# Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives

## Elsi  Kaiser

Department of Linguistics, University of Southern California, Los Angeles, CA, USA

We report three experiments on reference resolution in Dutch. The results of two off-line experiments and an eye-tracking study suggest that the interpretation of different referential forms—in particular, "emphatic" strong pronouns, weak pronouns, and demonstrative pronouns—cannot be satisfactorily explained in terms of a single feature of the antecedent. The findings show that while the different preferences of demonstratives pronouns and nonemphatic personal pronouns correlate with the antecedent's grammatical role, the distinction between strong/emphatic personal pronouns and weak personal pronouns cannot be satisfactorily explained by grammatical role. The results suggest that the strong form is sensitive to the presence of contrastive, switched topics. These findings indicate that the form-specific multiple-constraints approach (e.g., Kaiser & Trueswell, 2008) can be extended to the strong/weak distinction and contrast sensitivity. We also discuss the implications of these results for the nature of the form-function mapping in anaphoric paradigms from a Gricean perspective.

*Keywords:* Reference resolution; Visual-world eye tracking; Dutch.

---

---

## INTRODUCTION

Reference resolution is a requirement for successful comprehension and a task that we all perform extremely frequently and rapidly. One of the core questions facing theories of reference resolution is how a comprehender faced with a "semantically impoverished" referring expression such as a pronoun locates the intended antecedent in the discourse model. Already-mentioned entities are presumably represented in the participants' mental-discourse model and, at a particular point in the discourse, are associated with particular properties, such as being highly salient, being newly introduced, being the agent or patient of an action, and/or contrasting with another referent in the discourse model, and so on. These properties are not mutually exclusive: for example, a discourse entity can be both highly salient and contrastive. During the past few decades, a considerable amount of linguistic and psycholinguistic research has focused in particular on questions related to salience, i.e., the question of what referential forms are used for antecedents of different salience levels (i.e., referents that are more or less prominent in comprehenders' mental models of the discourse). A large body of work indicates that informationally impoverished referential forms (e.g., pronouns) are used for highly salient entities, whereas fuller forms (e.g., noun phrases – NP) are used for less salient referents.

However, for some referential forms—such as *strong or emphatic pronouns*—the relevant linguistic and cognitive factors guiding interpretation are not yet well understood. Broadly speaking, research on emphatic referential forms cross-linguistically has led to the observation that in some languages, particular referential forms are used when some kind of special emphasis is involved (e.g., Rigau, 1986, on overt pronouns in Spanish). It seems that *strong* forms are used for entities, which are construed as being emphasised or as having some kind of special status in the discourse, whereas *weak* forms are used for referents that are not being emphasised. However, questions remain about the precise nature of this emphasis or "special status". Existing research—based mainly on corpus data and native-speaker intuitions—has led to conflicting claims regarding the question of whether interpretation of strong pronouns is sensitive to *antecedent salience*—as has been suggested for "regular" pronouns and demonstrative pronouns—or whether the interpretation of emphatic/strong pronouns is also sensitive to the presence of another factor, namely *contrast*.

This issue has implications for our understanding of the nature of the form-function mapping in anaphoric paradigms; in particular, the question of what kinds of representations guide the process of pronoun interpretation. Existing research suggests that the notion of salience plays an important role, but there are also indications that a unified concept of salience may not be sufficient to explain all aspects of reference resolution (see e.g., Brown-Schmidt, Byron, &

Tanenhaus, 2005; Kaiser, 2003; Kaiser & Trueswell, 2008; see also Kehler, 2002; Kehler, Kertz, Rohde, & Elman, 2008, for work on the effects of semantic coherence relations).

The three experiments in this paper aim to shed light on these issues. They investigate the processing of so-called *emphatic or strong pronouns* in Dutch, as compared to "regular" personal pronouns as well as demonstrative pronouns, which in Dutch can be used anaphorically to refer to human antecedents. Due to the nature of its anaphoric paradigm, Dutch is well suited for investigating the different claims regarding strong and weak anaphoric forms. Broadly speaking, this research contributes to our understanding of the extent to which reference resolution can(not) be regarded as a process guided by referent salience, by looking at how contrast contributes to this process and how it interacts with salience. Additionally, this research furthers our understanding of a typologically well established but psycholinguistically under-researched linguistic phenomenon.

