Accuracy ~ NP-Pronoun Gender Match * NP Type
+ (NP-Pronoun Gender Match + NP Type |participant)+ (1 |Item),
family =binomial)
```

Response times, on the other hand, were not considered, as participants had not been explicitly instructed to respond to the questions as fast as possible.

4. Results

4.1 ERP data

4.1.1 ERP responses, 300-500 ms (LAN)

A visual examination of the EEG grand averages of the left-frontal scalp electrodes indicated no LAN effect for the predicate constructions overall or the mismatch conditions specifically, but rather a sustained negativity first visible around 700 ms post stimulus onset (discussed in detail in subsection 4.1.3):

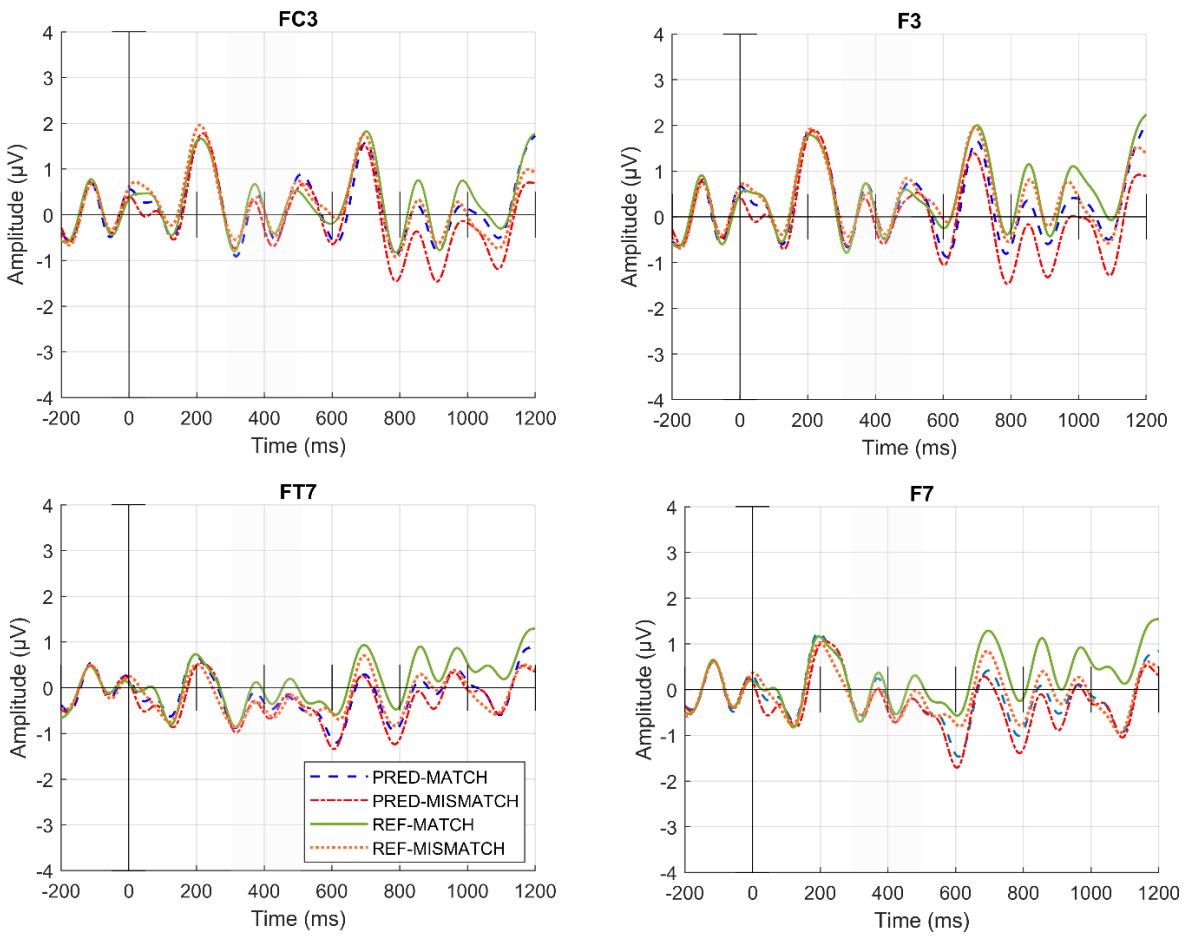

Figure 5. Grand average ERPs of select electrodes in left anterior region at pronoun onset, all conditions. Grey marks the LAN latency window.

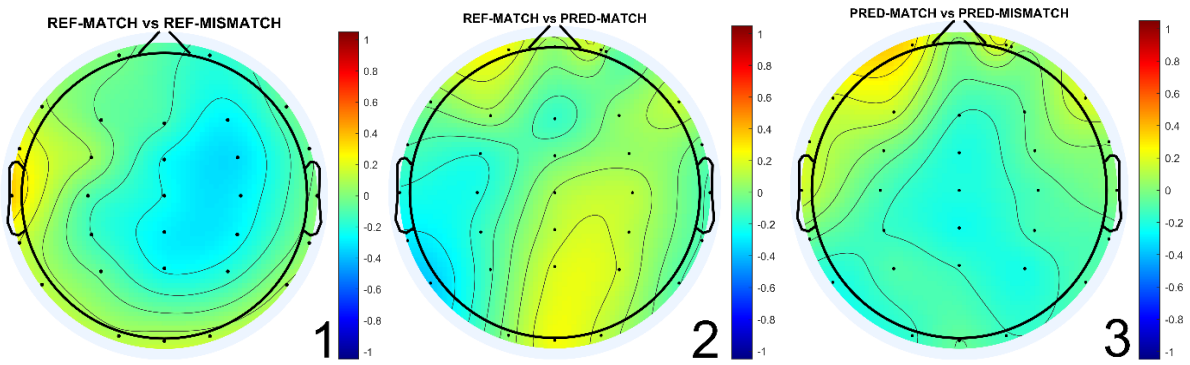

Figure 6. Topographic plots, 300-400 ms post-stimulus onset, showing the spatially distributed (amplitude) difference between REF-MATCH and REF-MISMATCH (1), REF-MATCH and PRED-MATCH (2) and PRED-MATCH and PRED-MISMATCH (3).

The LMM analysis confirmed this indication; no effect of NP-PRONOUN GENDER MATCH or NP TYPE was found to affect processing in the predefined LAN time window and ROI. This was shown by the insignificant differences and interactions between factors and their levels against both a REF-MATCH and PRED-MATCH intercept:

Fixed effects	**Estimate**	**SE**	**Df**	**t value**	**Pr(>|t|)**	**Sign.**
REF, MATCH (intercept)	2.360e-01	1.626e-01	1.258e+02	1.451	0.149	
MISMATCH	-7.094e-02	1.970e-01	3.656e+03	-0.360	0.719	
PRED	-5.826e-02	1.971e-01	3.663e+03	-0.296	0.768	
MISMATCH:PRED (interaction)	4.593e-03	2.787e-01	3.659e+03	0.016	0.987	
PRED, MATCH (intercept)	1.778e-01	1.625e-01	1.254e+02	1.094	0.276	
MISMATCH	-6.635e-02	1.972e-01	3.656e+03	-0.336	0.737	
REF	5.826e-02	1.971e-01	3.663e+03	0.296	0.768	
MISMATCH:REF (interaction)	-4.593e-03	2.787e-01	3.659e+03	-0.016	0.987	

Table 2. LMM results for Left Anterior region and time window; Model: `Mean Amplitude ~ NP-Pronoun Gender Match * NP-Type + (1 | Participant) + (1 | Item)`. The interaction of NP-PRONOUN GENDER MATCH*NP-TYPE and its effect on EEG amplitude.

As the LMM analysis indicated no significant effects of either factor on the mean amplitude in the LAN time window or ROI, no follow-up EMM pairwise comparisons were conducted.

4.1.2 ERP Responses, 500-1000 ms (P600)

4.1.2.1 P600 Mean Amplitude

A visual inspection of the grand averages in the posterior region across conditions suggested a pronounced positive ERP response to the PRED-MISMATCH condition compared to its PRED-MATCH counterpart, the baseline (REF-MATCH) and control condition (REF-MISMATCH). The positivity appeared similar to a P600-like response, peaking between 500-1000 ms post stimulus onset with a clear posterior distribution:

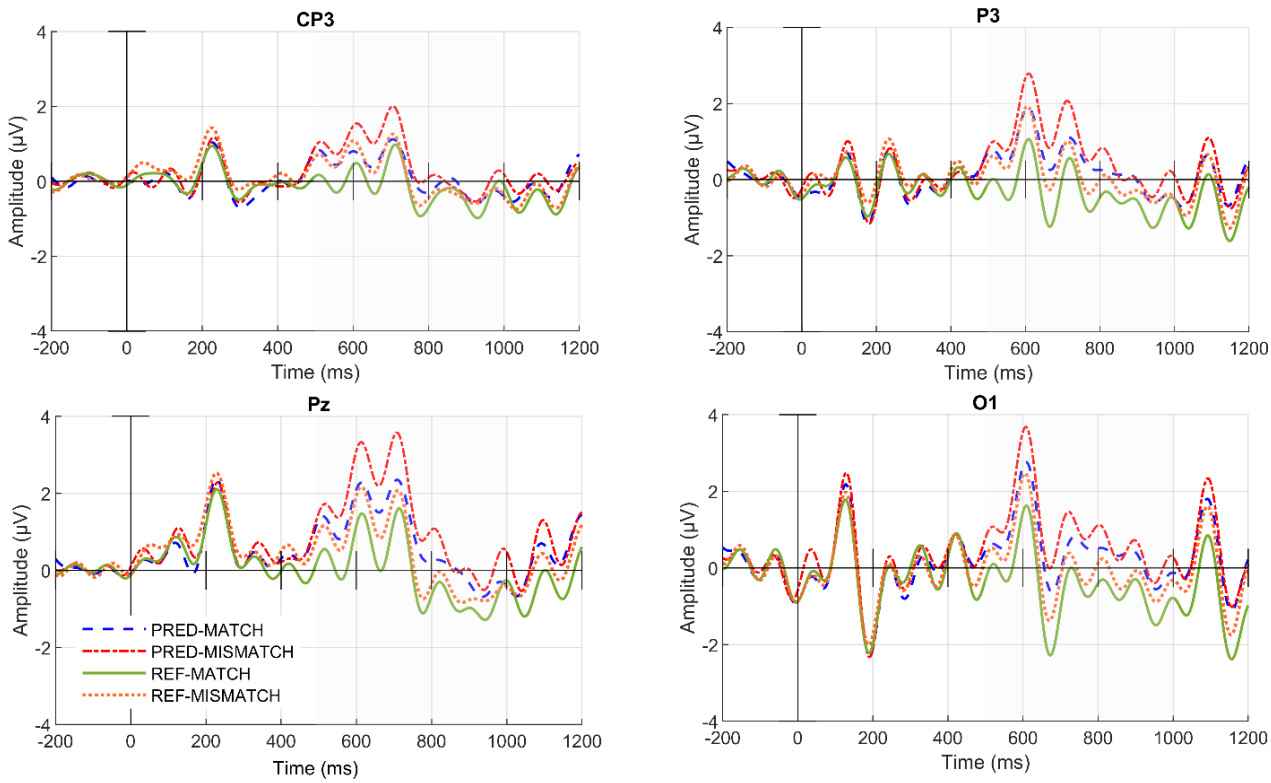

Figure 7. Grand Average ERPs on select electrodes in the posterior region at pronoun onset, all conditions. Grey marks the P600 latency window.

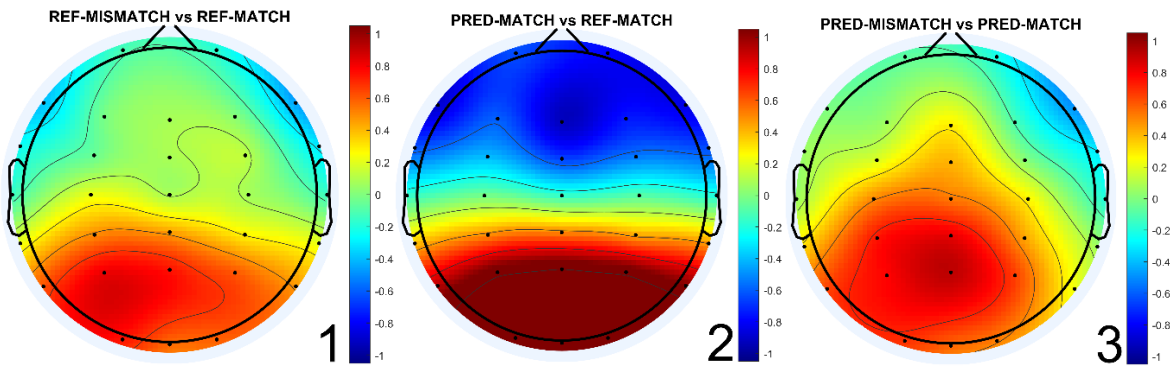

Figure 8. Topographic plots in the 600-700-ms time window, showing the spatially distributed amplitude difference between REF-MISMATCH and REF-MATCH (1), PRED-MATCH and REF-MATCH(2), PRED-MISMATCH and PRED-MATCH (3).

An LMM analysis of the mean amplitudes in the predefined P600 time window (500-1000 ms) and ROI was then conducted. The LMM analysis calculates the effect on EEG amplitude when one level in the intercept reference condition (in this case, the baseline REF-MATCH condition) is altered for an alternate level (REF-MISMATCH and PRED-MATCH). The cumulative effect of altering both intercept factor levels (REF, MATCH) for their alternate levels (PRED, MISMATCH) is thought of as the interaction between factors. That is, the interaction indicates whether the outcome of altering the NP-PRONOUN GENDER MATCH factor is dependent on the level of the NP-TYPE-factor. As H2 predicted an interaction between the factors, a first LMM model including an interaction parameter was run (Model: `Mean Amplitude ~ NP-Pronoun Gender Match * NP-Type + (1 | participant) + (1 | item)`). The model indicated no significant interaction effect (p=0.8997). The plot in Figure 9 illustrates how the NP-pronoun gender mismatch renders an increase in amplitude, regardless of NP-TYPE:

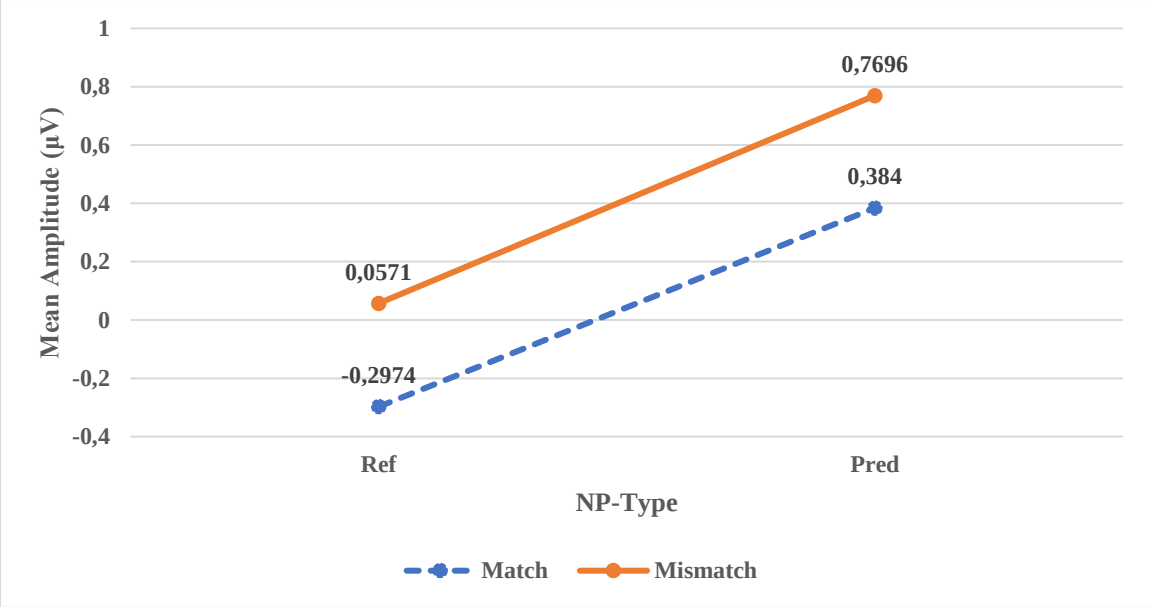

Figure 9. The effects of NP-TYPE and NP-PRONOUN GENDER MATCH on mean amplitude (µV) in the P600-time window.

As the *t*-value and *p*-value to the interaction of factors in the first LMM model suggested that the interaction effect was redundant ($t<\pm2$, $p=0.8997$), the model was simplified by removing the interaction parameter altogether. The first and second models were compared using a Likelihood ratio test, which indicated that the simpler model was justified ($p>0.05$, indicating no significant difference between models):

Models	**npar**	**AIC**	**BIC**	**logLik**	**deviance**	**Chisq**	**Df**	**Pr(>Chisq)**
`Mean Amplitude ~ NP-Pronoun Gender Match + NP-Type+ (1 | Participant) + (1 | Item)`	6	21079	21116	-10533	21067			
`Mean Amplitude ~ NP-Pronoun Gender Match * NP-Type+ (1 | Participant) + (1 | Item)`	7	21081	21125	-10533	21067	0.0159	1	0.8997

Table 3. Likelihood ratio test between LMM models on P600 data, with and without factor interaction parameter.

Model fit and the assumption of normality of the second model were then evaluated and confirmed visually (Winter, 2020, p. 110). The LMM assumes that the residuals are normally distributed. By plotting the residuals (i.e. the errors, or the discrepancies between the observed

and predicted data) and comparing how these hold up against models of normal distribution, model fit can be assessed:

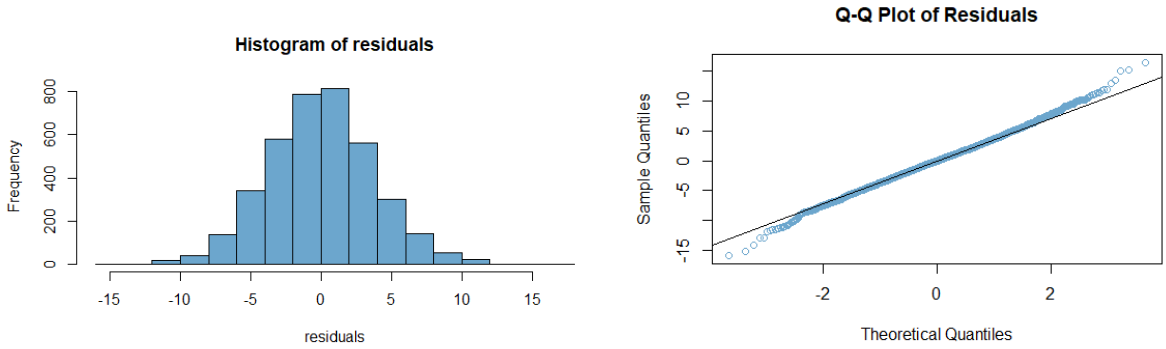

Figure 10. The normal distribution of residuals of the fitted P600 LM model, as visualised in a histogram (left) and a quantile-quantile (QQ) plot (right).