## Background: Pronouns in Dutch

In Dutch, both masculine and feminine personal pronouns have two versions, sometimes referred to as strong and weak forms. Both the weak and the strong feminine forms of "she"—the weak form *ze* and the strong form *zij*—are used in colloquial and Standard Dutch. The feminine singular forms are in fact homophonous with the plural pronoun ("they"); this paper focuses on the usage of these forms as singular personal pronouns. The masculine form *hij* "he" is used in both Standard and colloquial Dutch. There is also a weak masculine form *ie* "he", but it is a clitic restricted to the spoken language and cannot occur sentence-initially in subject position, in contrast to the other forms[1] (Examples (1a) and (1b); Donaldson, 1997, p. 56; see also Cardinaletti & Starke, 1996, p. 43). Thus, *hij* "he" is not a reduced variant of another form, but it is the "smallest" referential form that can be used to refer to a masculine referent in sentence-initial subject position, relative to other options such as anaphoric demonstratives (discussed below), definite and indefinite full nouns, and so on.

(1a) Hij/*ie gelooft er niets van. (Haeseryn et al., 1997, p. 253)

He believes there none of.

---

[1] Although both *ze* and *zij* occur in subject position, in some other syntactic configurations only the strong form is possible. For Example, *ze* cannot occur in coordinated structures, unlike *zij* and *hij* (example adapted from Cardinaletti & Starke, 1996, p. 43; see also Haeseryn et al., 1997, p. 252). Thus, masculine *hij* resembles *zij* in its syntactic behaviour.

    (a) {Zij/*ze} en Jan willen niet komen.    (b) Hij en Jan willen niet komen.

    "She and Jan do not want to come".    "He and Jan do not want to come".

"He doesn't believe any of it".

(1b) Zij/ze gelooft er niets van. (Haeseryn et al., 1997, p. 253)

She believes there none of.

"She doesn't believe any of it".

It is worth noting that my use of the terms "strong" and "weak" to refer to the feminine form *ze* and *zij* should not be taken to imply a binomial or dichotomous view of referring expressions as a whole, because these forms (just like the masculine form *hij*) are part of a larger set of referring expressions that includes not only pronouns but also demonstratives, full nouns, and so on. The terms "strong" and "weak" are borrowed from existing work in this area and are used here as relative terms *within the personal pronoun category* to distinguish *zij* and the more attenuated *ze*.

## The contrast account and the salience account

Given that Standard Dutch offers two forms that can be used for feminine referents (*ze* and *zij*) in subject position but only one for masculine referents (*hij*), we are faced with the question of whether the weak form *ze* and the strong from *zij* differ in their referential properties, and how their referential properties compare to those of the masculine form *hij*. How do comprehenders, when faced with one of these forms, process them in order to locate the intended antecedent? Loosely speaking, two main views—which are not necessarily mutually exclusive—are represented in the literature. Some researchers have claimed that strong forms are used to pick out *lower-salience referents* (e.g., Cardinaletti & Starke, 1996, 1999), whereas others argue that strong forms are sensitive to the presence of *contrast*, possibly among other factors (e.g., Haeseryn et al., 1997, on Dutch; Pajusalu, 1995, 1997, on Estonian).

### The contrast account

The contrast account is largely based on work by Pajusalu (1995, 1997) on the Finno-Ugric language Estonian. Pajusalu uses corpus data to argue that the strong vs. weak pronoun distinction in Estonian is sensitive to the presence vs. absence of contrast. On this theory, when the antecedent stands in a salient opposition or some kind of contrastive relation to another entity in the domain of discourse, the strong form tends to be used (see also Kaiser, 2005b; Kaiser, 2010, for further work on contrast in Estonian). If we extend this to Dutch, comprehenders' interpretation of these two forms would proceed as follows. The strong from *zij* is more likely than the weak form *ze* to be interpreted as referring to salient antecedents that contrast with

other entities in the discourse, whereas the weak form *ze* tends to refer to noncontrastive salient antecedents.

In fact, the role of contrast is mentioned in Haeseryn et al.'s (1997, p. 252) descriptive grammar of Dutch as being one of the factors that guides use of strong forms. Although Haeseryn et al. do not provide detailed analyses or hypotheses, they note that strong pronouns can be used in contexts involving contrast, and present Example (2a) with the strong form of the second-person pronoun *jou* "you" (object position). Similarly, Donaldson's (1997, p. 55) Dutch grammar includes the example in (2b), with the strong form of the second-person pronoun *jij*, this time in subject position, in a context where *jij* and Marie are being contrasted.

(2a) Hij bedoelt *jou* niet, maar Mark. (Italics in original)

He means *you* not, but Mark.

"He isn't referring to *you*, he means Mark".

(2b) Heb jij je vrouw gesproken of heeft Marie je vrouw gesproken?

Did you your wife speak or did Marie your wife speak?

"Did you speak to your wife or did Marie speak to your wife?"

### The salience account

However, not everyone agrees that the strong and weak forms are differentiated by their sensitivity to contrast. Cardinaletti and Starke (1996, 1999) argue that strong pronouns like *zij* and weak pronouns like *ze* (which they call "deficient"; Cardinaletti & Starke, 1996, p. 43) differ in their morphosyntactic structure and the prominence/salience of their antecedents. According to their view, strong pronouns "are able to refer to a non-prominent discourse referent" (Cardinaletti & Starke, 1999, p. 154) in contrast to weak pronouns, which "must have an antecedent prominent in the discourse" (1999, p. 154). Cardinaletti and Starke's arguments are not specific to Dutch or to any other language for that matter (they intend them to be general claims about strong and weak pronouns in all human languages), but crucially *ze* vs. *zij* exhibit the appropriate linguistic behaviour to be characterised as weak ("deficient") vs. strong in their system. Thus, applying this account to Dutch leads to the prediction that comprehenders will differentiate *ze* vs. *zij* on the basis of salience, with *ze* being used to pick out more salient referents than *zij*.

Cardinaletti and Starke's claim that strong and weak forms refer to antecedents with different prominence levels resembles some aspects of the

widespread idea that there exists a connection between referential form and referent salience and that referential forms can be hierarchically organised in terms of how salient/prominent their antecedents are (e.g., Ariel, 1990; Givón, 1983; see also Aissen, 2003, on prominence hierarchies; see the Givenness Hierarchy by Gundel, Hedberg, & Zacharski, 1993, for a somewhat different approach using a form-function mapping that hinges on an implicational scale). Many researchers assume that the referential forms of a language are ranked along a salience hierarchy. Part of the standard hierarchy is given in (1); forms further to the left side are used to refer to more salient referents. (In languages like Spanish and Italian that allow subjects to be omitted/null, the "null" form is ranked above pronouns.) Thus, according to this view, the central antecedent property that guides reference resolution is level of salience.

(3) pronoun > demonstrative > full Noun Phrase…

The underlying assumptions of many salience-hierarchy approaches lead to the prediction that strong pronouns are used for less salient antecedents than weak pronouns. For example, according to Ariel (2001), referring expressions that are more attenuated/phonologically reduced have more salient antecedents than less reduced referring expressions. Thus, if we assume that *ze* and *zij* count as referentially distinct forms such that the weak form *ze* is ranked above the strong form *zij* on the salience hierarchy (…*weak pronoun > strong pronoun > demonstrative > full NP*…), the prediction is that salience is what distinguishes these two forms.

The question of what influences referent salience—i.e., how prominently a particular entity is represented in people's mental models of the discourse—has been heavily researched, and existing work shows that a number of factors play a role. Perhaps the most often-observed finding concerns grammatical role, specifically that entities realised in the (agentive) subject position are more salient than entities realised in object or oblique positions (e.g., Brennan, Friedman, & Pollard, 1987; Chafe, 1976; Crawley & Stevenson, 1990; see also Aissen, 2003; Yang & van Bergen, 2007). On the basis of this claim it seems reasonable to expect that, if interpretation of *ze* vs. *zij* is guided by salience and if the difference between the salience requirements of *ze* and *zij* is sufficient to be detected in their grammatical-role preferences, then *ze* may show a stronger preference for referring back to subjects than *zij* (assuming other factors are constant). (We focus here on pronouns realised in the subject position; parallelism considerations come into play for object pronouns, see e.g., Chambers & Smyth, 1998; Smyth, 1994). Furthermore, we may find that a subject preference emerges earlier with *ze* than with *zij* during real-time language comprehension.