The plots in Figure 10 indicate that the model estimates maximum likelihood; the residuals follow a typical bell-shaped curve (Fig. 10, left) and mostly align in a straight diagonal line (Fig. 10, right). The results of the simplified LMM can thus be taken as a good explanation of the data, accurately depicting the significance of the observed effects.

The results of the LMM for the P600 ROI and latency window revealed a significant increase in positivity via manipulation for the NP-TYPE level against both intercepts ($p<.001$). Moreover, it indicated a significant effect of manipulation of the NP-PRONOUN GENDER MATCH factor, in both the REF and PRED conditions ($p=0.00273$). An EMM analysis was then run on the model, which provides a good overview of the significant contrasts between conditions:

Contrast	Estimate	SE	Df	t ratio	p-value	Sign.
PRED, MATCH – PRED, MISMATCH	-0.370	0.123	3655	-2.998	0.0145	*
PRED, MATCH – REF, MATCH	0.697	0.123	3663	5.646	<.0001	***
PRED, MATCH – REF, MISMATCH	0.327	0.174	3656	1.875	0.2392	
PRED, MISMATCH – REF, MATCH	1.067	0.175	3661	6.106	<.0001	***
PRED, MISMATCH – REF, MISMATCH	0.697	0.123	3663	5.646	<.0001	***
REF, MATCH – REF, MISMATCH	-0.370	0.123	3655	-2.998	0.0145	*

Table 4. Results of EMM analysis on the P600 LMM; pairwise comparison on the effect of NP-TYPE and NP-PRONOUN GENDER MATCH. p-values have been adjusted for multiple comparisons (Tukey method, 4 estimates). Significance codes: '***' $p<0.001$ '**' $p<0.01$ '*' $p<0.05$ '.' $p<0.1$

The EMM calculates the differences in mean amplitudes between all conditions and indicates whether the differences are significant or not. The results in Table 4 show a significant contrast between conditions PRED-MATCH and REF-MATCH, in terms of increased positivity for the former (as indicated by the positive Estimate value; $p<.0001$). Moreover, a similar contrast was found between conditions PRED-MISMATCH and REF-MISMATCH ($p<.0001$), as well as for the maximally different conditions PRED-MISMATCH and REF-MATCH ($p<.0001$). There was, however, no significant contrast between the REF-MISMATCH and PRED-MATCH conditions ($p=0.2392$). Furthermore, the amplitude contrast between REF-MATCH and REF-MISMATCH, as well as between PRED-MATCH and PRED-MISMATCH reached significance ($p=0.0145$). In sum, both PRED conditions generated significant P600 effects against the baseline REF-MATCH condition, whereas the PRED-MATCH response was not significantly more positive compared to the REF-MISMATCH condition. There was also a significant effect of NP-PRONOUN GENDER MATCH on the P600 mean amplitude in both the PRED and REF condition. The implications of these results are discussed in further detail in Chapter 5.

4.1.2.2 P600 Effect Significance Onset

As outlined in Chapter 3, P600 effect latency differences were investigated by dividing the P600 time window into smaller 200-ms intervals, in order to see where the P600 effect first appeared significant across and between conditions. As the general P600 mean amplitude

analysis presented in 4.1.2.1 indicated no effect of interaction between factors, the interaction parameter was excluded here as well.

LMM analyses were run on the individual datasets (500-700 ms, 600-800 ms, 700-900 ms and 800-1000 ms), as well as on a post-hoc dataset (400-600 ms). Follow-up EMM analyses were then conducted on the maximally converging models without overparameterizations (as indicated by model output inspection and Likelihood ratio tests). The mean amplitude contrasts between conditions in the latency intervals can be summarised as in Table 5:

	400-600 ms		500-700 ms		600-800 ms		700-900 ms		800-1000 ms	
Condition	**Contrast (µV)**	***p*-value**	**Contrast (µV)**	***p*-value**	**Contrast (µV)**	***p*-value**	**Contrast (µV)**	***p*-value**	**Contrast (µV)**	***p*-value**
REF, MATCH – PRED, MATCH	-0.408	0.0556 .	-0.793	0.0001**	**-0.898**	<.0001***	-0.724	<.0001***	-0.532	0.0076**
REF, MATCH – PRED, MISMATCH	-0.689	0.0391*	-1.283	0.0001**	**1.378**	<.0001***	-1.029	<.0001***	-0.767	0.0047**
REF, MATCH – REF, MISMATCH	-0.281	0.2746	**-0.490**	0.0310 *	-0.479	0.0265*	-0.305	0.1571	-0.234	0.4434
REF, MISMATCH – PRED, MATCH	-0.127	0.8987	-0.302	0.4293	-0.419	0.3839	-0.419	0.1470	-0.298	0.5601
REF, MISMATCH – PRED, MISMATCH	-0.408	0.0556 .	-0.793	0.0001**	**-0.898**	<.0001***	-0.724	<.0001***	-0.532	0.0076**
PRED, MATCH – PRED, MISMATCH	-0.281	0.2746	**-0.490**	0.0310*	-0.479	0.0265*	-0.305	0.1571	-0.234	0.4434

Table 5. Estimated contrast (µV) between experimental conditions at different latency intervals at posterior ROI. Estimate and probability values are from EMM analysis outputs on individual the LMM models. Bold indicates maximal significant mean amplitude contrast. Significance codes: '***'$p<0.001$ '**'$p<0.01$ '*'$p<0.05$ '.'$p<0.1$ ''$p<1$. (with inspiration from Gouvea et al., 2010, p. 164) [23]

[23] Fitted LMM models upon which EMM analyses were run:
400-600 ms: Mean Amplitude ~ NP-Pronoun Gender Match + NP-TYPE + (NP-Pronoun Gender Match + NP-TYPE | participant) + (1 | item)
500-700 ms: Mean Amplitude ~ NP-Pronoun Gender Match + NP-TYPE + (NP-Pronoun Gender Match + NP-TYPE |participant) + (NP-Pronoun Gender Match + NP-TYPE |item)
600-800 ms: Mean Amplitude ~ NP-Pronoun Gender Match + NP-TYPE + (NP-Pronoun Gender Match + NP-TYPE |participant) + (1|item)
700-900 ms: Mean Amplitude ~ NP-Pronoun Gender Match + NP-TYPE + (1+ NP-Pronoun Gender Match |participant) + (1|item)
800-1000 ms: Mean Amplitude ~ NP-Pronoun Gender Match + NP-TYPE + (NP-Pronoun Gender Match + NP-TYPE |participant) + (1|item)

In sum, the contrast between REF-MATCH (baseline) and PRED-MISMATCH was significant already in the 400-600 ms window, was most pronounced in the 600-800 ms time window, and stayed significant throughout the entire P600 latency window. The contrast between REF-MISMATCH and PRED-MATCH never reached significance, but the contrasts between MATCH and MISMATCH, for both the PRED and REF conditions, were significant between 500-800 ms, with a peak amplitude contrast in the 500-700 ms time window. The contrast between REF-MATCH and PRED-MATCH appeared significant in the 500-700 latency window and was most pronounced in the 600-800 ms time window. The effect stayed significant until 1000 ms post stimulus onset. The P600 effect for PRED-MATCH relative to the baseline REF-MATCH condition appeared significant in the same latency window as the P600 effect for the REF-MISMATCH condition against the REF-MATCH baseline.

4.1.3 Post-hoc: ERP Responses, 700-1100 ms (Late Anterior Negativity)

As mentioned in 4.1.1, the visual inspection of the EEG grand averages for all electrodes indicated a sustained late negative deflection, emerging at anterior electrode sites somewhere after 700 ms post stimulus onset. Its latency resembled that of the Late (L)AN component (hereafter: L-(L)AN) (Steinhauer, 2010). However, the negativity was observed not only at left-lateralized electrode sites, but appeared bilateral or slightly right-lateralized, at least for the PRED-MATCH condition relative to REF-MATCH:

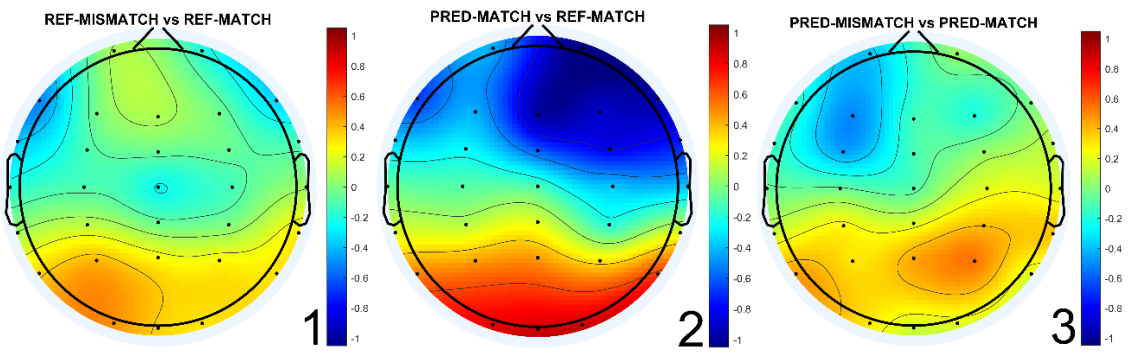

Figure 11. Topographic plots, 900-1000 ms post-stimulus onset, showing the spatially distributed amplitude difference between REF-MATCH and REF-MISMATCH (1), REF-MATCH and PRED-MATCH (2) and PRED-MATCH and PRED-MISMATCH (3).

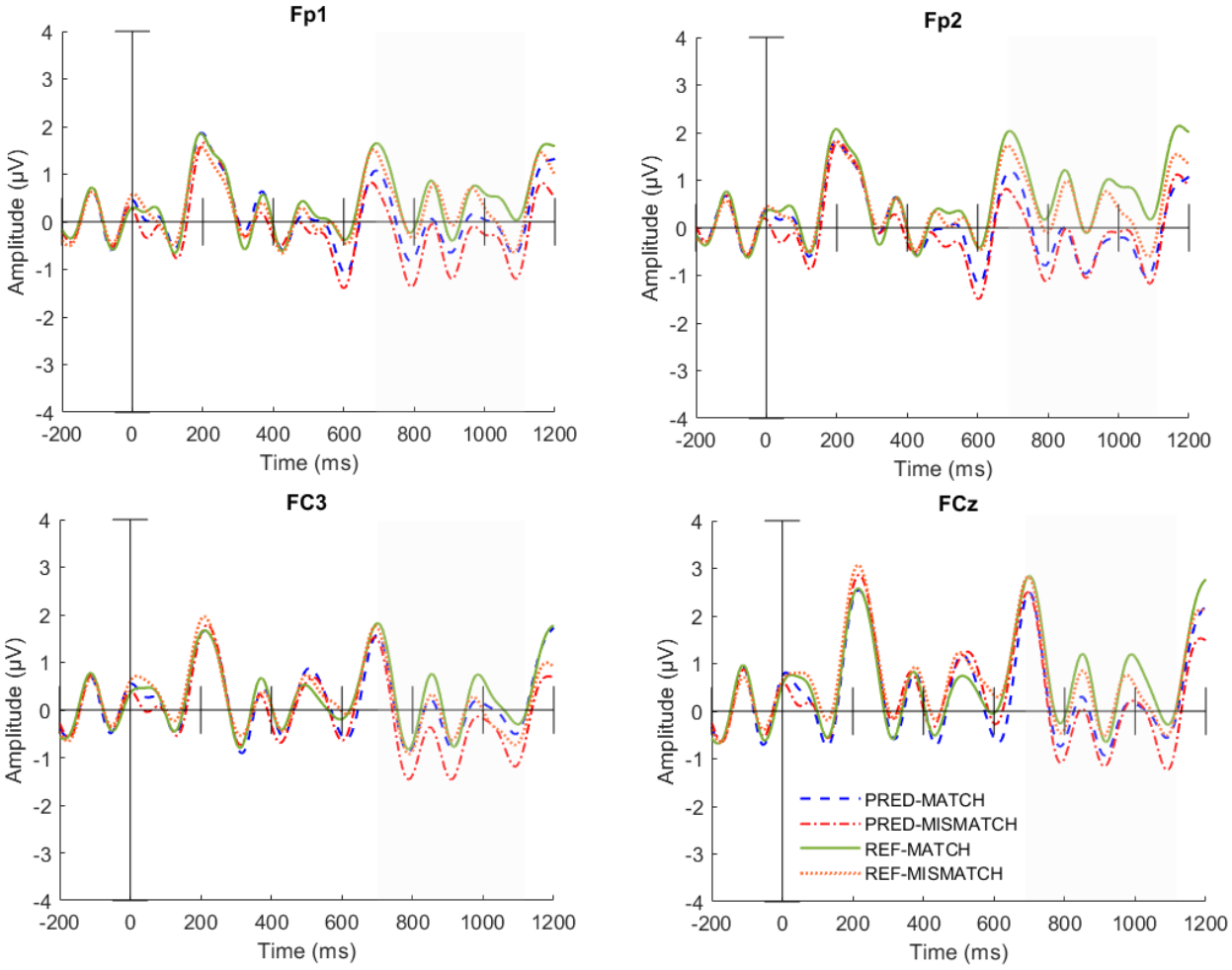

Figure 12. Grand average ERPs of select electrodes in the Anterior region at pronoun onset, all conditions. Grey marks the L-AN latency window (700-1100 ms)

Based on this observation, post-hoc LMM and EMM analyses were conducted in the 700-1100 ms time window, on a broadly anterior region (electrodes included in the ROI: FP1, FP2, Fz, FCz, FC3, FC4, F3, F4, F7, F8, FT7, FT8). A first LMM model was first run on the dataset with the following LMM syntax, including an interaction argument:

(18) `Mean Amplitude ~ NP-Pronoun Gender Match * NP-Type + (1 | Participant) + (NP-Pronoun Gender Match + NP-Type | Item)`

As for the P600 dataset, the model output indicated no significant interaction between factors (p=0.7792). It was thus simplified by removing the interaction argument, and the simplification was evaluated with a Likelihood ratio test:

Models	n par	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
`Mean Amplitude ~ NP-Pronoun Gender Match + NP-Type + (1 | Participant) + (Np-Pronoun Gender Match + Np-Type | Item)`	11	22148	21117	-11063	22126			
`Mean Amplitude ~ NP-Pronoun Gender Match * NP-Type + (1 | Participant) + (Np-Pronoun Gender Match + Np-Type | Item)`	12	22150	22225	-11063	22126	0.0707	1	0.7903

Table 6. Likelihood ratio test results of LMM models on L-AN data, with and without factor interaction parameter.

The simpler model without the interaction parameter was justified (p=0.7903), indicating no significant difference between models. Assumption of normality and goodness-of-fit was then evaluated visually using a histogram and Q-Q-plot:

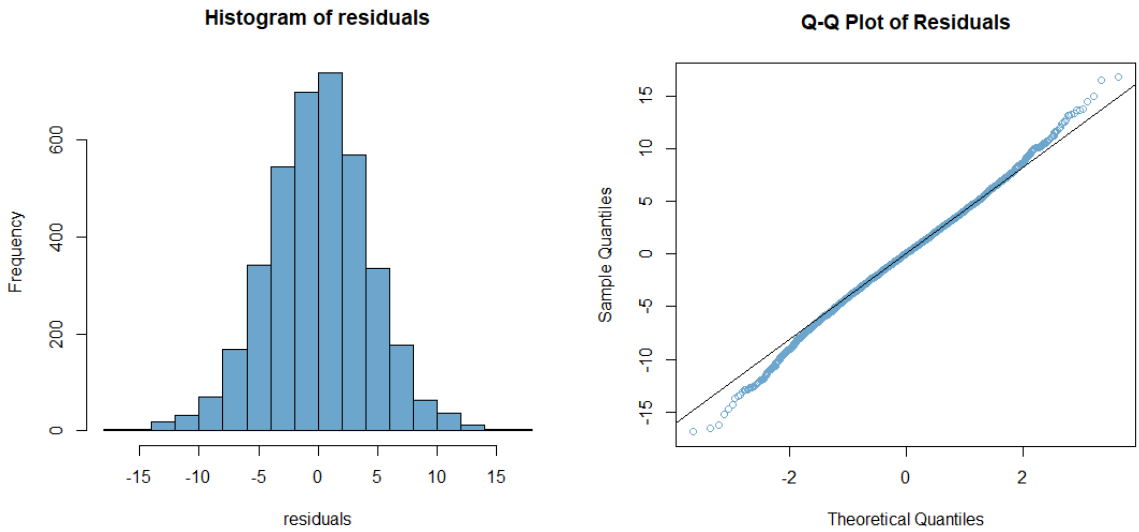

Figure 13. The normal distribution of residuals of the fitted L-AN LM model, as visualised in a histogram (left) and a quantile-quantile (QQ) plot (right).

As deemed by visual inspection, the residuals mostly align in accordance with standard normal distribution. Thus, the residuals/errors in the L-AN dataset can be considered normally

distributed and the LMM and subsequent EMM model can be trusted as good indicators of accurate representations of the results.