In addition to the idea that weak pronouns refer to more salient referents than strong pronouns, it is also commonly agreed that pronouns are used for more salient referents than demonstratives, full NPs, and other forms (e.g., Ariel, 2001). As shown in (3), pronouns are ranked above demonstratives and other forms in the hierarchy. As a result, according to this type of salience approach, the prediction is that the masculine pronoun *hij* (in sentence-initial subject position) will show a preference for subjects over more oblique arguments. This is because, in salience terms, *hij* is the highest-ranked referring expression that can be used in subject position for a masculine referent, and so it is predicted to be used for the most salient antecedents (such as subjects).

### Stress/accent

When considering the salience account and the contrast account, the question of stress/accent often comes up. One might ask: if there is a correlation between strong vs. weak and stressed vs. unstressed (with strong pronouns correlating with stress/accent), could this explain the referential properties of strong and weak pronouns? Existing work suggests that: (1) it is not clear whether such a correlation exists, and (2) even if such a correlation was to exist, it would not clearly distinguish the contrast account from the salience account.

According to Cardinaletti and Starke (1999), cross-linguistically, the strong vs. deficient distinction cannot be reduced to prosodic or semantic focus. They claim that deficient pronouns can be prosodically focused/stressed in the right contexts (but see Zwart, 1993) and that strong pronouns do not have to be stressed. Looking specifically at Dutch, Haeseryn et al. (1997, p. 252) note that while weak pronouns must be unstressed, strong forms can be stressed or unstressed. Thus, the strong/weak pronoun distinction does not seem to map directly to the prosodically stressed vs. unstressed distinction. (A similar lack of a straightforward correlation has been observed in Estonian; see Pajusalu, 1995, on Estonian strong and weak pronouns.)

Furthermore, even if the strong vs. weak distinction was to correlate with the presence vs. absence of stress, this would *not* resolve the contrast vs. salience question concerning the interpretation of strong/stressed forms. Existing work on stressed pronouns in English is characterised by conflicting views, with some researchers arguing in favour of a contrast-based account (e.g., de Hoop, 2003) and others adopting a more salience-oriented approach (e.g., Kameyama, 1999; see also Nakatani, 1993). Thus, even when a form is clearly stressed, this does not provide us with an unequivocal answer to the "salience or contrast" question. This suggests that the question of what guides the interpretation of strong and weak pronouns needs to be

investigated regardless of whether or not strong pronouns are stressed/accented.

## Dutch demonstratives

In addition to personal pronouns, the distal demonstrative *die* ("that") can be used anaphorically to refer to masculine and feminine human antecedents in Dutch. (Like English "that", *die* can also be used as a pronominal modifier, e.g., *that man*, but we do not focus on that usage here.) In contrast to *ze/zij*, there is less controversy regarding the referential properties of anaphoric *die*. According to Haeseryn et al. (1997, p. 306), demonstratives are used to refer to entities that have just been introduced into the conversation (e.g., referents mentioned for the first time in the sentence preceding the demonstrative-containing sentence), while pronouns are used to refer back to "old information", entities that have already been the topic of conversation for a while (see also Geerts, Haeseryn, de Rooij, & van den Toorn, 1984). Similarly, Rullmann (2001) concluded, based on a corpus study, that pronouns prefer topical or discourse-old antecedents and demonstratives tend to refer back to nontopics and/or new information (see also Comrie, 1997). These findings converge with work on anaphoric demonstratives in other languages (e.g., Bosch, Rozario, & Zhao, 2003, on German; Kaiser, 2003; Kaiser & Trueswell, 2008, on Finnish; Kibrik, 1996; Krasavina & Chiarcos, 2007, on Russian).Thus, there is a fairly clear consensus concerning the demonstrative *die*: it prefers lower-salience, nontopical referents—as suggested by the salience hierarchy shown in (3). If we follow the widespread view that objects and obliques are lower in salience than subjects, the expectation is that *die* will prefer objects and obliques over subjects, as in Example (4), assuming other salience-influencing factors are controlled.

(4) Mark kwam Arthur$_i$ tegen. Die$_i$ droeg en regenjas.

Mark ran into Arthur$_i$. That$_i$ was wearing a raincoat. (Rullmann, 2001, from Geerts et al., 1984)

## Overview of experiments

The experiments presented here aim to investigate the salience account and the contrast account, two hypotheses that have been put forth concerning the interpretation of strong and weak pronouns. The three experiments reported in this paper investigate how comprehenders process the strong and weak forms in the feminine paradigm in Dutch (*ze/zij*) and compare their properties to those of the demonstrative *die* and the masculine pronoun *hij*. The referential properties of strong and weak pronominal forms are not yet well

understood; and in addition to shedding light on their interpretation, this research will also have broader implications for our understanding of the nature of the form-function mapping in anaphoric paradigms and the extent to which reference resolution can(not) be explained in terms of a unified notion of referent salience. I report the results of a sentence-completion study and an eye-tracking experiment that tested whether strong forms are used for relatively lower-salience referents (**Experiments 1a and 1b**), and a forced-choice sentence-completion study that investigated whether strong forms are used for contrastive topics (**Experiment 2**).

# EXPERIMENT 1A: INVESTIGATING EFFECTS OF GRAMMATICAL ROLE

As a first step in investigating the referential properties of weak and strong referential forms in Dutch, I conducted a sentence-completion study with two main aims: (1) to test the existing claims regarding the salience differences between the demonstrative *die* and the pronoun *hij* in an experimental setting by looking at how likely they are to be interpreted as referring to preceding subjects vs. objects, and (2) to investigate whether the interpretation of *zij* and *ze* is also influenced by the grammatical role of potential antecedents. The sentence-completion study investigated how frequently, when presented with a transitive sentence followed by a prompt pronoun/demonstrative, participants treated the prompt pronoun as referring to the preceding subject vs. object.

The predictions are straightforward for the demonstrative *die* and the masculine pronoun *hij*. Based on existing work, *die* is predicted to be used for lower-salience referents and the masculine pronoun *hij* for higher-salience referents. In terms of grammatical role, the prediction is that *hij* and *die* will both show significant effects of grammatical role, with *hij* being used mostly for the preceding subject and *die* for the preceding object. For the strong feminine pronoun *zij* and the weak feminine pronoun *ze*—given the widespread view that entities in subject position are more salient than entities in object position—a finding that *ze* has a stronger subject preference than *zij* would (at least at first glance) fit with the view that *ze* is used for more salient antecedents than *zij*. Explicit testing of the effects of contrast (i.e., the idea that *zij* refers to entities that contrast with something else; *ze* is used for salient but noncontrastive entities) is tackled in Experiment 2.

## Method

Forty adult native Dutch speakers participated in a standard sentence-completion task over the internet. The experiment consisted of 16 targets and 48 fillers. Target items consisted of SVO sentences followed by the first word

of the next sentence: either *hij*, *die*, *ze*, or *zij* (see Examples (5a) and (5b)). Four conditions were tested in a Latin-Square design: (1) two masculine characters followed by *hij*; (2) two masculine characters followed by *die*; (3) two feminine characters followed by *ze*; and (4) two feminine characters followed by *zij*.[2]

(5a) De brandweerman kneep de bokser speels. Hij/Die . . .

The fireman pinched the boxer jokingly. He . . .

(5b) De serveerster kneep de onderwijzeres speels. Ze/Zij . . .

The waitress pinched the teacher jokingly. She . . .

Participants were asked to provide natural-sounding continuations. The subject and object nouns in the critical items were selected so as to be maximally clear in their gender properties; they were morphologically marked for gender (e.g., *leraar/lerares* "male high school teacher/female high school teacher") or their gender was otherwise clear (e.g., king vs. queen). The verbs were action/agent-patient verbs (as defined by Stevenson, Crawley, & Kleinman, 1994). The continuations were coded according to which of the referents in the preceding sentence (subject, object, or unclear) the participants chose as the referent of the anaphor. When participants interpreted the demonstrative *die* as a discourse deictic (e.g., "this was a silly thing to do") or a prenominal modifier (e.g., "this teacher"), the continuation was coded as "other". Continuations which treated the feminine pronouns *ze/zij* as plural "they" were also coded as "other". (In Dutch, "she" and "they" are identical: both are denoted with *ze/zij*.)

## Results and discussion

Table 1 shows percentages of different continuation types for each of the four referential forms. Overall, *hij* triggered a high rate of subject continuations (75.6%), whereas *die* resulted in a high rate of object continuations (81.3%). Resembling *hij*, the short feminine form *ze* triggered mostly subject continuations (63.1%); and the long feminine form *zij* resulted in 50% subject continuations.

First, to assess the referential preferences of the different forms statistically, one-sample *t*-tests were used to test whether the proportions of subject continuations differ from chance. These analyses were conducted

---

[2] I did not test *die* preceded by two feminine characters due to reasons of experiment length, and because preliminary investigations suggested that *die* with two feminine referents resembles *die* with two masculine referents.