The LMM results could confirm a significant effect of NP-TYPE alternation against a REF-MATCH intercept, meaning that (at least) PRED-MATCH yielded a significant increase in negativity against the baseline condition (p<3.98e-06). The fitted model indicated no significant effect of gender MISMATCH relative to the baseline REF-MATCH intercept on the L-AN negativity amplitude (p=0.097382). Releveling the model did not reveal any significant effect of gender congruency against a PRED-MATCH intercept, either (p=0.0687). The EMM analysis that was run on the fitted model illustrates further which contrasts reached significance:

contrast	**Estimate**	**SE**	**df**	**t ratio**	**p-value**	**sign.**
PRED, MATCH – REF, MATCH	-0.674	0.145	158	-4.658	<.0001	***
PRED, MATCH – PRED, MISMATCH	0.258	0.155	160	1.662	0.3475	
PRED, MATCH – REF, MISMATCH	-0.416	0.218	159	-1.912	0.2270	
REF, MATCH – PRED, MISMATCH	0.932	0.207	158	4.503	0.0001	**
REF, MATCH – REF, MISMATCH	0.258	0.155	160	1.662	0.3475	
PRED, MISMATCH – REF, MISMATCH	-0.674	0.145	158	-4.658	<.0001	***

Table 7. Results of EMM analysis on LMM for the L-AN latency window and ROI; pairwise comparison on the effect of NP-TYPE and NP-PRONOUN GENDER MATCH on mean amplitude. p-values have been adjusted for multiple comparisons (Tukey method, 4 estimates). Significance codes: '***' p<0.001 '**' $p<0.01$ '*' $p<0.05$ '.' $p<0.1$ ' ' $p<1$.

The EMM model summarised in Table 7 indicates significant contrasts between the REF-MATCH (baseline) and PRED-MATCH conditions (p<.0001), the REF-MISMATCH and PRED-MISMATCH conditions (p<.0001) and the REF-MATCH and PRED-MISMATCH conditions (p=0.0001). However, it shows no significant effects of NP-PRONOUN GENDER on the Late AN amplitude within the NP-TYPE levels – that is, for neither REF (p=0.3475) nor PRED (p=0.3475). The model does not indicate any significant amplitude difference between the REF-MISMATCH and PRED-MATCH conditions, either (p=0.2270). In sum, these results indicate a main effect of NP-TYPE (PRED) on the L-(L)AN amplitude 700-1100 ms post pronoun onset

in the anterior ROI. These results, in light of an overview of the L-(L)AN component, are discussed in section 5.1.

4.2 Behavioural data

Each participant answered a total of 40 comprehension questions to the experimental stimuli. Table 8 shows a summary of response accuracy by condition:

	NP-PRONOUN GENDER MATCH	
NP-TYPE	*Match*	*Mismatch*
Referential	91.5 (*SD*=28.0)	92.8 (*SD*=25.9)
Predicate	96.1 (*SD*=19.5)	90.0 (*SD*=30.0)

Table 8. Mean response accuracy rate (%) per condition

As is clear from the descriptive results in Table 8, response accuracy rates were overall high, but the condition with the highest accuracy rate was PRED-MATCH, with a mean accuracy rate of 96.1% (0.961; *SD*=0.195). Although accuracy rates were seemingly higher for MISMATCH than MATCH in the REF condition, this pattern was not significant:

Fixed effects	**Estimate**	**SE**	**z value**	**Pr(>|z|)**	**Sign.**
REF, MATCH (intercept)	2.9520	0.4058	7.274	3.49e-13	***
MISMATCH	-0.1425	0.4492	-0.317	0.7511	
PRED	0.8952	0.4971	1.801	0.0717	.
MISMATCH:PRED (interaction)	-1.2503	0.5878	-2.127	0.0334	*

Table 9. Results of GLMM analysis on comprehension question; the interaction of MATCH*NP-TYPE and its effect on Response Accuracy[24]. Estimate values correspond to log-odds estimates. Significance codes: '***'$p<0.001$ '**'$p<0.01$ '*'$p<0.05$ '.'$p<0.1$ ''$p<1$.

[24] Model: `Accuracy Score ~ NP-Pronoun Gender Match * NP-Type + (NP-Pronoun Gender Match + NP-Type | Participant) + (1 | Item)`

The GLMM results indicate that there was no significant increase in response accuracy between the REF and PRED levels (p=0.0717), but that there was a significant interaction between MISMATCH and PRED, i.e. that the accuracy rate for the MISMATCH level was dependent on the PRED level, as illustrated in Figure 14:

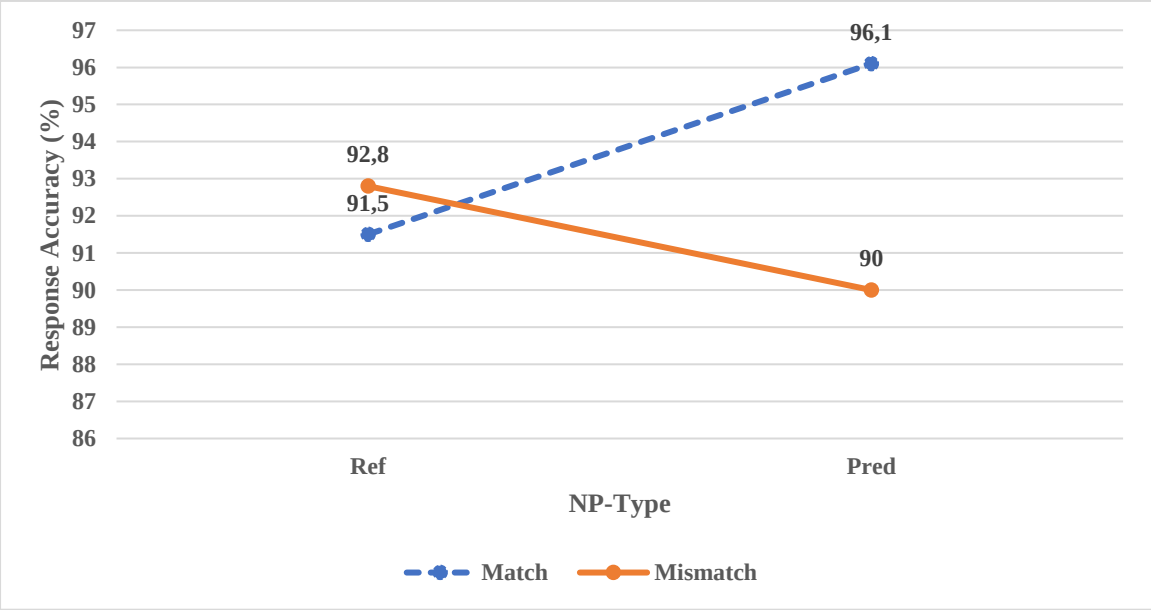

Figure 14. The effect of NP-TYPE and NP-PRONOUN GENDER MATCH on Response Accuracy (%).

The interaction effect illustrated in the plot suggests that the MISMATCH level yielded a lower response accuracy than MATCH in the PRED condition. Nonetheless, the negative correlation between the effect of MISMATCH and accuracy rates in the PRED condition was merely by 6 percentage points, indicating that participants on average answered ~2.4 more questions incorrect in the PRED-MISMATCH condition than in the PRED-MATCH condition.

5. Discussion

This chapter discusses the results presented in Chapter 4. The first section revisits the hypotheses the study aimed to test and discusses the electrophysiological results. The second section briefly discusses the behavioural results. Some limitations of the study are discussed in the third section.

5.1 Electrophysiological Results

This study set out to test two competing hypotheses on how predicate nominals are stored and accessed during co-referential resolution:

> *H1*. Predicate NPs are accessible as potential antecedents for anaphora in the mental model, due to the morphosyntactic form that they share with referential NPs.
>
> *H2*. Predicate NPs are not accessible as potential antecedents for anaphora in the mental model, due to their predicative function.

For the possibility of predicate NPs being considered potential antecedents during online anaphor resolution (H1), it was hypothesized that a gender congruency factor between the complement NP and a subsequent anaphor would modulate ERP amplitudes for referential NPs and predicate NPs alike. For predicate NPs *not* being regarded as possible antecedents for anaphor resolution (H2), it was hypothesized that gender congruency would only have an effect on ERP amplitudes for the referential NPs, and not for the predicate NPs. Instead, only a general increase in P600 and LAN amplitudes between the predicate-NP conditions and the referential gender congruent condition was expected. The hypotheses were thus formulated with specific test implications that could be tested in an EEG paradigm.

There are a few interesting aspects to discuss in light of the results presented in Chapter 4. First of all, no experimental condition elicited any Left Anterior Negativity in the predefined component latency window, i.e. 300-500 ms post stimulus onset. The LAN effect was absent within both the PRED and REF condition levels (MATCH vs. MISMATCH), but also between the NP-TYPE condition levels (REF vs. PRED). The lack of a LAN response is perhaps not entirely

surprising, given the inconsistency of (early) LAN responses reported in previous EEG research on morphosyntactic agreement violations (e.g., Allen et al., 2003 and Osterhout et al., 1996, as cited in Tanner & Van Hell, 2014, p. 290). Along the reasoning of some scholars, the absence of a LAN effect could potentially be accredited to the lack of explicit, morphological markers of agreement (in this case gender) between the agreeing elements (see e.g. Molinaro et al., 2011, p. 925-926). Indeed, most role nouns in this study did not contain any morphological marker for gender. Instead, the gender of the nouns was determined by their stereotypical or definitional gender, the latter usually only encoded lexicosemantically. Molinaro et al. (2011) claim that agreement errors dependent on 'opaque' gender markings are not likely to elicit a LAN, in general; and, if elicited, it is likely not left-lateralized, but likely broadly anterior (Molinaro et al., 2011, p. 925). In light of e.g. the back-association hypothesis (Barkley et al., 2015) and the fact that gender incongruency in fact has been found to generate early LAN effects in pronoun resolution paradigms before (e.g. King & Kutas, 1998), there were good reasons for predicting the presence of a LAN effect in the current study. However, the results are ultimately in line with those of many other studies reporting no biphasic LAN-P600 effects.

While no LAN effect was found in the predefined latency window, a post-hoc analysis indicated a notable, late, broadly anterior, sustained negativity, significant between 700-1100 ms post stimulus onset. In other words, while the expected, LAN/P600 pattern was not found, a P600/L-AN pattern was found instead. While such sustained negativities can be difficult to disassociate from for example the Nref component, the Nref is typically associated with referential ambiguity (Van Berkum, 2009, p. 286, 301). It is not commonly elicited in unambiguous referential contexts - not for unambiguous pronouns in licit licensing contexts, nor in situations when context effectively can cancel referential ambiguity between two licit antecedents, or when there is no appropriate antecedent at all (Van Berkum et al., 2007, p. 162). Nonetheless, Nieuwland (2014) did report an Nref response to mismatching pronouns with only one possible antecedent. However, in two out of three experimental settings, the Nref was not observed alongside any P600, the electrophysiological response most commonly reported for reference resolution failure (p. 17; Van Berkum et al., 2007). The Nref was thus interpreted as the participants inferring a referent outside the current discourse altogether, rather than trying to build co-reference with the existing, gender-mismatching antecedent (Nieuwland, 2014, pp.17-18). Contrastively, the negativity observed in the current study was elicited by illicit (predicate) antecedents, and, importantly, alongside a significant P600 effect. Therefore, it is

not obvious that it should be considered an Nref effect; instead, it might be indicative of general working-memory load.

The biphasic P600/L-AN pattern found in this study is not unique; Steinhauer et al. (2010), for example, reported a similar pattern in their study on NPIs and truth-conditional processing. Building on previous research of a working-memory load interpretation on anterior negativities (e.g. Shao & Neville, 1998), they suggested two interpretations for the L-LAN in their data: that it reflects "(i) a search for a licensor and (ii) maintenance of unintegrated material" (Steinhauer et al., 2010, p. 1538). In their study, the failed attempt to integrate an unlicensed NPI into the current discourse first yielded a P600 response; then, the P600 was followed by a L-LAN, which was taken to either reflect the load on working memory of having to search through the parsed discourse for an appropriate licensor, or a similar load of having to keep track of the unintegrated NPI, or both (Steinhauer et al., 2010, p. 1539). Other scholars draw similar conclusions; Molinaro et al. (2011) suggest that anterior, bilaterally distributed negativities are specifically indicative of working-memory load (p.916; see also Van Berkum et al., 2007, p. 162). The Steinhauer et al. explanation of the L-AN appears somewhat similar to the Barkley et al. (2015) explanation of the LAN, in that both components supposedly reflect a backward-looking process of searching through the parsed discourse for an appropriate dependency resolver (for Steinhauer et al.: an NPI licensor, for Barkley et al.: an appropriate antecedent). This interpretation of the L-AN would suggest that the predicate condition in this study, regardless of the gender-feature match, triggered a search through the previous discourse for an appropriate antecedent.

Given that both hypotheses predicted a main effect of gender mismatch on the LAN/P600 amplitudes for the referential NPs, the lack of any L-AN effect for the referential mismatching NP is noteworthy. The effect of gender mismatch did, as reported in Chapter 4, modulate a P600 response for the gender-mismatching referential NP relative to its gender-matching counterpart. The gender incongruency should render the referential NP unavailable as antecedent to the anaphoric pronoun and could therefore be expected to yield a similar L-AN response as the one observed for the illicit predicate antecedents. It is therefore possible that some methodological issues are responsible for the lack of an L-AN effect for the gender-incongruent referential NP. These issues are discussed in more detail in section 5.3.

It is also important to stress that the L-AN analysis was conducted post-hoc, based on a visual inspection of the grand average EEG plots. In other words, the data was analysed via

an inductive, rather than deductive, approach, and the results cannot directly relate to the test implications of either hypothesis. Moreover, testing significance for something only after it is observed can increase the risk of generating false positives, and the main conclusions of this study are therefore based on the P600-analyses, while still recognising the present L-AN effect.

The data from the P600 latency window and ROI showed significant positive deflections for both predicate conditions, at least compared to the baseline condition (containing a referential, gender-congruent antecedent to the pronoun). In line with previous accounts on the P600, these deflections should be interpreted as both predicate complements being more difficult to process than the referential equivalents. In other words, it seems as if the pronouns in the predicate conditions were overall difficult to integrate into the current discourse, most likely as the predicate NPs were recognised as implausible antecedents in the pronoun-resolution setting. The observed P600 effect, reflecting integration difficulty, therefore entails that predicate nominals are not *as* salient or accessible as referential nominals during online pronoun resolution (cf, Ariel, 1990; Gundel et al., 1993). These results are perhaps not unexpected; nominal predicates are, as has been discussed, non-referential and do not introduce any new referent to the discourse.

Although it might appear evident that predicate NPs should be realised as impossible antecedents during online pronoun resolution, the results indicated a significant effect of gender congruency on the observed EEG amplitude in the P600 latency window and ROI for the predicate NPs, as shown in Table 4. Much like for the referential-NP condition, the predicate NP that was gender incongruent with the subsequent pronoun elicited a significant increase in positivity relative to not only the baseline condition, but also its gender-congruent counterpart. This could be rephrased as the P600 amplitude being significantly *reduced* for the predicate condition when a gender-match was present, as compared to the mismatch condition. These interesting results are in alignment with previous studies reporting reduced P600 amplitudes for agreement attraction or intrusion effects (e.g. Xiang et al., 2009) and could thus be interpreted as the gender component contributing to some kind of illusory licensing of the personal pronoun (cf. Badecker & Straub, 2002; Kennison, 2003). Contrary to Xiang et al. (2009), as well as the test implications for H1, the predicate condition with an intrusive, feature-matching licensor still rendered a significant P600 effect against baseline. Notwithstanding the reduced amplitude in the gender-matching condition, the present and significant main effect of NP-type on the P600 amplitude still seems to imply that both predicate nominals were difficult

to process, and that a subsequent pronoun was difficult -if at all possible- to integrate. Moreover, the working-memory load account that Steinhauer et al. (2010) suggested for the L-AN assimilates quite well with the observed P600-effect for the predicate conditions; that is, the attempt to resolve the pronoun failed (as reflected by the P600), and the pronoun then had to be kept in memory, unintegrated (as reflected by the L-AN). Nonetheless, the significantly less pronounced P600 effect for the gender congruent condition for the predicate NP as compared to the incongruent condition still indicates that the nominal predicate, at least initially, was considered as a potential antecedent to the personal pronoun. These results are ultimately in line with the predictions made for H1.

The follow-up analyses of the P600 in different latency windows found that the two different P600 effects (elicited by gender-congruency and NP-type manipulation, respectively) did in fact not differ with regard to when they first appeared significant, between conditions. Focusing on the potential difference in P600 onset in the predicate and referential conditions, the effect onset analysis could show that the P600 effect elicited by gender incongruency between the intruding NP and the pronoun appeared significant in the same latency window, for both the referential and predicate condition (500-700 ms). That is, the effect of the gender-congruency manipulation on when the P600 effect became significant did not depend on NP-type, and the two P600 effects can thus be regarded similar with reference to their onset.

Although many scholars recognise that the P600 can vary with regard to onset latency (e.g. Brouwer et al., 2012; Gouvea et al., 2010; Delogu et al., 2019; Steinhauer et al., 2010), exactly what cognitive processes such latency variations reflect is not entirely clear. Gouvea et al. (2010) suggest a positive correlation between the P600 onset and the ease with which entities necessary to compute dependency-resolution are retrieved. They had found that the P600 appeared later in a condition with illegal subject-verb agreement relative to an ambiguous garden-path condition. That is, the 'obvious' incongruence seemed to *delay* the P600 onset, which they reasoned was because "correct subject verb agreement provides a more effective retrieval cue than incorrect agreement" (p. 176). However, Gouvea et al. also note that the P600, as a response to agreement-violations, will not onset until the violation has been detected (p. 180), and that a delayed P600 onset "reflects the difficulty of completing the diagnosis or reanalysis of the anomalous structure" (Gouvea et al., 2010, p. 154). By that account, the P600 effect onset will be delayed if the general violation detection is delayed. Although not statistically supported, the results of the effect significance onset analysis hint at a potentially

earlier P600 onset for the NP-type manipulation than for the gender-congruency manipulation, as the former appeared near-significant already in the 400-600 ms window. This could potentially provide further support for the predicate NPs being quickly recognised as poor antecedents to the pronoun, realised as an early onset of the P600. The manipulation for NP type could then potentially be considered a 'more effective retrieval cue' than the gender-congruency manipulation (cf. Gouvea et al., 2010). I will, however, refrain from drawing any strong conclusions based solely on the P600 significance onset analysis, as the differences observed did not reach significance.