TABLE 1
How many times different continuation types occurred for each of the four anaphoric forms

|     | *Subject* | *Object* | *Unclear* | *Other* |
| --- | --- | --- | --- | --- |
| Hij | 121<br>75.6% (4.6) | 23<br>14.4% (3.6) | 16<br>10% (2.8) | 0<br>0% (0) |
| Die | 1<br>0.6% (0.6) | 130<br>81.3% (3.5) | 8<br>5% (2.0) | 21<br>13.1% (3.8) |
| Ze | 101<br>63.1% (4.6) | 32<br>20% (3.5) | 20<br>12.5% (2.7) | 7<br>4.4% (1.8) |
| Zij | 80<br>50% (4.7) | 53<br>33.1% (5.3) | 18<br>11.3% (2.0) | 9<br>5.6% (1.9) |

*Note*: Percentages computed out of a total of 160 observations per anaphoric form. Standard errors are given in parentheses.

on the proportion of subject and object continuations only (with "other" and "unclear" excluded), such that the hypothesised mean was 0.5. Here, as well as for all other results reported in this paper, analyses were conducted both on raw proportion data and on arcsine-transformed data, to help compensate for the fact that proportion data are bounded between 1 and 0. The statistics are reported for the analyses of the raw data. The transformed data yielded the same significance patterns, except where stated.

Participants' continuations show that the pronoun *hij* and the demonstrative *die* have clear referential biases. As predicted, *hij* triggers a higher-than-chance proportion of subject continuations [one-sample $t$-test: $t_1(39) = 8.254$, $p < .001$, $t_2(15) = 9.076$, $p < .001$]. In contrast, *die* results in lower-than-chance proportion of subject continuations [one-sample $t$-test: $t_1(39) = -29.651$, $p < .001$, $t_2(15) = -71$, $p < .001$]. The short feminine pronoun *ze* resembles *hij* in triggering a higher-than-chance proportion of subject continuations [one-sample $t$-test: $t_1(39) = 5.84$, $p < .001$, $t_2(15) = 5.898$, $p < .001$]. The long form *zij* exhibits some hints of a higher-than-chance proportion of subject continuations [one-sample $t$-test: $t_1(39) = 1.922$, $p = .062$, $t_2(15) = 1.591$, $p = .133$; arcsine transformed: $t_1(39) = 2.199$, $p < .05$, $t_2(15) = 1.594$, $p = .132$]—the effect is not significant by items but significant by subjects in the arcsine-transformed analyses (marginal in the untransformed analyses). Thus, with the long form *zij*, the patterns are less clear.

To compare the conditions to each other, I conducted one-way ANOVAs, followed by Bonferroni-corrected contrasts. Analyses were run on the proportion of subject continuations and object continuations. The ANOVAs revealed significant main effects of anaphoric form on the proportion of subject

continuations, $F_1(3, 117) = 91.877$, $p < .001$, $F_2(3, 45) = 82.618$, $p < .001$, and object continuations, $F_1(3, 117) = 61.556$, $p < .001$, $F_2(3, 45) = 67.117$, $p < .001$. Bonferroni-corrected contrasts were used to further investigate the differences between the four conditions. The analyses show that *hij* triggers a significantly higher rate of subject continuations than *die*, mean difference $= 0.75$, 95% CI for difference $= \pm 0.127$ by subjects ($\pm 0.1$ by items), and that *die* results in a significantly higher rate of object continuations than *hij*, mean difference $= 0.699$, 95% CI $= \pm 0.16$ ($\pm 0.132$). Compared to the demonstrative *die*, the two feminine forms, *ze* and *zij*, also resulted in a significantly higher proportion of subject continuations [*ze* vs. *die*: mean difference $= 0.625$, 95% CI $= \pm 0.127$ ($\pm 0.125$); *zij* vs. *die*: mean difference $= 0.494$, 95% CI $= \pm 0.13$ ($\pm 0.162$)] and a lower proportion of object continuations than *die* [*ze* vs. *die*: mean difference $= -0.613$, 95% CI $= \pm 0.16$ ($\pm 0.13$); *zij* vs. *die*: mean difference $= -0.481$, 95% CI $= \pm 0.194$ ($\pm 0.19$)].