In summary, both predicate-NP conditions gave rise to a significant P600 effect relative to the referential-NP conditions. This P600 effect can be taken to reflect that the predicate NPs were realised as relatively poor antecedents for pronoun resolution purposes. Moreover, the subsequent significant L-AN for both the gender congruent and incongruent predicate condition relative to both referential conditions could be interpreted as the predicate conditions resulting in a taxing working memory load of keeping track of the unintegrated item (the personal pronoun). However, despite the nominal predicates being realised as poor antecedents to anaphoric pronouns (as reflected by the substantial P600 effect for both predicate conditions), gender-congruency between the predicate NP and a subsequent anaphor did give rise to a temporary agreement-attraction effect. Thus, contrary to the predictions of H2, the effect of gender-congruency manipulation on the P600 amplitude was the same in both the predicate and referential conditions. Hence, the results at least suggest that even non-referential, predicate NPs are considered among the set of potential antecedents during the initial stages of online reference resolution. Their predicate function, at least initially, does not seem to render them unavailable candidates as referents for anaphora.

5.2 Behavioural Results

As mentioned in Chapter 3, the comprehension questions in this study mainly served the purpose of keeping the participants' attention to the stimuli, wherefore the response accuracy rates were not treated as a main dependent variable. Hence, no predictions were made as to how accuracy rates would vary as dependent on either independent variable. The behavioural results should therefore not be treated as main findings from this study, as the study was not designed with a behavioural-variable triangulation in mind.

The behavioural results indicate that the gender mismatch in the predicate condition yielded significantly lower response accuracy rates than the gender-matching condition. These results could point to the gender incongruent predicate condition being more difficult to process than the gender congruent predicate condition, as psycholinguistic research often uses response accuracy scores as a diagnostic of processing difficulty/ease. However, it is important to remember that the number of questions per condition was low (n=40), and that both the mean accuracy rates and standard deviations were quite high (see Table 8). In other words, participants did overall not seem to have much difficulty answering the questions correctly, but accuracy rates did vary between participants. It would therefore be dubious to claim that the gender-mismatch substantially hindered sentence-comprehensibility for the participants; replication in a larger setting with an explicit behavioural-measure triangulation design would be needed to make such claims.

5.3 Limitations of the Study

When the main results and findings have been discussed, it is also of relevance to mention and discuss some of the limitations of this study. A first note regards the stereotypically gendered nouns, which, as mentioned, were not controlled with norming procedures prior to running the experiment. As the degree of gender bias for the stereotypically gendered nouns was based on reports from previous studies, at least one almost 30 years old, it is possible that some nouns used in this study are not as prominently biased for gender as they once have been, or that their gender stereotype has changed. Nevertheless, as the results did find significant effects of gender congruency manipulation (in that the gender mismatch rendered a significant positivity in the 500-1000 ms time window for both the referential and predicate conditions), the stereotypically gendered nouns do seem to have had an overall impact on the modulation of the P600 effect.

A second note regards the verbs used to construct the referential condition sentences, which were not controlled for their Implicit Causality (IC) bias prior to the experiment. It is possible that any IC inconsistencies that were overlooked in the stimuli sentences may have affected the EEG patterns for the referential conditions extraneously. For example, as mentioned in section 5.1, there was no significant effect of gender congruency manipulation on the L-AN amplitude for the referential condition. For the predicate NPs, the presence of the

L-AN response was interpreted as indicative of the working memory load of having to search through prior discourse for a suitable antecedent, as well as the load of keeping track of an unintegrated discourse item (the personal pronoun). As the intruding NP in the referential gender-mismatching condition should technically also be an inappropriate antecedent, it is surprising that this condition did not give rise to any significant L-AN effect – especially as there was a significant effect of gender-mismatch on the P600 amplitude for the same condition.

While somewhat speculative, a possible explanation for the lack of any L-AN response for the referential gender-congruent condition could be the potential presence of IC bias inconsistencies in the material, as the IC bias of the select transitive verbs was not explicitly controlled for. If both referential NP conditions contained different incongruencies that are known to affect ERP responses (for the referential gender-matching condition: IC bias inconsistencies; for the referential gender-*mis*matching condition: IC bias inconsistency *and* antecedent-pronoun gender incongruency), the potential contrast in L-AN amplitude between these conditions may have been attenuated. However, IC inconsistencies have mainly been found to modulate P600 responses (e.g. Van Berkum et al., 2007), and the current explanation for the lack of a L-AN response for the referential gender-incongruent condition is thus highly tentative. A follow-up study in which the IC bias of the experimental material is explicitly controlled for would thus be appropriate to conduct. The results of such a study could then also give further insight into whether IC inconsistencies modulate late anterior negativities in the first place.

A third note regards the possibility of alternative interpretations of the reduced P600 in the gender-matching predicate condition. Hitherto, there has been no mention of the fact that the predicate gender-matching condition in fact featured an additional incongruency that participants had to process. Although there was a gender-feature match between the predicate NP and the pronoun, there was an earlier gender-feature mismatch between the proper noun subject and the predicative NP, as in *Dylan was a lady but she*…. The incremental unfolding of the sentences allowed the participants to construct and store mental representations of the different discourse entities as they were encountered, as well as update these as more information became available. As a predicate NP provides information about its predicand, it is possible that the gender-mismatch between the subject and the predicative NP triggered a revision of the assumed gender of the subject, altogether. That is, the confusion of *Dylan* being a lady could possibly have forced the participants to update their mental representation of

Dylan: from first having perceived *Dylan* as male, the participants would then consider him female. A subsequent feminine anaphoric pronoun would then perhaps not be too difficult to integrate into the discourse, as the pronoun then would agree with the, now updated, mental representation of *Dylan*. However, the predicate gender-matching condition still yielded a significant P600 response against the referential equivalent, indicating that pronoun resolution still was taxing in this condition. In other words, although the mental representation of *Dylan* was potentially updated, it did not seem to allow for unproblematic pronoun-resolution. Nonetheless, an analysis of the ERPs after the onset of the predicate noun would be needed to further discuss this alternate interpretation of the reduced P600 for the predicate gender-matching condition. As of now, the fact remains that the results indicate that the presence of a predicate NP which was gender-congruent with a later-occurring anaphoric pronoun effectively modulated the amplitude of a P600 effect visible after pronoun onset.

6. Conclusion

This study set out to investigate to what extent predicate NPs are stored and accessed from the mental model as potential antecedents during online anaphoric dependency computations. The study built on theories on a mental model, discourse accessibility and salience, reference resolution and agreement attraction, as well as theories of two well-researched EEG components commonly observed in studies with similar focus (the P600 and LAN components). Furthermore, the study related to a relatively new research field in English linguistics (that of semantically incorporated nouns) and investigated a related but hitherto under-research, even un-investigated, grammatical phenomena: nominal predicates. With inspiration from a previous study (Wittenberg & Trotzke, 2021), it was hypothesized that nominal predicates were either disregarded as potential antecedents in the mental model due to their predicate function, or that their morphosyntactic form would render them accessible enough to (illicitly) license an anaphoric pronoun. The hypotheses were spelled out with specific test implications that were tested using an EEG approach.

The results seem to indicate that predicate NPs, to some extent, are considered among a potential set of possible antecedents during online pronoun resolution, as shown by the fact that gender congruency between a predicate NP and a subsequent pronoun modulated the amplitude of a P600 component visible 500-1000 ms after pronoun onset. Both predicate NP conditions, however, did indeed yield a substantial P600 effect relative to the baseline condition, indicating that the pronouns were difficult to resolve and/or integrate into the current discourse. Although the study failed to replicate the biphasic LAN/P600 response commonly reported in studies on morphosyntactic violations, the analyses identified a sustained anterior negativity, a L-AN effect, significant after 700 ms post pronoun onset in the predicate conditions. In alignment with previous accounts of anterior negativities, the L-AN was interpreted as indicative of working memory load, possibly an extension of the fact that pronoun resolution failed for the predicate NPs and that the pronoun had to be kept unintegrated in the mental model. However, no L-AN effect was found for the referential gender-incongruent condition.

The results ultimately give rise to new questions how (morpho-)syntactic and semantic information interact during online pronominal resolution, as well as what electrophysiological

responses they give rise to. The observed results thereby extend previous indications of structurally unavailable entities being regarded potential antecedents during online reference resolution (e.g. Kennison, 2003) to a new field, namely that of non-referential entities. A possible extension to the current study would be to conduct a cross-linguistic examination/comparison of how predicate NPs in other languages than English are treated during online pronoun resolution. Nominal predicates in Swedish, for example, are not realised as full determiner phrases and are therefore morphosyntactically distinct from referential NPs in that regard. Comparing Swedish referential and predicate NPs in a similar experimental paradigm could thus further investigate to what extent the morphosyntactic form of nominal predicates contribute to the presence/absence of any agreement attraction effects during online pronoun resolution. Furthermore, extending the current study with an explicit focus on what type of anterior negativities predicate NPs give rise to in pronoun resolution paradigms would shed more light on the difference between, for example, the Nref and the L-(L)AN.

References

Aarts, B. (2014). “predicand”. In *The Oxford Dictionary of English Grammar*. Oxford University Press. Retrieved August 2, 2023, from https://www-oxfordreference-com.ludwig.lub.lu.se/view/10.1093/acref/9780199658237.001.0001/acref-9780199658237-e-1892.

Adger, D. (2003). *Core Syntax: A Minimalist Approach*. Oxford University Press.

Adger, D. & Ramchand, G. (2003). Predication and Equation. *Linguistic Inquiry*, *24*(3), 325-359. https://www.jstor.org/stable/4179238

Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An ERP study. *Language and Cognitive Processes*, *18*(4), 405–430. https://doi.org/10.1080/01690960244000054

Ariel, M. (1990). *Accessing NP antecedents*. Routledge (Croom Helm Linguistics Series.)

Ariel, M. (1991). The function of accessibility in a theory of grammar. *Journal of Pragmatics*, *16*(5), 443-463. https://doi.org/10.1016/0378-2166(91)90136-L

Badecker, W. & Straub, K. The Processing Role of Structural Constraints on the Interpretation of Pronouns and Anaphors. *Journal of Experimental Psychology*: *Learning, Memory and Cognition*, *28*(4), 748-769. https://doi.org/10.1037/0278-7393.28.4.748

Barkley, C., Kluender, R. & Kutas, M. (2015). Referential processing in the human brain: an Event-Related Potential (ERP) study. *Brain Research*, *1629*, 143-159. https://doi.org/10.1016/j.brainres.2015.09.017

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01.

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*, *5*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bock, K., & Miller, C. (1991). Broken agreement. *Cognitive Psychology*, *23*, 45–93. https://doi.org/10.1016/0010-0285(91)90003-7

Bornkessel-Schlesewsky, I. & Schlesewsky, M. (2008). An alternatice perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews, 59*(1), 55-73. https://doi.org/10.1016/j.brainresrev.2008.05.003

Bowers, J. The Syntax of Predication. *Linguistic Inquiry*, *24*(4), 591-656. https://www.jstor.org/stable/4178835

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *NeuroReport: For Rapid Communication of Neuroscience Research*, 18(*17*), 1851–1854. https://doi.org/10.1097/WNR.0b013e3282f1a999

Butler, C. (2003). *Structure and Function: a guide to three major structural-functional theories. Part 1, Approaches to the simplex clause.* John Benjamins.

Caffarra, S., Mendoza, M. & Davidson, D. (2019). Is the LAN effect in morphosyntactic processing an ERP artefact? *Brain and Language*, *191*, 9-16. https://doi.org/10.1016/j.bandl.2019.01.003

Caramazza, A., Grober, E., Garvey, C. & Yates, J. (1977). Comprehension of Anaphoric Pronouns. *Journal of Verbal Learning and Verbal Behavior*, *16*, 601-709. https://doi.org/10.1016/S0022-5371(77)80022-4.

Carreiras, M., Garnham, A., Oakhill, J. & Cain, K. (1996). The Use of Stereotypical Gender Information in Constructing a Mental Model: Evidence from English and Spanish. *The Quarterly Journal of Experimental Psychology*, *49*(3), 639-663. https://doi.org/10.1080/713755647

Chomsky, N. (1981). *Lectures on government and binding: the Pisa lectures*. Foris.

Ciaccio, L.A., Bürki, A. & Clahsen, H. (2023). Inter-individual variability in morphological processing: An ERP study on German plurals. https://doi.org/10.1016/j.jneuroling.2023.101138

Coulson, S., King, J. W., and Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, *13*(1), 21–58. https://doi.org/10.1080/016909698386582

Delogu, F., Brouwer, H. & Crocker, M. W. (2019). Event-related potentials index lexical retrieval (N400) and integration (P600) during language comprehension. *Brain and Cognition*, *135*, 1-14. https://doi.org/10.1016/j.bandc.2019.05.007

Delorme, A. & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134,* 9-21.

Dillon, B., Mishler, A., Sloggett, S. & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*, 85-103. http://dx.doi.org/10.1016/j.jml.2013.04.003

Duffy, S.A. & Keir, J.A. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition, 32*(4), 551-559. https://doi.org/10.3758/BF03195846

Erlich, K. (1980). Comprehension of pronouns. *Quarterly Journal of Experimental Psychology*, *32*, 247-255.

Federmeier, K. D., Kluender, R. & Kutas, M. (2003). Aligning Linguistic and Brain Views on Language Comprehension. In A. Zani, A. Mado Proverbio & M. I. Posner (Eds.), *The Cognitive Electrophysiology of Mind and Brain* (pp. 143-168). Academic Press.

Friederici, A.D., Pfeifer, E. & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*(3), 183–192. https://doi.org/10.1016/0926-6410(93)90026-2

García-Sierra, A., Silva-Pereyra, J., Alatorre-Cruz, G. C. & Wig, N. (2021). An Event-Related Brain Potential (ERP) Study of Complex Anaphora in Spanish. *Frontiers in Psychology*, *12*(625314), 1-15. https://doi.org/10.3389/fpsyg.2021.625314

Garnham, A. (2001). *Mental Models and the Interpretation of Anaphora*. Psychology Press.

Garnham, A., Gabriel, U., Sarrasin, O., Gygax, P. & Oakhill, J. (2012). Gender Representation in Different Languages and Grammatical Marking on Pronouns: When Beauticians, Musicians, and Mechanics Remain Men. *Discourse Processes*, *49*(6), 481-500. https://doi.org/10.1080/0163853X.2012.688184

Garnham, A. & Oakhill, J. (1985). On-line resolution of anaphoric pronouns: Effects of inference making and verb semantics. *British Journal of Psychology*, *76*, 385-393. https://doi.org/10.1111/j.2044-8295.1985.tb01961.x

Gosselke Berthelsen, S., Horne, M., Shtyrov, Y. & Roll., M. (2020). Different neural mechanisms for rapid acquisition of words with grammatical tone in learners from tonal and non-tonal backgrounds: ERP evidence. *Brain Research*, *1729*, 1-15. https://doi.org/10.1016/j.brainres.2019.146614

Gouvea, A. C., Phillips, C., Kazanina, N. & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, *25*(2), 149-188. https://doi.org/10.1080/01690960902965951

Green, J.N. (1988). Spanish. In M. Harris & N. Vincent (Eds.), *Romance Languages* (pp. 79-130). Croom Helm.

Grice, P. (1991). *Studies in the Way of Words*. Harvard University Press.

Gundel, J.K., Hedberg, N. & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, *69*(2), 274-307. https://doi.org/10.2307/416535

Hall, K. & Yoshida, M. (2021). Coreference and parallelism. *Language, Cognition and Neuroscience*, *36*(3), 296-319. https://doi.org/10.1080/23273798.2020.1827154

Hagoort, P., Brown, C. & Groothusen, J. (1993). The syntactic positive shift (SPS) as an erp measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), 439-483. https://doi.org/10.1080/01690969308407585

Haviland, S.E. & Clark, H.H. (1974). What's New? Acquiring New Information as a Process in Comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512-521. https://doi.org/10.1016/S0022-5371(74)80003-4

Heinat, F. & Klingvall, E. (2020). Set focus and anaphoric reference: An ERP study. *Brain and Language*, *206*, 1-6. https://doi.org/10.1016/j.bandl.2020.104808

Huang, Y. (2000). *Anaphora.* Oxford University Press.

Huang, Y. (2015). Anaphora. In J.D. Wright (Ed) *International Encyclopedia of the Social & Behavioral Sciences* (2nd Edition), pp. 674-679. Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.53001-3

Huddleston, R. & Pullum, G.K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Johnson-Laird, P.N. (1980). Mental Models in Cognitive Science. *Cognitive Science*, 4, 71-115. https://onlinelibrary.wiley.com/doi/pdf/10.1207/s15516709cog0401_4

Johnson-Laird, P.N. (1981). Comprehension as the Construction of Mental Models. *Philosophical Transactions of the Royal Society of London*, *295*(1077), 353-374. https://www.jstor.org/stable/2395747

Kaan, E., Harris, A., Gibson, E. & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159-201. https://doi.org/10.1080/016909600386084

Kaan, E. & Swaab, T. Y. (2003). Electrophysiological evidence for serial sentence processing: a comparison between non-preferred and ungrammatical continuations. *Cognitive Brain Research*, *17*(3), 621-635. https://doi.org/10.1016/S0926-6410(03)00175-7

Kaiser, E., Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition*, *112*(1), 55–80. https://doi.org/10.1016/j.cognition.2009.03.010

Kaiser, E. (2013). Experimental paradigms in psycholinguistics. In R.J. Podesva & D. Sharma (Eds.), *Research Methods in Linguistics,* 135-168. Cambridge: Cambridge University Press.

Kemmerer, D. (2015). *Cognitive Neuroscience of Language*. New York/London: Psychology Press.

Kennison, S. M. (2003). Comprehending the pronouns *her*, *him*, and *his*: Implications for theories of referential processing. *Journal of Memory and Language*, *49*, 335-352. https://doi.org/10.1016/S0749-596X(03)00071-8

Kennison, S. M. & Trofe, J. L. (2003). Comprehending Pronouns: A Role for Word-Specific Gender Stereotype Information. *Journal of Psycholinguistic Research*, *32*(3), 355-378. https://doi.org/10.1023/A:1023599719948

King, J. W., and Kutas, M. (1998). He really is a nurse: ERPs and anaphoric coreference. *Psychophysiology*, *35*(Suppl. 1), S47.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363-394. https://doi.org/10.1037/0033-295X.85.5.363

Kluender, R. & Kutas, M. (1993). Bridging the Gap: Evidence from ERPs on the Processing of Unbounded Dependencies. *Journal of Cognitive Neuroscience*, *5*(2), 196-214. https://doi.org/10.1162/jocn.1993.5.2.196

Kreiner, H., Sturt, P. & Garrod, S. (2008). Processing definitional and stereotypical gender in reference resolution: Evidence from eye-movements. *Journal of Memory and Language*, *58*, 239-261. https://doi:10.1016/j.jml.2007.09.003

Kuperberg, G. R., Sitnikova, T., Caplan, D. & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*, 117-129. https://doi.org/10.1016/S0926-6410(03)00086-7

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B. (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, *82*(13), 1–26. doi:10.18637/jss.v082.i13.

Lenth, R. (2018). *Emmeans: Estimated marginal means, aka least-squares means* [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=emmeans

Luck, S.J. (2014). *An introduction to the Event-Related Potential Technique*. Cambridge, MA: The MIT Press.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in human neuroscience*, *8*, 213.

Molinaro, N., Barber, H. A. & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, *47*(8), 908-930. https://doi.org/10.1016/j.cortex.2011.02.019

Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3*(2), 151–165. https://doi.org/10.1162/jocn.1991.3.2.151

Nieuwland, M. S. (2014). "Who's he?" Event-related brain potentials and unbound pronouns. *Journal of Memory and Language*, *76*, 1-28. https://doi.org/10.1016/j.jml.2014.06.002

Oakhill, J., Garnham, A. & Reynolds, D. (2005). Immediate activation of stereotypical gender information. *Memory & Cognition,* *33*(6), 972-983. https://doi.org/10.3758/BF03193206

Osterhout, L., Bersick, M. & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition, 25*(3), 273-285. https://doi.org/10.3758/BF03211283

Osterhout, L. & Holcomb, P. J. (1992). Event-Related Brain Potentials Elicited by Syntactic Anomaly. *Journal of Memory and Language*, *31*(6), 785-806. https://doi.org/10.1016/0749-596X(92)90039-Z

Osterhout, L., Mckinnon, R. , Bersick, M., & Corey, V. (1996). On the language-specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family? *Journal of Cognitive Neuroscience*, *8*, 507–526. https://doi.org/10.1162/jocn.1996.8.6.507

Osterhout, L. & Mobley, L. A. (1995). Event-Related Brain Potentials Elicited by Failure to Agree. *Journal of Memory and Language*, *34*(6), 739-773. https://doi.org/10.1006/jmla.1995.1033

Oxford University Press. (2022). ”pronoun, n.”. In *Oxford English Dictionary Online*. Retrieved January 31, 2023 from https://www-oed-com.ludwig.lub.lu.se/view/Entry/152548?redirectedFrom=pronoun

Oxford University Press. (2023a). “-ess, suffix[1]”. In *Oxford English Dictionary Online*. Retrieved July 14, 2023 from https://doi.org/10.1093/OED/5503524576

Oxford University Press. (2023b). “man, n.[1]”. In *Oxford English Dictionary Online*. Retrieved July 14, 2023 from https://doi-org.ludwig.lub.lu.se/10.1093/OED/1034670116

Peirce, J., Gray, J., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behaviour Research Methods.* https://doi:10.3758/s13428-018-01193-y

Reali, C., Esaulova, Y., Öttl, A. & von Stockhausen, L. (2015). Role descriptions induce gender mismatch effects in eye movements during reading. *Frontiers in Psychology*, *6*(1607), 1-13. https://doi:10.3389/fpsyg.2015.01607

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reinhart, T. (1983). *Anaphora and semantic interpretation*. London: Croom Helm.

Roy, I. (2013). *Nonverbal Predication: Copular Sentences at the Syntax-Semantics Interface*. Oxford University Press.

Saeed, J.I. (2016). *Semantics* (4th ed). Wiley-Blackwell.

Sauermann, A. & Gagarina, N. (2017). Grammatical Role Parallelism Influences Ambiguous Pronoun Resolution in German. *Frontiers in Psychology*, *8*(1205), 1-8. https://doi:10.3389/fpsyg.2017.01205

Scholten, J., Aguilar-Guevara, A., 2010. Assessing the discourse referential properties of weak definite NPs. *Linguistics in the Netherlands*, *27*(1), 115-128. https://doi.org/10.1075/avt.27.10sch

Shao, J., & Neville, H. (1998). Analyzing semantic processing using event-related potentials. *Newsletter for the Center for Research in Language*, *11*(5), 3–20.

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. https://doi.org/10.1016/S0022-5371(74)80064-2

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, *23*(3), 197–229. https://doi.org/10.1007/BF02139085

Steinhauer, K., Drury, J. E., Portner, P., Walenski, M. & Ullman, M. T. (2010). Syntax, concepts and logic in the temporal dynamics of language comprehension: Evidence from event-related potentials. *Neuropsychologia*, *48*(6), 1525-1542. https://doi.org/10.1016/j.neuropsychologia.2010.01.013

Swaab, T.Y., Ledoux, K. Camblin, C.C. & Boudewyn, M.A. (2011). In S.J. Luck & E.S. Kappenman (Eds.) *Oxford handbook of event-related potential components*. Oxford Academic.

Tanner, D. (2015). On the left anterior negativity (LAN) in electrophysiological studies of morphosyntactic agreement: A Commentary on "Grammatical agreement processing in reading: ERP findings and future directions" by Molinaro et al., 2014. *Cortex*, *66*, 149-155. https://doi.org/10.1016/j.cortex.2014.04.007

Tanner, D. & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, *56*, 289-301. https://doi.org/10.1016/j.neuropsychologia.2014.02.002

The MathWorks Inc. (2022). MATLAB version: 9.12.0 (R2022a), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com

Van Berkum, J.J.A. (2004). Sentence comprehension in a wider discourse: can we use erps to keep track of things?. In: Carreiras Jr., M & Clifton, C. (Eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond*. Psychology Press.

Van Berkum, J.J.A. (2009). The Neuropragmatics of 'Simple' Utterance Comprehension: an ERP Review. In U. Sauerland & K. Yatsushiro (Eds.), *Semantics and Pragmatics: From Experiment to Theory* (pp. 276-317). Palgrave.

Van Berkum, J.J.A., Koorneef, A.W., Otten, M. & Nieuwland, M.S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, *1146*, 158-171. https://doi.org/10.1016/j.brainres.2006.06.091

Van Berkum, J.J.A., Zwitserlood, P., Bastiaansen, M.C.M., Brown, C.M. & Hagoort, P. (2004). So who's "he" anyway? Differential ERP and ERSP effects of referential success, ambiguity and failure during spoken language comprehension. *Annual meeting of the Cognitive Neuroscience Society (CNS-2004)*, San Francisco, April 18–20.

van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.

van Geenhoven, V. (1998). *Semantic Incorporation and Indefinite Descriptions: Semantic and Syntactic Aspects of Noun Incorporation in West Greenlandic*. Dissertation. Stanford: CSLI.

van Geenhoven, V. (2002). Raised Possessors and Noun Incorporation in West Greenlandic. *Natural Language & Linguistic Theory*, *20*(4), 759-821. https://www.jstor.org/stable/4047999

van Herten, M., Kolk, H. H. J. & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22*(2), 241-255. https://doi.org/10.1016/j.cogbrainres.2004.09.002

Von Heusinger, K. & Schumacher, P.B. (2019). Discourse prominence: Definition and application. *Journal of Pragmatics*, *154*, 117-127. https://doi.org/10.1016/j.pragma.2019.07.025

Wagers, M., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*, 206–237. https://doi.org/10.1016/j.jml.2009.04.002

Wilbourn, M. P., & Kee, D. W. (2010). Henry the nurse is a doctor too: Implicitly examining children's gender stereotypes for male and female occupational roles. *Sex Roles: A Journal of Research*, *62*(9-10), 670–683. https://doi.org/10.1007/s11199-010-9773-7

Winter, B. (2020). *Statistics for Linguistis: An Introduction Using R.* Routledge.

Wittenberg, E. & Trotzke, A. (2021). Semantic incorporation and discourse prominence: Experimental evidence from English pronoun resolution. *Journal of Pragmatics*, *186*, 87-99. https://doi.org/10.1016/j.pragma.2021.09.019

Xiang, M., Dillon, B. & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain & Language*, *108*, 40-55. https://doi.org/10.1016/j.bandl.2008.10.002

Appendix A

Lists of nouns and verbs used to construct the experimental material.

Stereotypically gendered nouns		**Definitionally gendered nouns**	
Female	**Male**	**Female**	**Male**
Nurse	Surgeon	Governess	Bachelor
Paralegal	Lawyer	Princess	Prince
File-clerk	Judge	Baroness	Baron
Florist	Butcher	Girl	Boy
Secretary	Politician	Queen	King
Housekeeper	Blacksmith	Milkmaid	Milkman
Librarian	Scientist	Housewife	Gentleman
Receptionist	Senator	Chairwoman	Chairman
Obstetrician	Engineer	Cowgirl	Cowboy
Cosmetologist	Minister	Widow	Pope
Hairdresser	Barber	Midwife	Businessman
Cleaner	Farmer	Girl scout	Boy scout
Beautician	Electrician	Ballerina	Bellboy
Figure-skater	Plumber	Woman	Man
Babysitter	Doctor	Wife	Husband
Model	Robber	Hostess	Host
Manicurist	Firefighter	Saleswoman	Salesman
Dressmaker	Technician	Lady	Lord
Fortune-teller	Magician	Maid	Butler
Cashier	Janitor	Duchess	Duke
Flight-attendant	Pilot	Stewardess	Repairman
Typist	Astronaut	Bride	Groom
Dietician	Detective	Spinster	Deacon
Kindergarten-teacher	Dentist	Landlady	Landlord
Nanny	Mechanic	Priestess	Priest
Caterer	Chef	Choirgirl	Mailman
Ballet-dancer	Paratrooper	Matriarch	Patriarch
Wedding-planner	Tailor	Policewoman	Policeman
Cheerleader	Mayor	Nun	Monk
Canteen-assistant	Chauffeur	Mother	Father
Advice-columnist	Bartender	Stepmother	Stepfather
Stripper	Clown	Waitress	Waiter
Beauty-consultant	Dean	Headmistress	Headmaster
Caregiver	Diplomat	Actress	Actor
Daycare-manager	Hunter	Countess	Count
Stenographer	Lumberjack	Congresswoman	Congressman

Feminist	Ballplayer/footballer	Clergywoman	Clergyman
Groupie	General	Newswoman	Newsman
Gymnast	Drummer	Noblewoman	Nobleman
Interior-decorator	Explorer	Camerawoman	Cameraman

(Semi-)copular verbs	
Be	Train as
Become	Turn into
Work as	Was considered
Remain	Was perceived as
Look like	Was hired as
Act like	Was employed as
Feel like	Was educated as
Behave like	Continued as
Stay	Persevered as
Intern as	Freelance as
Excel as	Wind up as
Seem like	Struggle as
Appear as	Was content as
Volunteered as	Was born

Transitive verbs							
Advise	Chase down	Debate against	Fancy	Identified	Laugh with	Pray for	Teach
Amaze	Call	Dine with	Fight with	Imitate	Live with	Party with	Texted
Asked	Call for	Dance with	Fool	Impress	Listen to	Quiz	Telephone
Avoid	Complain to	Dazzle	Go out with	Interview	Marry	Question	Talk to
Argue with	Comfort	Desire	Gossip about	Intimidate	Meet up with	Reassure	Tutor
Arm-wrestle with	Court	Drank with	Greet	Interrogate	Meet	Run into	Train with
Arrest	Cook dinner for	Disagree with	Hire	Invite	Negotiate with	Spy on	Train
Admire	Congratulate	Dislike	Hurt	Ignore	Notice	Sublet to	Traced down
Apologize	Consult	Eavesdrop on	Have	Insult	Need	See	Travel to
Be introduced to	Cuddle with	Embrace	Help	Join	Nominate	Surprise	Travel with
Be friendly with	Chat with	Encourage	Host	Joke around with	Oppose	Spoke with	Upset

Be related to	Care for	Entertain	Hug	Kiss	Observe	Spot	Trick
Be married to	Creeped on	Employ	Hate	Kick	Owe money	Say hello to	Warn
Befriend	Criticize	Encounter	Fire	Know	Offend	Sing for	Visit
Bother	Compete with	Have a crush on	Hang out with	Know of	Perform for	Shado w	Work with
Bring	Commute with	Frighten	Have tea with	Love	Propose to	Study with	Wrote music with
Badmouth	Charm	Fund	Have lunch with	Long for	Pine for	Stop by	Wave to
Banter with	Date	Follow	Have dinner with	Look for	Play soccer with	Sweet-talk	Want to meet with
Bully	Discover	Flirt with	Have drinks with	Like	Pray with	Sugges ted	Work for
Book	Dress	Fell in love with	Have eyes for	Laugh at	Pay	Suppor t	Yell at

Appendix B

A list of all experimental stimuli sentences in all conditions, with their corresponding comprehension questions.

Condition	**Sentence**	**Question**
REF-MATCH	Henry saw a nurse at the hospital and she laughed a lot.	
REF-MISMATCH	Henry saw a surgeon at the hospital and she laughed a lot.	
PRED-MATCH	Henry was a nurse at the hospital and she laughed a lot.	
PRED-MISMATCH	Henry was a surgeon at the hospital and she laughed a lot.	
REF-MATCH	Mary said hello to a lawyer and then he took off immediately.	
REF-MISMATCH	Mary said hello to a paralegal and then he took off immediately.	
PRED-MATCH	Mary was employed as a lawyer and then he took off immediately.	
PRED-MISMATCH	Mary was employed as a paralegal and then he took off immediately.	
REF-MATCH	Mike hugged a file-clerk but later she complained about it.	
REF-MISMATCH	Mike hugged a judge but later she complained about it.	
PRED-MATCH	Mike remained a file-clerk but later she complained about it.	
PRED-MISMATCH	Mike remained a judge but later she complained about it.	
REF-MATCH	Lisa kicked a butcher and obviously he didn't like that.	
REF-MISMATCH	Lisa kicked a florist and obviously he didn't like that.	
PRED-MATCH	Lisa worked as a butcher and obviously he didn't like that.	
PRED-MISMATCH	Lisa worked as a florist and obviously he didn't like that.	

REF-MATCH	Andrew spotted a secretary and luckily she was able to help.	
REF-MISMATCH	Andrew spotted a politician and luckily she was able to help.	
PRED-MATCH	Andrew looked like a secretary and luckily she was able to help.	
PRED-MISMATCH	Andrew looked like a politician and luckily she was able to help.	
REF-MATCH	Olga traced down a chef but later he left the country.	Was it a chef that Olga traced down?
REF-MISMATCH	Olga traced down a caterer but later he left the country.	Was it a caterer that Olga traced down?
PRED-MATCH	Olga acted like a chef but later he left the country.	Was it a chef that Olga acted like?
PRED-MISMATCH	Olga acted like a caterer but later he left the country.	Was it a caterer that Olga acted like?
REF-MATCH	Steve knew of a librarian but apparently she didn't read much.	
REF-MISMATCH	Steve knew of a scientist but apparently she didn't read much.	
PRED-MATCH	Steve felt like a librarian but apparently she didn't read much.	
PRED-MISMATCH	Steve felt like a scientist but apparently she didn't read much.	
REF-MATCH	Anna interviewed a senator but unfortunately he didn't get the job.	
REF-MISMATCH	Anna interviewed a receptionist but unfortunately he didn't get the job.	
PRED-MATCH	Anna volunteered as a senator but unfortunately he didn't get the job.	
PRED-MISMATCH	Anna volunteered as a receptionist but unfortunately he didn't get the job.	
REF-MATCH	John trained a ballet-dancer after college but she quit long ago.	
REF-MISMATCH	John trained a paratrooper after college but she quit long ago.	
PRED-MATCH	John was employed as a ballet-dancer after college but she quit long ago.	
PRED-MISMATCH	John was employed as a paratrooper after college but she quit long ago.	
REF-MATCH	Matilda debated against a minister but later he completely switched careers.	Was it a surgeon that Matilda debated against?

REF-MISMATCH	Matilda debated against a cosmetologist but later he completely switched careers.	Was it a surgeon that Matilda debated against?
PRED-MATCH	Matilda was a minister but later he completely switched careers.	Was it a surgeon that Matilda was?
PRED-MISMATCH	Matilda was a cosmetologist but later he completely switched careers.	Was it a surgeon that Matilda was?
REF-MATCH	Russel consulted a hairdresser and luckily she was very good.	Was it a gravedigger that Russel consulted?
REF-MISMATCH	Russel consulted a barber and luckily she was very good.	Was it a gravedigger that Russel consulted?
PRED-MATCH	Russel trained as a hairdresser and luckily she was very good.	Was it a gravedigger that Russel trained as?
PRED-MISMATCH	Russel trained as a barber and luckily she was very good.	Was it a gravedigger that Russel trained as?
REF-MATCH	Jane helped a farmer and understandably he was very thankful.	Did Jane confront a farmer?
REF-MISMATCH	Jane helped a cleaner and understandably he was very thankful.	Did Jane confront a cleaner?
PRED-MATCH	Jane remained a farmer and understandably he was very thankful.	Did Jane confront a farmer?
PRED-MISMATCH	Jane remained a cleaner and understandably he was very thankful.	Did Jane confront a cleaner?
REF-MATCH	Ricky befriended a beautician and later she promised to call.	
REF-MISMATCH	Ricky befriended an electrician and later she promised to call.	
PRED-MATCH	Ricky remained a beautician and later she promised to call.	
PRED-MISMATCH	Ricky remained an electrician and later she promised to call.	
REF-MATCH	Emma tutored a plumber and apparently he became very successful.	Was it a cashier that Emma tutored?
REF-MISMATCH	Emma tutored a figure-skater and apparently he became very successful.	Was it a cashier that Emma tutored?
PRED-MATCH	Emma worked as a plumber and apparently he became very successful.	Was it a cashier that Emma worked as?
PRED-MISMATCH	Emma worked as a figure-skater and apparently he became very successful.	Was it a cashier that Emma worked as?
REF-MATCH	David needed a babysitter at home but she didn't get paid.	
REF-MISMATCH	David needed a doctor at home but she didn't get paid.	

PRED-MATCH	David acted like a babysitter at home but she didn't get paid.	
PRED-MISMATCH	David acted like a doctor at home but she didn't get paid.	
REF-MATCH	Linda identified a robber and then he spent four years in jail.	