However, the two feminine forms, *ze* and *zij,* do not differ significantly from each other in terms of the proportion of subject or object continuations. The feminine short form *ze* also does not differ reliably from the masculine *hij* in the proportion of subject or object continuations. (In the by-subjects analysis of the proportion of subject continuations in arcsine-transformed data, there appears to be a hint of *ze* triggering less subject continuations than *hij*; *ze* vs. *hij*: mean difference $= -0.196$, 95% CI $= \pm 0.196$. However, this is only marginal in the by-subjects analysis of the raw data [*ze* vs. *hij*: mean difference $= -0.125$, 95% CI $= \pm 0.136$ by subjects ($\pm 0.151$ by items)] and not significant in the transformed and raw-data items analyses ($ps > .14$), and thus does not provide a reliable indication of *ze* and *hij* differing in this respect.)

In contrast to the short form *ze*, the long form *zij* did result in more object continuations and fewer subject continuations than the masculine form *hij* [subject continuations: mean difference $= -0.256$, 95% CI $= \pm 0.135$ ($\pm 0.179$); object continuations: mean difference $= 0.188$, 95% CI $= \pm 0.131$ ($\pm 0.201$; marginal by items)].[3]

The results for *die*, *hij*, and *ze* confirm the prediction that the demonstrative *die* is used for lower-salience referents whereas *hij* and *ze* prefer more salient antecedents. In fact, the numbers in Table 1 show that *die* is more extreme in its preferences ($< 1\%$ subject choices) than *hij* and *ze* (14.4% and 20% object choices, respectively). This asymmetry in "strength of preference" fits with a corpus-based observation by Comrie (1997), namely

---

[3] These differences are significant in both the by-subjects and by-items analyses of the proportion of *subject continuations* in both the raw and the arcsine-transformed data ($ps < .005$); they are also significant in the by-subjects analyses of the *object continuations* in both the raw and the transformed data ($ps < .005$). They are marginal in the by-items analysis of the object continuations in the raw ($p = .077$) and transformed data ($p = .062$).

that pronouns *can* be used to refer to the less-preferred antecedent even though this is not as frequent as reference to the preferred antecedent, whereas demonstratives are a more marked option and are used to exclude one of the potential antecedents. In other words, we observe an asymmetry in the strength of referential biases: *reference to the less-preferred antecedent* is more frequent with pronouns (reference to objects occurs occasionally) than with demonstratives (reference to subjects is extremely rare).

The results for the strong pronoun *zij* present a more complex picture. On the one hand, the long form *zij* does not differ significantly from the short form *ze*, but on the other hand, *zij* results in more object interpretations than the masculine *hij*. This, combined with the numerical differences between *ze* and *zij* (*ze*: 63.1% subject continuations vs. *zij* with 50% subject continuations) brings up the possibility that perhaps *zij* is used to refer to antecedents that are slightly less salient than the ones that *ze* and *hij* pick out (but not as low in salience as the referents that *die* is used for), but that the methodology used in Experiment 1a was not sufficiently sensitive to detect this. In order to test whether this "intermediate salience" idea for *zij* has any merit, Experiment 1b used a more sensitive method, visual-world eye-gaze tracking.

The interpretation of the *ze/zij* conditions is complicated by the fact that the absence of a significant subject preference with *zij*, coupled with a subject preference with *ze*, could also be explained by the contrast account. If *zij* (but not *ze*) is sensitive to the presence of contrast, it could presumably be used to refer back to whichever argument, subject or object, is construed as contrasting with something else. Because the sentence-completion approach used in Experiment 1a was an open-ended task, it potentially allowed participants to construe the mentioned characters as contrasting with other entities.[4] Thus, the reasons for the finding that the proportion of subject continuations triggered by *zij* does not differ from chance are not yet clear.

To help answer these remaining questions, **Experiment 1b** tested effects of antecedent grammatical role while minimising the possibility of contrastive construals. If the results nevertheless show that *zij* falls between *ze* and *die* in its subject preference (as measured by the proportion and timing of looks to the preceding subject vs. object), this would be strong evidence for the grammatical-role-based salience account. (The role of contrast is investigated directly in the final study, Experiment 2.)

---

[4] A possible way of investigating this issue directly would be to analyse participants' continuations to see whether the *zij* conditions are more likely to involve contrastive construals than the *ze* conditions. However, preliminary attempts revealed that such coding could not be done reliably, due to the open-endedness of the task and the limited amount of context available in this kind of task. If one were to change the methodology slightly and ask participants to write a longer continuation consisting of several sentences and/or sentences that could *precede* the test sentence, then it might be easier to detect potential effects of contrast.