REF-MISMATCH	Linda identified a model and then he spent four years in jail.	
PRED-MATCH	Linda looked like a robber and then he spent four years in jail.	
PRED-MISMATCH	Linda looked like a model and then he spent four years in jail.	
REF-MATCH	Mark kissed a model but unfortunately she didn't feel better.	
REF-MISMATCH	Mark kissed a robber but unfortunately she didn't feel better.	
PRED-MATCH	Mark became a model but unfortunately she didn't feel better.	
PRED-MISMATCH	Mark became a robber but unfortunately she didn't feel better.	
REF-MATCH	Erica questioned a doctor and luckily he exuded a lot of confidence.	Was it Randy that questioned a doctor?
REF-MISMATCH	Erica questioned a babysitter and luckily he exuded a lot of confidence.	Was it Randy that questioned a babysitter?
PRED-MATCH	Erica behaved like a doctor and luckily he exuded a lot of confidence.	Was it Randy that behaved like a doctor?
PRED-MISMATCH	Erica behaved like a babysitter and luckily he exuded a lot of confidence.	Was it Randy that behaved like a babysitter?
REF-MATCH	Steven fancied a manicurist but sadly she had to leave the country.	
REF-MISMATCH	Steven fancied a firefighter but sadly she had to leave the country.	
PRED-MATCH	Steven excelled as a manicurist but sadly she had to leave the country.	
PRED-MISMATCH	Steven excelled as a firefighter but sadly she had to leave the country.	
REF-MATCH	Jessica advised a technician for three years and he learnt so much.	
REF-MISMATCH	Jessica advised a dressmaker for three years and he learnt so much.	
PRED-MATCH	Jessica persevered as a technician for three years and he learnt so much.	

PRED-MISMATCH	Jessica persevered as a dressmaker for three years and he learnt so much.	
REF-MATCH	Daniel follows a fortune-teller on Instagram and she is very outspoken.	
REF-MISMATCH	Daniel follows a magician on Instagram and she is very outspoken.	
PRED-MATCH	Daniel perseveres as a fortune-teller on Instagram and she is very outspoken.	
PRED-MISMATCH	Daniel perseveres as a magician on Instagram and she is very outspoken.	
REF-MATCH	Joanne talked to a janitor in Leeds and he admitted that it was a great job.	Did the sentence you just read contain the word "admitted"?
REF-MISMATCH	Joanne talked to a cashier in Leeds and he admitted that it was a great job.	Did the sentence you just read contain the word "admitted"?
PRED-MATCH	Joanne interned as a janitor in Leeds and he admitted that it was a great job.	Did the sentence you just read contain the word "admitted"?
PRED-MISMATCH	Joanne interned as a cashier in Leeds and he admitted that it was a great job.	Did the sentence you just read contain the word "admitted"?
REF-MATCH	Michael photographed a flight-attendant but unfortunately she was quite shy.	
REF-MISMATCH	Michael photographed a pilot but unfortunately she was quite shy.	
PRED-MATCH	Michael was a flight-attendant but unfortunately she was quite shy.	
PRED-MISMATCH	Michael was a pilot but unfortunately she was quite shy.	
REF-MATCH	Claire joked around with an astronaut and obviously he really loved it.	
REF-MISMATCH	Claire joked around with a typist and obviously he really loved it.	
PRED-MATCH	Claire was perceived as an astronaut and obviously he really loved it.	
PRED-MISMATCH	Claire was perceived as a typist and obviously he really loved it.	
REF-MATCH	James interrogated a typist and fortunately she was very forthcoming.	
REF-MISMATCH	James interrogated an astronaut and fortunately she was very forthcoming.	

PRED-MATCH	James became a typist and fortunately she was very forthcoming.	
PRED-MISMATCH	James became an astronaut and fortunately she was very forthcoming.	
REF-MATCH	Eleanor dates a detective but unfortunately he is quite mean.	
REF-MISMATCH	Eleanor dates a dietitian but unfortunately he is quite mean.	
PRED-MATCH	Eleanor works as a detective but unfortunately he is quite mean.	
PRED-MISMATCH	Eleanor works as a dietitian but unfortunately he is quite mean.	
REF-MATCH	Martin flirted with a cashier but eventually she called the police.	
REF-MISMATCH	Martin flirted with a janitor but eventually she called the police.	
PRED-MATCH	Martin acted like a cashier but eventually she called the police.	
PRED-MISMATCH	Martin acted like a janitor but eventually she called the police.	
REF-MATCH	Anne worked with a dentist but apparently he was quite lazy.	
REF-MISMATCH	Anne worked with a kindergarten-teacher but apparently he was quite lazy.	
PRED-MATCH	Anne was hired as a dentist but apparently he was quite lazy.	
PRED-MISMATCH	Anne was hired as a kindergarten-teacher but apparently he was quite lazy.	
REF-MATCH	Benjamin hired a nanny and luckily she knew what to do.	Was it a nanny that Benjamin hired?
REF-MISMATCH	Benjamin hired a mechanic and luckily she knew what to do.	Was it a mechanic that Benjamin hired?
PRED-MATCH	Benjamin continued as a nanny and luckily she knew what to do.	Was it a nanny that Benjamin continued as?
PRED-MISMATCH	Benjamin continued as a mechanic and luckily she knew what to do.	Was it a mechanic that Benjamin continued as?
REF-MATCH	Maria worked for a blacksmith but sadly he wasn't that good.	Was it a butler that Maria worked for?
REF-MISMATCH	Maria worked for a housekeeper but sadly he wasn't that good.	Was it a butler that Maria worked for?
PRED-MATCH	Maria was hired as a blacksmith but sadly he wasn't that good.	Was it a butler that Maria was hired as?

PRED-MISMATCH	Maria was hired as a housekeeper but sadly he wasn't that good.	Was it a butler that Maria was hired as?
REF-MATCH	Jim called for an obstetrician and of course she was quite thrilled.	
REF-MISMATCH	Jim called for an engineer and of course she was quite thrilled.	
PRED-MATCH	Jim became an obstetrician and of course she was quite thrilled.	
PRED-MISMATCH	Jim became an engineer and of course she was quite thrilled.	
REF-MATCH	Pamela looked for a tailor but unfortunately he couldn't help out.	Did Pamela look for a tailor?
REF-MISMATCH	Pamela looked for a wedding-planner but unfortunately he couldn't help out.	Did Pamela look for a wedding-planner?
PRED-MATCH	Pamela looked like a tailor but unfortunately he couldn't help out.	Did Pamela look like a tailor?
PRED-MISMATCH	Pamela looked like a wedding-planner but unfortunately he couldn't help out.	Did Pamela look like a wedding-planner?
REF-MATCH	Kevin booked a wedding-planner yesterday and she did so much to help.	
REF-MISMATCH	Kevin booked a tailor yesterday and she did so much to help.	
PRED-MATCH	Kevin turned into a wedding-planner yesterday and she did so much to help.	
PRED-MISMATCH	Kevin turned into a tailor yesterday and she did so much to help.	
REF-MATCH	Angela yelled at a paratrooper yesterday but he didn't complain.	
REF-MISMATCH	Angela yelled at a ballet-dancer yesterday but he didn't complain.	
PRED-MATCH	Angela turned into a paratrooper yesterday but he didn't complain.	
PRED-MISMATCH	Angela turned into a ballet-dancer yesterday but he didn't complain.	
REF-MATCH	Dwight studied with a figure-skater in Bristol and she loved that lifestyle.	
REF-MISMATCH	Dwight studied with a plumber in Bristol and she loved that lifestyle.	
PRED-MATCH	Dwight freelanced as a figure-skater in Bristol and she loved that lifestyle.	

PRED-MISMATCH	Dwight freelanced as a plumber in Bristol and she loved that lifestyle.	
REF-MATCH	Nancy shadowed a mayor in Texas but he was very unpleasant about it.	
REF-MISMATCH	Nancy shadowed a cheerleader in Texas but he was very unpleasant about it.	
PRED-MATCH	Nancy wound up as a mayor in Texas but he was very unpleasant about it.	
PRED-MISMATCH	Nancy wound up as a cheerleader in Texas but he was very unpleasant about it.	
REF-MATCH	Jacob met a caterer at the party and she was also a great dancer.	
REF-MISMATCH	Jacob met a chef at the party and she was also a great dancer.	
PRED-MATCH	Jacob felt like a caterer at the party and she was also a great dancer.	
PRED-MISMATCH	Jacob felt like a chef at the party and she was also a great dancer.	
REF-MATCH	Lily had dinner with a true magician last week and he agreed to take a selfie.	
REF-MISMATCH	Lily had dinner with a true fortune-teller last week and he agreed to take a selfie.	
PRED-MATCH	Lily felt like a true magician last week and he agreed to take a selfie.	
PRED-MISMATCH	Lily felt like a true fortune-teller last week and he agreed to take a selfie.	
REF-MATCH	Will was introduced to a kindergarten-teacher in London and she looked really nice.	
REF-MISMATCH	Will was introduced to a dentist in London and she looked really nice.	
PRED-MATCH	Will interned as a kindergarten-teacher in London and she looked really nice.	
PRED-MISMATCH	Will interned as a dentist in London and she looked really nice.	
REF-MATCH	Barbara dated an engineer in the US but he was very jealous.	Was it in the US that Barbara dated an engineer?
REF-MISMATCH	Barbara dated an obstetrician in the US but he was very jealous.	Was it in the US that Barbara dated an obstetrician?
PRED-MATCH	Barbara stayed an engineer in the US but he was very jealous.	Was it in the US that Barbara stayed an engineer?

PRED-MISMATCH	Barbara stayed an obstetrician in the US but he was very jealous.	Was it in the US that Barbara stayed an obstetrician?
REF-MATCH	Lucas married a canteen-assistant and luckily she was an easy-going person.	
REF-MISMATCH	Lucas married a chauffeur and luckily she was an easy-going person.	
PRED-MATCH	Lucas trained as a canteen-assistant and luckily she was an easy-going person.	
PRED-MISMATCH	Lucas trained as a chauffeur and luckily she was an easy-going person.	
REF-MATCH	Elaine lived with a surgeon but then he left to study French.	
REF-MISMATCH	Elaine lived with a nurse but then he left to study French.	
PRED-MATCH	Elaine trained as a surgeon but then he left to study French.	
PRED-MISMATCH	Elaine trained as a nurse but then he left to study French.	
REF-MATCH	Marcus courted a housekeeper but secretly she didn't enjoy it.	
REF-MISMATCH	Marcus courted a blacksmith but secretly she didn't enjoy it.	
PRED-MATCH	Marcus continued as a housekeeper but secretly she didn't enjoy it.	
PRED-MISMATCH	Marcus continued as a blacksmith but secretly she didn't enjoy it.	
REF-MATCH	Susan knew a politician for the Labour party and he always behaved kindly.	
REF-MISMATCH	Susan knew a secretary for the Labour party and he always behaved kindly.	
PRED-MATCH	Susan was hired as a politician for the Labour party and he always behaved kindly.	
PRED-MISMATCH	Susan was hired as a secretary for the Labour party and he always behaved kindly.	
REF-MATCH	Jonathan was friendly with a paralegal at a big firm and she was very confident.	Was it a paralegal that Jonathan was friendly with?
REF-MISMATCH	Jonathan was friendly with a lawyer at a big firm and she was very confident.	Was it a lawyer that Jonathan was friendly with?
PRED-MATCH	Jonathan excelled as a paralegal at a big firm and she was very confident.	Was it as a paralegal that Jonathan excelled?

PRED-MISMATCH	Jonathan excelled as a lawyer at a big firm and she was very confident.	Was it as a lawyer that Jonathan excelled?
REF-MATCH	Karen funded a scientist at the university but then he disappeared very suddenly.	
REF-MISMATCH	Karen funded a librarian at the university but then he disappeared very suddenly.	
PRED-MATCH	Karen excelled as a scientist at the university but then he disappeared very suddenly.	
PRED-MISMATCH	Karen excelled as a librarian at the university but then he disappeared very suddenly.	
REF-MATCH	Malcolm loved a cosmetologist for five years but then she had had enough.	
REF-MISMATCH	Malcolm loved a minister for five years but then she had had enough.	
PRED-MATCH	Malcolm remained a cosmetologist for five years but then she had had enough.	
PRED-MISMATCH	Malcolm remained a minister for five years but then she had had enough.	
REF-MATCH	Erin offended a judge on the trial but he seemed bored most of the time.	
REF-MISMATCH	Erin offended a file-clerk on the trial but he seemed bored most of the time.	
PRED-MATCH	Erin appeared as a judge on the trial but he seemed bored most of the time.	
PRED-MISMATCH	Erin appeared as a file-clerk on the trial but he seemed bored most of the time.	
REF-MATCH	Oscar spoke with a dressmaker in Manchester and then she had to leave.	
REF-MISMATCH	Oscar spoke with a technician in Manchester and then she had to leave.	
PRED-MATCH	Oscar struggled as a dressmaker in Manchester and then she had to leave.	
PRED-MISMATCH	Oscar struggled as a technician in Manchester and then she had to leave.	
REF-MATCH	Charlotte telephoned a mechanic and fortunately he could fix the problem.	
REF-MISMATCH	Charlotte telephoned a nanny and fortunately he could fix the problem.	
PRED-MATCH	Charlotte interned as a mechanic and fortunately he could fix the problem.	

PRED-MISMATCH	Charlotte interned as a nanny and fortunately he could fix the problem.	
REF-MATCH	Arthur wanted to meet with a dietitian on the firm but she had other plans.	
REF-MISMATCH	Arthur wanted to meet with a detective on the firm but she had other plans.	
PRED-MATCH	Arthur wanted to stay a dietitian on the firm but she had other plans.	
PRED-MISMATCH	Arthur wanted to stay a detective on the firm but she had other plans.	
REF-MATCH	Kathy fired a barber at the salon and he wasn't happy about that.	
REF-MISMATCH	Kathy fired a hairdresser at the salon and he wasn't happy about that.	
PRED-MATCH	Kathy stayed a barber at the salon and he wasn't happy about that.	
PRED-MISMATCH	Kathy stayed a hairdresser at the salon and he wasn't happy about that.	
REF-MATCH	Pete hung out with a cheerleader in public but she didn't appreciate it.	
REF-MISMATCH	Pete hung out with a mayor in public but she didn't appreciate it.	
PRED-MATCH	Pete behaved like a cheerleader in public but she didn't appreciate it.	
PRED-MISMATCH	Pete behaved like a mayor in public but she didn't appreciate it.	
REF-MATCH	Emily paid a chauffeur and luckily he got everywhere on time.	
REF-MISMATCH	Emily paid a canteen-assistant and luckily he got everywhere on time.	
PRED-MATCH	Emily behaved like a chauffeur and luckily he got everywhere on time.	
PRED-MISMATCH	Emily behaved like a canteen-assistant and luckily he got everywhere on time.	
REF-MATCH	Nathan noticed a cleaner but in reality she didn't work there.	
REF-MISMATCH	Nathan noticed a farmer but in reality she didn't work there.	
PRED-MATCH	Nathan seemed like a cleaner but in reality she didn't work there.	

PRED-MISMATCH	Nathan seemed like a farmer but in reality she didn't work there.	
REF-MATCH	Janet met up with a firefighter in training but he didn't know what to do.	
REF-MISMATCH	Janet met up with a manicurist in training but he didn't know what to do.	
PRED-MATCH	Janet seemed like a firefighter in training but he didn't know what to do.	
PRED-MISMATCH	Janet seemed like a manicurist in training but he didn't know what to do.	
REF-MATCH	Robert suggested a receptionist but unfortunately she was totally unqualified for the job.	Did the sentence you just read contain the word "fortunately"?
REF-MISMATCH	Robert suggested a senator but unfortunately she was totally unqualified for the job.	Did the sentence you just read contain the word "fortunately"?
PRED-MATCH	Robert seemed like a receptionist but unfortunately she was totally unqualified for the job.	Did the sentence you just read contain the word "fortunately"?
PRED-MISMATCH	Robert seemed like a senator but unfortunately she was totally unqualified for the job.	Did the sentence you just read contain the word "fortunately"?
REF-MATCH	Elizabeth asked an electrician and fortunately he could solve any problem.	
REF-MISMATCH	Elizabeth asked a beautician and fortunately he could solve any problem.	
PRED-MATCH	Elizabeth was born an electrician and fortunately he could solve any problem.	
PRED-MISMATCH	Elizabeth was born a beautician and fortunately he could solve any problem.	
REF-MATCH	Thomas hated a florist but eventually she left for beauty school.	
REF-MISMATCH	Thomas hated a butcher but eventually she left for beauty school.	
PRED-MATCH	Thomas freelanced as a florist but eventually she left for beauty school.	
PRED-MISMATCH	Thomas freelanced as a butcher but eventually she left for beauty school.	
REF-MATCH	Patricia fell in love with a pilot and not long after he went on a mission.	
REF-MISMATCH	Patricia fell in love with a flight-attendant and not long after he went on a mission.	

PRED-MATCH	Patricia was educated as a pilot and not long after he went on a mission.	
PRED-MISMATCH	Patricia was educated as a flight-attendant and not long after he went on a mission.	
REF-MATCH	Jordan complained to a stripper but then she had to leave.	
REF-MISMATCH	Jordan complained to a clown but then she had to leave.	
PRED-MATCH	Jordan continued as a stripper but then she had to leave.	
PRED-MISMATCH	Jordan continued as a clown but then she had to leave.	
REF-MATCH	Florence laughed at a clown but then he got into trouble.	Was it a pianist that Florence laughed at?
REF-MISMATCH	Florence laughed at a stripper but then he got into trouble.	Was it a pianist that Florence laughed at?
PRED-MATCH	Florence freelanced as a clown but then he got into trouble.	Was it a pianist that Florence freelanced as?
PRED-MISMATCH	Florence freelanced as a stripper but then he got into trouble.	Was it a pianist that Florence freelanced as?
REF-MATCH	Mohammed bantered with an advice-columnist and luckily she was quite successful.	Did Mohammed banter with an advice-columnist?
REF-MISMATCH	Mohammed bantered with a bartender and luckily she was quite successful.	Did Mohammed banter with a bartender?
PRED-MATCH	Mohammed wound up as an advice-columnist and luckily she was quite successful.	Did Mohammed wind up as an advice-columnist?
PRED-MISMATCH	Mohammed wound up as a bartender and luckily she was quite successful.	Did Mohammed wind up as a bartender?
REF-MATCH	Mia wrote music with a bartender and then he got very famous.	
REF-MISMATCH	Mia wrote music with an advice-columnist and then he got very famous.	
PRED-MATCH	Mia was considered a bartender and then he got very famous.	
PRED-MISMATCH	Mia was considered an advice-columnist and then he got very famous.	
REF-MATCH	Archie partied with a gymnast but then she got tired of it.	Was it Harry that partied with a gymnast?
REF-MISMATCH	Archie partied with a drummer but then she got tired of it.	Was it Harry that partied with a drummer?
PRED-MATCH	Archie was perceived as a gymnast but then she got tired of it.	Was it Harry that was perceived as a gymnast?

PRED-MISMATCH	Archie was perceived as a drummer but then she got tired of it.	Was it Harry that was perceived as a drummer?
REF-MATCH	Rosie trained with a drummer for a week and he absolutely loved it.	
REF-MISMATCH	Rosie trained with a gymnast for a week and he absolutely loved it.	
PRED-MATCH	Rosie turned into a drummer for a week and he absolutely loved it.	
PRED-MISMATCH	Rosie turned into a gymnast for a week and he absolutely loved it.	
REF-MATCH	Harry spied on a caregiver for two years but she didn't mind.	Was it Harry that spied on a caregiver?
REF-MISMATCH	Harry spied on a diplomat for two years but she didn't mind.	Was it Harry that spied on a diplomat?
PRED-MATCH	Harry volunteered as a caregiver for two years but she didn't mind.	Was it Harry that volunteered as a caregiver?
PRED-MISMATCH	Harry volunteered as a diplomat for two years but she didn't mind.	Was it Harry that volunteered as a diplomat?
REF-MATCH	Daisy sublet to a diplomat for a while and he was very poor.	
REF-MISMATCH	Daisy sublet to a caregiver for a while and he was very poor.	
PRED-MATCH	Daisy struggled as a diplomat for a while and he was very poor.	
PRED-MISMATCH	Daisy struggled as a caregiver for a while and he was very poor.	
REF-MATCH	Ronald had a crush on a beauty-consultant in college but now she has completely moved on.	
REF-MISMATCH	Ronald had a crush on a dean in college but now she has completely moved on.	
PRED-MATCH	Ronald was perceived as a beauty-consultant in college but now she has completely moved on.	
PRED-MISMATCH	Ronald was perceived as a dean in college but now she has completely moved on.	
REF-MATCH	Harper visited a dean last year and he enjoyed it a lot.	
REF-MISMATCH	Harper visited a beauty-consultant last year and he enjoyed it a lot.	

PRED-MATCH	Harper was considered a dean last year and he enjoyed it a lot.	
PRED-MISMATCH	Harper was considered a beauty-consultant last year and he enjoyed it a lot.	
REF-MATCH	Elijah apologized to a feminist and later she really appreciated it.	
REF-MISMATCH	Elijah apologized to a footballer and later she really appreciated it.	
PRED-MATCH	Elijah was born a feminist and later she really appreciated it.	
PRED-MISMATCH	Elijah was born a footballer and later she really appreciated it.	
REF-MATCH	Maya upset a footballer and luckily he didn't change.	
REF-MISMATCH	Maya upset a feminist and luckily he didn't change.	
PRED-MATCH	Maya was content as a footballer and luckily he didn't change.	
PRED-MISMATCH	Maya was content as a feminist and luckily he didn't change.	
REF-MATCH	Ted warned a groupie and of course she took it seriously.	
REF-MISMATCH	Ted warned a general and of course she took it seriously.	
PRED-MATCH	Ted was considered a groupie and of course she took it seriously.	
PRED-MISMATCH	Ted was considered a general and of course she took it seriously.	
REF-MATCH	Rose desired a general but later he disappeared without a trace.	
REF-MISMATCH	Rose desired a groupie but later he disappeared without a trace.	
PRED-MATCH	Rose appeared as a general but later he disappeared without a trace.	
PRED-MISMATCH	Rose appeared as a groupie but later he disappeared without a trace.	
REF-MATCH	Marshall cooked dinner for a stenographer and luckily she did a great job.	
REF-MISMATCH	Marshall cooked dinner for a lumberjack and luckily she did a great job.	
PRED-MATCH	Marshall was employed as a stenographer and luckily she did a great job.	

PRED-MISMATCH	Marshall was employed as a lumberjack and luckily she did a great job.	
REF-MATCH	Molly cuddled with a lumberjack and of course he really loved it.	
REF-MISMATCH	Molly cuddled with a stenographer and of course he really loved it.	
PRED-MATCH	Molly was educated as a lumberjack and of course he really loved it.	
PRED-MISMATCH	Molly was educated as a stenographer and of course he really loved it.	
REF-MATCH	Oliver stopped by a daycare-manager and fortunately she was really knowledgeable.	Was it Nancy that stopped by a daycare-manager?
REF-MISMATCH	Oliver stopped by a hunter and fortunately she was really knowledgeable.	Was it Nancy that stopped by a hunter?
PRED-MATCH	Oliver was content as a daycare-manager and fortunately she was really knowledgeable.	Was it Nancy that was content as a daycare-manager?
PRED-MISMATCH	Oliver was content as a hunter and fortunately she was really knowledgeable.	Was it Nancy that was content as a hunter?
REF-MATCH	Clara sang for a hunter but later he had to leave.	
REF-MISMATCH	Clara sang for a daycare-manager but later he had to leave.	
PRED-MATCH	Clara volunteered as a hunter but later he had to leave.	
PRED-MISMATCH	Clara volunteered as a daycare-manager but later he had to leave.	
REF-MATCH	Theodore chased an interior-decorator for two weeks but she could eventually get away.	Was it an interior-decorator that Theodore chased?
REF-MISMATCH	Theodore chased an explorer for two weeks but she could eventually get away.	Was it an explorer that Theodore chased?
PRED-MATCH	Theodore struggled as an interior-decorator for two weeks but she could eventually get away.	Was it as an interior-decorator that Theodore struggled?
PRED-MISMATCH	Theodore struggled as an explorer for two weeks but she could eventually get away.	Was it as an explorer that Theodore struggled?
REF-MATCH	Lizzie considered an explorer for the job but then he did something else.	
REF-MISMATCH	Lizzie considered an interior-decorator for the job but then he did something else.	

PRED-MATCH	Lizzie was educated as an explorer for the job but then he did something else.	
PRED-MISMATCH	Lizzie was educated as an interior-decorator for the job but then he did something else.	
REF-MATCH	William waved to a governess and luckily she was very friendly.	
REF-MISMATCH	William waved to a bachelor and luckily she was very friendly.	
PRED-MATCH	William behaved like a governess and luckily she was very friendly.	
PRED-MISMATCH	William behaved like a bachelor and luckily she was very friendly.	
REF-MATCH	Jennifer charmed a bachelor and afterwards he went out clubbing.	Was it Jennifer that charmed a bachelor?
REF-MISMATCH	Jennifer charmed a governess and afterwards he went out clubbing.	Was it Jennifer that charmed a governess?
PRED-MATCH	Jennifer felt like a bachelor and afterwards he went out clubbing.	Was it Jennifer that felt like a bachelor?
PRED-MISMATCH	Jennifer felt like a governess and afterwards he went out clubbing.	Was it Jennifer that felt like a governess?
REF-MATCH	Richard was related to a duchess in York but then she disappeared without a trace.	
REF-MISMATCH	Richard was related to a duke in York but then she disappeared without a trace.	
PRED-MATCH	Richard was born a duchess in York but then she disappeared without a trace.	
PRED-MISMATCH	Richard was born a duke in York but then she disappeared without a trace.	
REF-MATCH	Sarah had drinks with a duke in Norfolk and he enjoyed it immensely.	
REF-MISMATCH	Sarah had drinks with a duchess in Norfolk and he enjoyed it immensely.	
PRED-MATCH	Sarah remained a duke in Norfolk and he enjoyed it immensely.	
PRED-MISMATCH	Sarah remained a duchess in Norfolk and he enjoyed it immensely.	
REF-MATCH	Charles frightened a girl and of course she found that upsetting.	Did the sentence you just read contain the word "upset"?
REF-MISMATCH	Charles frightened a boy and of course she found that upsetting.	Did the sentence you just read contain the word "upset"?
PRED-MATCH	Charles was educated as a girl and of course she found that upsetting.	Did the sentence you just read contain the word "upset"?

PRED-MISMATCH	Charles was educated as a boy and of course she found that upsetting.	Did the sentence you just read contain the word "upset"?
REF-MATCH	Karen encouraged a boy but still he felt completely inadequate.	
REF-MISMATCH	Karen encouraged a girl but still he felt completely inadequate.	
PRED-MATCH	Karen looked like a boy but still he felt completely inadequate.	
PRED-MISMATCH	Karen looked like a girl but still he felt completely inadequate.	
REF-MATCH	Christopher embraced a queen and obviously she just went on and on about it.	
REF-MISMATCH	Christopher embraced a king and obviously she just went on and on about it.	
PRED-MATCH	Christopher seemed like a queen and obviously she just went on and on about it.	
PRED-MISMATCH	Christopher seemed like a king and obviously she just went on and on about it.	
REF-MATCH	Betty owed money to a king in Africa but he never spoke of it.	
REF-MISMATCH	Betty owed money to a queen in Africa but he never spoke of it.	
PRED-MATCH	Betty was content as a king in Africa but he never spoke of it.	
PRED-MISMATCH	Betty was content as a queen in Africa but he never spoke of it.	
REF-MATCH	Anthony dined with a waitress but then she had to go back to work.	
REF-MISMATCH	Anthony dined with a waiter but then she had to go back to work.	
PRED-MATCH	Anthony was perceived as a waitress but then she had to go back to work.	
PRED-MISMATCH	Anthony was perceived as a waiter but then she had to go back to work.	
REF-MATCH	Margaret congratulated a waiter and naturally he felt quite lucky.	
REF-MISMATCH	Margaret congratulated a waitress and naturally he felt quite lucky.	
PRED-MATCH	Margaret trained as a waiter and naturally he felt quite lucky.	

PRED-MISMATCH	Margaret trained as a waitress and naturally he felt quite lucky.	
REF-MATCH	Donald crept on a woman and unfortunately she took offense by that.	
REF-MISMATCH	Donald crept on a man and unfortunately she took offense by that.	
PRED-MATCH	Donald was considered a woman and unfortunately she took offense by that.	
PRED-MISMATCH	Donald was considered a man and unfortunately she took offense by that.	
REF-MATCH	Ashley admired a man on TV but unfortunately he quit last year.	
REF-MISMATCH	Ashley admired a woman on TV but unfortunately he quit last year.	
PRED-MATCH	Ashley excelled as a man on TV but unfortunately he quit last year.	
PRED-MISMATCH	Ashley excelled as a woman on TV but unfortunately he quit last year.	
REF-MATCH	Steven dressed a bride for the festivities and she was very happy.	
REF-MISMATCH	Steven dressed a groom for the festivities and she was very happy.	
PRED-MATCH	Steven looked like a bride for the festivities and she was very happy.	
PRED-MISMATCH	Steven looked like a groom for the festivities and she was very happy.	
REF-MATCH	Donna ran into a groom at the wedding but he seemed very upset.	
REF-MISMATCH	Donna ran into a bride at the wedding but he seemed very upset.	
PRED-MATCH	Donna was content as a groom at the wedding but he seemed very upset.	
PRED-MISMATCH	Donna was content as a bride at the wedding but he seemed very upset.	
REF-MATCH	Joshua eavesdropped on a saleswoman but then she got up and left.	
REF-MISMATCH	Joshua eavesdropped on a salesman but then she got up and left.	
PRED-MATCH	Joshua was hired as a saleswoman but then she got up and left.	

PRED-MISMATCH	Joshua was hired as a salesman but then she got up and left.	
REF-MATCH	Michelle competed with a salesman at work and he was very skilled.	
REF-MISMATCH	Michelle competed with a saleswoman at work and he was very skilled.	
PRED-MATCH	Michelle continued as a salesman at work and he was very skilled.	
PRED-MISMATCH	Michelle continued as a saleswoman at work and he was very skilled.	
REF-MATCH	Josh invited a baroness from Spain but she was quite an unpleasant person.	
REF-MISMATCH	Josh invited a baron from Spain but she was quite an unpleasant person.	
PRED-MATCH	Josh was perceived as a baroness from Spain but she was quite an unpleasant person.	
PRED-MISMATCH	Josh was perceived as a baron from Spain but she was quite an unpleasant person.	
REF-MATCH	Carol avoided a baron at the party but he didn't understand why.	
REF-MISMATCH	Carol avoided a baroness at the party but he didn't understand why.	
PRED-MATCH	Carol felt like a baron at the party but he didn't understand why.	
PRED-MISMATCH	Carol felt like a baroness at the party but he didn't understand why.	
REF-MATCH	Tim chatted with a hostess at the event and she was very organized.	Was it at the event that Tim chatted with a hostess?
REF-MISMATCH	Tim chatted with a host at the event and she was very organized.	Was it at the event that Tim chatted with a host?
PRED-MATCH	Tim appeared as a hostess at the event and she was very organized.	Was it at the event that Tim appeared as a hostess?
PRED-MISMATCH	Tim appeared as a host at the event and she was very organized.	Was it at the event that Tim appeared as a host?
REF-MATCH	Amanda drank with a host at the party but he regretted it afterwards.	
REF-MISMATCH	Amanda drank with a hostess at the party but he regretted it afterwards.	
PRED-MATCH	Amanda was hired as a host at the party but he regretted it afterwards.	

PRED-MISMATCH	Amanda was hired as a hostess at the party but he regretted it afterwards.	
REF-MATCH	Jason greeted a countess but unfortunately she was indifferent about it.	
REF-MISMATCH	Jason greeted a count but unfortunately she was indifferent about it.	
PRED-MATCH	Jason wound up as a countess but unfortunately she was indifferent about it.	
PRED-MISMATCH	Jason wound up as a count but unfortunately she was indifferent about it.	
REF-MATCH	Stephanie was married to a count and thankfully he was quite charming.	
REF-MISMATCH	Stephanie was married to a countess and thankfully he was quite charming.	
PRED-MATCH	Stephanie was born a count and thankfully he was quite charming.	
PRED-MISMATCH	Stephanie was born a countess and thankfully he was quite charming.	
REF-MATCH	Jeffrey had a wife but eventually she had to leave.	
REF-MISMATCH	Jeffrey had a husband but eventually she had to leave.	
PRED-MATCH	Jeffrey persevered as a wife but eventually she had to leave.	
PRED-MISMATCH	Jeffrey persevered as a husband but eventually she had to leave.	
REF-MATCH	Rebecca longed for a husband but unfortunately he was nowhere to be found.	
REF-MISMATCH	Rebecca longed for a wife but unfortunately he was nowhere to be found.	
PRED-MATCH	Rebecca was a husband but unfortunately he was nowhere to be found.	
PRED-MISMATCH	Rebecca was a wife but unfortunately he was nowhere to be found.	
REF-MATCH	Ryan danced with a cowgirl and apparently she had a blast.	Did Ryan dance with a cowgirl?
REF-MISMATCH	Ryan danced with a cowboy and apparently she had a blast.	Did Ryan dance with a cowboy?
PRED-MATCH	Ryan was educated as a cowgirl and apparently she had a blast.	Was Ryan educated as a cowgirl?

PRED-MISMATCH	Ryan was educated as a cowboy and apparently she had a blast.	Was Ryan educated as a cowboy?
REF-MATCH	Beatrice pined for a cowboy for two years but he couldn't be trusted.	
REF-MISMATCH	Beatrice pined for a cowgirl for two years but he couldn't be trusted.	
PRED-MATCH	Beatrice stayed a cowboy for two years but he couldn't be trusted.	
PRED-MISMATCH	Beatrice stayed a cowgirl for two years but he couldn't be trusted.	
REF-MATCH	Gary negotiated with a policewoman for a while but she didn't back down.	
REF-MISMATCH	Gary negotiated with a policeman for a while but she didn't back down.	
PRED-MATCH	Gary struggled as a policewoman for a while but she didn't back down.	
PRED-MISMATCH	Gary struggled as a policeman for a while but she didn't back down.	
REF-MATCH	Amy called a policeman and fortunately he was very helpful.	Did Amy call a policeman?
REF-MISMATCH	Amy called a policewoman and fortunately he was very helpful.	Did Amy call a policewoman?
PRED-MATCH	Amy was employed as a policeman and fortunately he was very helpful.	Was Amy employed as a policeman?
PRED-MISMATCH	Amy was employed as a policewoman and fortunately he was very helpful.	Was Amy employed as a policewoman?
REF-MATCH	Eric cared for a nun for two years but then she changed quite radically.	Was it Eric that cared for a nun?
REF-MISMATCH	Eric cared for a monk for two years but then she changed quite radically.	Was it Eric that cared for a monk?
PRED-MATCH	Eric remained a nun for two years but then she changed quite radically.	Was it Eric that remained a monk?
PRED-MISMATCH	Eric remained a monk for two years but then she changed quite radically.	Was it Eric that remained a nun?
REF-MATCH	Nicole bothered a monk for three weeks but then he left very suddenly.	Did the sentence you just read contain the word "years"?
REF-MISMATCH	Nicole bothered a nun for three weeks but then he left very suddenly.	Did the sentence you just read contain the word "years"?
PRED-MATCH	Nicole continued as a monk for three weeks but then he left very suddenly.	Did the sentence you just read contain the word "years"?

PRED-MISMATCH	Nicole continued as a nun for three weeks but then he left very suddenly.	Did the sentence you just read contain the word "years"?
REF-MATCH	Scott dazzled an actress and of course she absolutely loved it.	
REF-MISMATCH	Scott dazzled an actor and of course she absolutely loved it.	
PRED-MATCH	Scott freelanced as an actress and of course she absolutely loved it.	
PRED-MISMATCH	Scott freelanced as an actor and of course she absolutely loved it.	
REF-MATCH	Helen encountered an actor in Hollywood and he was absolutely lovely.	Was it Alicia that encountered an actor?
REF-MISMATCH	Helen encountered an actress in Hollywood and he was absolutely lovely.	Was it Alicia that encountered an actress?
PRED-MATCH	Helen stayed an actor in Hollywood and he was absolutely lovely.	Was it Alicia that stayed an actor?
PRED-MISMATCH	Helen stayed an actress in Hollywood and he was absolutely lovely.	Was it Alicia that stayed an actress?
REF-MATCH	Brandon entertained a princess at the castle and she was a lovely person.	Was it at the ball that Brandon entertained a princess?
REF-MISMATCH	Brandon entertained a prince at the castle and she was a lovely person.	Was it at the ball that Brandon entertained a prince?
PRED-MATCH	Brandon was considered a princess at the castle and she was a lovely person.	Was it at the ball that Brandon was considered a princess?
PRED-MISMATCH	Brandon was considered a prince at the castle and she was a lovely person.	Was it at the ball that Brandon was considered a prince?
REF-MATCH	Samantha had eyes for a prince from Spain but he was not that nice.	
REF-MISMATCH	Samantha had eyes for a princess from Spain but he was not that nice.	
PRED-MATCH	Samantha seemed like a prince from Spain but he was not that nice.	
PRED-MISMATCH	Samantha seemed like a princess from Spain but he was not that nice.	
REF-MATCH	Samuel disagreed with a matriarch and luckily she handled that well.	Did the sentence you just read contain the word "unfortunately"?
REF-MISMATCH	Samuel disagreed with a patriarch and luckily she handled that well.	Did the sentence you just read contain the word "unfortunately"?

PRED-MATCH	Samuel turned into a matriarch and luckily she handled that well.	Did the sentence you just read contain the word "unfortunately"?
PRED-MISMATCH	Samuel turned into a patriarch and luckily she handled that well.	Did the sentence you just read contain the word "unfortunately"?
REF-MATCH	Katherine badmouthed a patriarch but luckily he didn't care.	Was it Josephine that badmouthed a patriarch?
REF-MISMATCH	Katherine badmouthed a matriarch but luckily he didn't care.	Was it Josephine that badmouthed a matriarch?
PRED-MATCH	Katherine was perceived as a patriarch but luckily he didn't care.	Was it Josephine that was perceived as a patriarch?
PRED-MISMATCH	Katherine was perceived as a matriarch but luckily he didn't care.	Was it Josephine that was perceived as a matriarch?
REF-MATCH	Greg insulted a widow back in the day but she was chill about it.	Did the sentence you just read contain the word "chill"?
REF-MISMATCH	Greg insulted a pope back in the day but she was chill about it.	Did the sentence you just read contain the word "chill"?
PRED-MATCH	Greg became a widow back in the day but she was chill about it.	Did the sentence you just read contain the word "chill"?
PRED-MISMATCH	Greg became a pope back in the day but she was chill about it.	Did the sentence you just read contain the word "chill"?
REF-MATCH	Christina invited a pope but unfortunately he was quite reserved.	
REF-MISMATCH	Christina invited a widow but unfortunately he was quite reserved.	
PRED-MATCH	Christina continued as a pope but unfortunately he was quite reserved.	
PRED-MISMATCH	Christina continued as a widow but unfortunately he was quite reserved.	
REF-MATCH	Alexander comforted a mother and naturally she felt very grateful.	
REF-MISMATCH	Alexander comforted a father and naturally she felt very grateful.	
PRED-MATCH	Alexander became a mother and naturally she felt very grateful.	
PRED-MISMATCH	Alexander became a father and naturally she felt very grateful.	
REF-MATCH	Rachel fought with a father for two years but then he completely broke down.	
REF-MISMATCH	Rachel fought with a mother for two years but then he completely broke down.	

PRED-MATCH	Rachel was content as a father for two years but then he completely broke down.	
PRED-MISMATCH	Rachel was content as a mother for two years but then he completely broke down.	
REF-MATCH	Patrick fooled a headmistress and of course she was very surprised.	
REF-MISMATCH	Patrick fooled a headmaster and of course she was very surprised.	
PRED-MATCH	Patrick was employed as a headmistress and of course she was very surprised.	
PRED-MISMATCH	Patrick was employed as a headmaster and of course she was very surprised.	
REF-MATCH	Heather gossiped about a headmaster but then he must have had enough.	Did the sentence you just read contain the word "gossiped"?
REF-MISMATCH	Heather gossiped about a headmistress but then he must have had enough.	Did the sentence you just read contain the word "gossiped"?
PRED-MATCH	Heather persevered as a headmaster but then he must have had enough.	Did the sentence you just read contain the word "persevered"?
PRED-MISMATCH	Heather persevered as a headmistress but then he must have had enough.	Did the sentence you just read contain the word "persevered"?
REF-MATCH	Jack insulted a spinster but despite all odds she did not care.	
REF-MISMATCH	Jack insulted a deacon but despite all odds she did not care.	
PRED-MATCH	Jack remained a spinster but despite all odds she did not care.	
PRED-MISMATCH	Jack remained a deacon but despite all odds she did not care.	
REF-MATCH	Diana prayed with a deacon and thankfully he wasn't weird about it.	
REF-MISMATCH	Diana prayed with a spinster and thankfully he wasn't weird about it.	
PRED-MATCH	Diana wound up as a deacon and thankfully he wasn't weird about it.	
PRED-MISMATCH	Diana wound up as a spinster and thankfully he wasn't weird about it.	
REF-MATCH	Dennis disliked a choirgirl and sadly she cried a lot about it.	

REF-MISMATCH	Dennis disliked a mailman and sadly she cried a lot about it.	
PRED-MATCH	Dennis interned as a choirgirl and sadly she cried a lot about it.	
PRED-MISMATCH	Dennis interned as a mailman and sadly she cried a lot about it.	
REF-MATCH	Ruth impressed a mailman for a while and he was so proud.	
REF-MISMATCH	Ruth impressed a choirgirl for a while and he was so proud.	
PRED-MATCH	Ruth was hired as a mailman for a while and he was so proud.	
PRED-MISMATCH	Ruth was hired as a choirgirl for a while and he was so proud.	
REF-MATCH	Aaron hurt a stewardess quite badly and she was rather upset.	
REF-MISMATCH	Aaron hurt a repairman quite badly and she was rather upset.	
PRED-MATCH	Aaron struggled as a stewardess quite badly and she was rather upset.	
PRED-MISMATCH	Aaron struggled as a repairman quite badly and she was rather upset.	
REF-MATCH	Olivia laughed with a repairman but then he had to leave.	
REF-MISMATCH	Olivia laughed with a stewardess but then he had to leave.	
PRED-MATCH	Olivia worked as a repairman but then he had to leave.	
PRED-MISMATCH	Olivia worked as a stewardess but then he had to leave.	
REF-MATCH	Zach employed a maid and luckily she was quite meticulous.	Did the sentence you just read contain the word "meticulous"?
REF-MISMATCH	Zach employed a butler and luckily she was quite meticulous.	Did the sentence you just read contain the word "meticulous"?
PRED-MATCH	Zach trained as a maid and luckily she was quite meticulous.	Did the sentence you just read contain the word "meticulous"?
PRED-MISMATCH	Zach trained as a butler and luckily she was quite meticulous.	Did the sentence you just read contain the word "meticulous"?

REF-MATCH	Kelly observed a butler for two weeks and he was very helpful.	Was it Pippa that observed a butler?
REF-MISMATCH	Kelly observed a maid for two weeks and he was very helpful.	Was it Pippa that observed a maid?
PRED-MATCH	Kelly interned as a butler for two weeks and he was very helpful.	Was it Pippa that interned as a butler?
PRED-MISMATCH	Kelly interned as a maid for two weeks and he was very helpful.	Was it Pippa that interned as a maid?
REF-MATCH	Ethan discovered a housewife in Miami but she was an awful person.	
REF-MISMATCH	Ethan discovered a gentleman in Miami but she was an awful person.	
PRED-MATCH	Ethan was considered a housewife in Miami but she was an awful person.	
PRED-MISMATCH	Ethan was considered a gentleman in Miami but she was an awful person.	
REF-MATCH	Judith amazed a gentleman back home but he never said anything.	Was it Clara that amazed a gentleman?
REF-MISMATCH	Judith amazed a housewife back home but he never said anything.	Was it Clara that amazed a housewife?
PRED-MATCH	Judith acted like a gentleman back home but he never said anything.	Was it Clara that acted like a gentleman?
PRED-MISMATCH	Judith acted like a housewife back home but he never said anything.	Was it Clara that acted like a housewife?
REF-MATCH	Keith imitated a landlady and then she filed a lawsuit.	
REF-MISMATCH	Keith imitated a landlord and then she filed a lawsuit.	
PRED-MATCH	Keith persevered as a landlady and then she filed a lawsuit.	
PRED-MISMATCH	Keith persevered as a landlord and then she filed a lawsuit.	
REF-MATCH	Megan argued with a landlord for two days but he eventually left the job.	
REF-MISMATCH	Megan argued with a landlady for two days but he eventually left the job.	
PRED-MATCH	Megan interned as a landlord for two days but he eventually left the job.	
PRED-MISMATCH	Megan interned as a landlady for two days but he eventually left the job.	
REF-MATCH	Kirk arm-wrestled with a girl-scout and luckily she had great fun.	

REF-MISMATCH	Kirk arm-wrestled with a boy-scout and luckily she had great fun.	
PRED-MATCH	Kirk turned into a girl-scout and luckily she had great fun.	
PRED-MISMATCH	Kirk turned into a boy-scout and luckily she had great fun.	
REF-MATCH	Hannah travelled with a boy-scout and of course he was always prepared.	
REF-MISMATCH	Hannah travelled with a girl-scout and of course he was always prepared.	
PRED-MATCH	Hannah excelled as a boy-scout and of course he was always prepared.	
PRED-MISMATCH	Hannah excelled as a girl-scout and of course he was always prepared.	
REF-MATCH	Carl ignored a chairwoman for many years but eventually she couldn't be bothered.	
REF-MISMATCH	Carl ignored a chairman for many years but eventually she couldn't be bothered.	
PRED-MATCH	Carl appeared as a chairwoman for many years but eventually she couldn't be bothered.	
PRED-MISMATCH	Carl appeared as a chairman for many years but eventually she couldn't be bothered.	
REF-MATCH	Martha desired a chairman for a while and he was quite dashing.	
REF-MISMATCH	Martha desired a chairwoman for a while and he was quite dashing.	
PRED-MATCH	Martha volunteered as a chairman for a while and he was quite dashing.	
PRED-MISMATCH	Martha volunteered as a chairwoman for a while and he was quite dashing.	
REF-MATCH	Harold commuted with a milkmaid and apparently she was very chatty.	
REF-MISMATCH	Harold commuted with a milkman and apparently she was very chatty.	
PRED-MATCH	Harold volunteered as a milkmaid and apparently she was very chatty.	
PRED-MISMATCH	Harold volunteered as a milkman and apparently she was very chatty.	
REF-MATCH	Teresa liked a milkman for a while but he was not very nice.	

REF-MISMATCH	Teresa liked a milkmaid for a while but he was not very nice.	
PRED-MATCH	Teresa continued as a milkman for a while but he was not very nice.	
PRED-MISMATCH	Teresa continued as a milkmaid for a while but he was not very nice.	
REF-MATCH	Sean prayed for a priestess in church and she enjoyed it a lot.	
REF-MISMATCH	Sean prayed for a priest in church and she enjoyed it a lot.	
PRED-MATCH	Sean was employed as priestess in church and she enjoyed it a lot.	
PRED-MISMATCH	Sean was employed as a priest in church and she enjoyed it a lot.	
REF-MATCH	Sophia joined a priest for a while but eventually he got too bored.	
REF-MISMATCH	Sophia joined a priestess for a while but eventually he got too bored.	
PRED-MATCH	Sophia stayed a priest for a while but eventually he got too bored.	
PRED-MISMATCH	Sophia stayed a priestess for a while but eventually he got too bored.	
REF-MATCH	Dylan proposed to a lady but unfortunately she struggled with depression.	
REF-MISMATCH	Dylan proposed to a lord but unfortunately she struggled with depression.	
PRED-MATCH	Dylan was a lady but unfortunately she struggled with depression.	
PRED-MISMATCH	Dylan was a lord but unfortunately she struggled with depression.	
REF-MATCH	Julia had lunch with a lord and fortunately he also had good manners.	
REF-MISMATCH	Julia had lunch with a lady and fortunately he also had good manners.	
PRED-MATCH	Julia looked like a lord and fortunately he also had good manners.	
PRED-MISMATCH	Julia looked like a lady and fortunately he also had good manners.	
REF-MATCH	Joe reassured a stepmother and eventually she felt much better.	
REF-MISMATCH	Joe reassured a stepfather and eventually she felt much better.	

PRED-MATCH	Joe persevered as a stepmother and eventually she felt much better.	
PRED-MISMATCH	Joe persevered as a stepfather and eventually she felt much better.	
REF-MATCH	Alice supported a stepfather and luckily he was absolutely thrilled.	
REF-MISMATCH	Alice supported a stepmother and luckily he was absolutely thrilled.	
PRED-MATCH	Alice became a stepfather and luckily he was absolutely thrilled.	
PRED-MISMATCH	Alice became a stepmother and luckily he was absolutely thrilled.	
REF-MATCH	Albert taught a midwife for two years but she dreamt of bigger things.	
REF-MISMATCH	Albert taught a businessman for two years but she dreamt of bigger things.	
PRED-MATCH	Albert freelanced as a midwife for two years but she dreamt of bigger things.	
PRED-MISMATCH	Albert freelanced as a businessman for two years but she dreamt of bigger things.	
REF-MATCH	Isabella arrested a businessman and then he became very successful.	Was it Pauline that arrested a businessman?
REF-MISMATCH	Isabella arrested a midwife and then he became very successful.	Was it Pauline that arrested a midwife?
PRED-MATCH	Isabella was educated as a businessman and then he became very successful.	Was it Pauline that was educated as a businessman?
PRED-MISMATCH	Isabella was educated as a midwife and then he became very successful.	Was it Pauline that was educated as a midwife?
REF-MATCH	Juan quizzed a congresswoman but later she lost all confidence.	Did the sentence you just read contain the word "senator"?
REF-MISMATCH	Juan quizzed a congressman but later she lost all confidence.	Did the sentence you just read contain the word "senator"?
PRED-MATCH	Juan excelled as a congresswoman but later she lost all confidence.	Did the sentence you just read contain the word "senator"?
PRED-MISMATCH	Juan excelled as a congressman but later she lost all confidence.	Did the sentence you just read contain the word "senator"?
REF-MATCH	Natalie intimidated a congressman but at least he was very confident.	
REF-MISMATCH	Natalie intimidated a congresswoman but at least he was very confident.	
PRED-MATCH	Natalie struggled as a congressman but at least he was very confident.	

PRED-MISMATCH	Natalie struggled as a congresswoman but at least he was very confident.	
REF-MATCH	Alan approached a noblewoman but it was obvious that she was rather insecure.	
REF-MISMATCH	Alan approached a nobleman but it was obvious that she was rather insecure.	
PRED-MATCH	Alan acted like a noblewoman but it was obvious that she was rather insecure.	
PRED-MISMATCH	Alan acted like a nobleman but it was obvious that she was rather insecure.	
REF-MATCH	Brittany sweet-talked a nobleman yesterday and he felt really good.	
REF-MISMATCH	Brittany sweet-talked a noblewoman yesterday and he felt really good.	
PRED-MATCH	Brittany behaved like a nobleman yesterday and he felt really good.	
PRED-MISMATCH	Brittany behaved like a noblewoman yesterday and he felt really good.	
REF-MATCH	Philip performed for a ballerina today and she loved every second.	
REF-MISMATCH	Philip performed for a bellboy today and she loved every second.	
PRED-MATCH	Philip worked as a ballerina today and she loved every second.	
PRED-MISMATCH	Philip worked as a bellboy today and she loved every second.	
REF-MATCH	Marie bullied a bellboy for two years but he still remained self-confident.	Did Marie bully a bellboy for two years?
REF-MISMATCH	Marie bullied a ballerina for two years but he still remained self-confident.	Did Marie bully a ballerina for two years?
PRED-MATCH	Marie was a bellboy for two years but he still remained self-confident.	Was Marie a bellboy for two years?
PRED-MISMATCH	Marie was a ballerina for two years but he still remained self-confident.	Was Marie a ballerina for two years?
REF-MATCH	Vincent listened to a newswoman and of course she had great things to say.	
REF-MISMATCH	Vincent listened to a newsman and of course she had great things to say.	
PRED-MATCH	Vincent freelanced as a newswoman and of course she had great things to say.	

PRED-MISMATCH	Vincent freelanced as a newsman and of course she had great things to say.	
REF-MATCH	Gabriella tricked a newsman but somehow he actually enjoyed it.	
REF-MISMATCH	Gabriella tricked a newswoman but somehow he actually enjoyed it.	
PRED-MATCH	Gabriella turned into a newsman but somehow he actually enjoyed it.	
PRED-MISMATCH	Gabriella turned into a newswoman but somehow he actually enjoyed it.	
REF-MATCH	Randy nominated a camerawoman but then she was criticized for it.	Did Randy interview a camerawoman?
REF-MISMATCH	Randy nominated a cameraman but then she was criticized for it.	Did Randy interview a cameraman?
PRED-MATCH	Randy volunteered as a camerawoman but then she was criticized for it.	Did Randy interview a camerawoman?
PRED-MISMATCH	Randy volunteered as a cameraman but then she was criticized for it.	Did Randy interview a cameraman?
REF-MATCH	Danielle opposed a cameraman but unfortunately he was a quite timid person.	
REF-MISMATCH	Danielle opposed a camerawoman but unfortunately he was a quite timid person.	
PRED-MATCH	Danielle appeared as a cameraman but unfortunately he was a quite timid person.	
PRED-MISMATCH	Danielle appeared as a camerawoman but unfortunately he was a quite timid person.	
REF-MATCH	Eugene had tea with a clergywoman and of course she was quite charming.	Was it tea that Eugene had with a clergywoman?
REF-MISMATCH	Eugene had tea with a clergyman and of course she was quite charming.	Was it tea that Eugene had with a clergyman?
PRED-MATCH	Eugene worked as a clergywoman and of course she was quite charming.	Was it a clergywoman that Eugene worked as?
PRED-MISMATCH	Eugene worked as a clergyman and of course she was quite charming.	Was it a clergyman that Eugene worked as?
REF-MATCH	Laura played soccer with a clergyman but unfortunately he didn't enjoy it.	
REF-MISMATCH	Laura played soccer with a clergywoman but unfortunately he didn't enjoy it.	
PRED-MATCH	Laura was educated as a clergyman but unfortunately he didn't enjoy it.	

PRED-MISMATCH	Laura was educated as a clergywoman but unfortunately he didn't enjoy it.	