# EXPERIMENT 1B: INVESTIGATING EFFECTS OF GRAMMATICAL ROLE DURING ON-LINE PROCESSING

Experiment 1b used visual-world eye tracking to test how the grammatical role of potential antecedents influences the real-time interpretation of Dutch anaphoric forms. This experiment builds on Experiment 1a by taking a closer look at the strength of the subject preference exhibited by the strong from *zij* and the weak from *ze*, in a context where the likelihood of contrastive interpretations is minimised.

Visual-world eye tracking allows us to investigate the temporal properties of on-line anaphor resolution using a natural dependent measure, namely eye movements. Eye movements provide a very useful measure for investigating reference resolution because of their automatic nature (e.g., Palmer, 1999) and the tight connections that have been found between eye movements and referential processing. Specifically, existing research has shown that eye movements to objects in a display or scene are closely time-locked to potential referents that a listener is considering as language unfolds over time (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; for a review see Tanenhaus & Trueswell, 2006).

In Experiment 1b, a visual-world setting was used where participants listened to sentences containing anaphors while looking at scenes that depicted the characters mentioned in the sentences. In this situation, we can analyse where participants look as they hear the anaphoric expressions. By analysing the timing and proportion of participants' eye movements to the potential antecedents in the scene, we can shed light on what entities participants consider as potential referents for anaphoric expressions during real-time processing. This kind of time-course information can potentially provide useful insights into when, relatively speaking, referential preferences emergence for the different forms being investigated, and which potential antecedents compete for selection before the final decision is reached. Importantly, the contexts in which the anaphoric forms occurred in Experiment 1b were also designed to minimise the possibility of participants construing any of the potential antecedents as contrastive, in order to facilitate interpretation of the results.

## Method

### Participants

Sixteen native Dutch-speaking participants, mainly students at the University of Nijmegen in the Netherlands, took part in the experiment. Participants received approximately $4 for participation in the experiment.

*Procedure*

An eye-movement-during-listening paradigm was employed in which participants heard descriptions of clip-art-generated pictures (similar to Altmann & Kamide, 1999; Arnold et al., 2000; Kaiser & Trueswell, 2004b, 2008). Participants were shown, on a computer screen, large colour pictures of simple scenes involving human or animal characters, and listened to a short pre-recorded story about the scene. Participants were told that in some cases, the story might not match the picture, and that in such cases, their task was to correct (by speaking out loud) the story according to what they saw in the picture.

A digital camera was used to record participants' eye movements during the experiment. On each trial, the participant was shown a large colour picture on the computer screen, and directly above the monitor was a SONY DVcam digital camcorder with audio-lock recording. The DVcam camcorder was centred directly above the monitor, and recorded the participant's face and eyes, the auditory stimuli and the participant's spoken responses. The pre-recorded sound files were played by a Dell laptop over external stereo speakers. Analyses of the eye movements and speech onsets, described below, were done by hand on the videotapes at a later date, using a SONY DSR-30 digital VCR with jog-shuttle control.

The Trueswell Lab has used this type of eye-gaze technique successfully on adult and child participants (see Kaiser & Trueswell, 2004b, 2008; Snedeker, Thorpe, & Trueswell, 2001; Snedeker & Trueswell, 2004; see also Snedeker & Yuan, 2008). Snedeker and Trueswell (2004) discuss the nature and validity of this kind of eye-gaze technique in depth, and provide a comparative analysis showing that a free-head video-based eye-gaze procedure produces data equivalent to that of an ISCAN head-mounted eye tracker. This method resembles preferential-looking studies with children, which result in high inter-coder reliability when frame-by-frame coding is used (Hirsh-Pasek & Golinkoff, 1996).

*Materials*

The visual stimuli for this experiment consisted of colour images presented full size on a standard computer screen. Participants were seated so that their faces were approximately 30 cm away from the computer screen. Pictures typically contained two to five animate entities (people or animals) and other objects that made up a coherent scene. These images were generated from clip-art images, and arranged and edited using Adobe Photoshop. For each picture, a brief story involving the characters shown in the picture was prepared. The same female native Dutch speaker's voice was used for recording all sound files. The recording was done in a sound-proof cabin using a Sony Electret Condenser microphone and a DAT recorder, and

recorded at a sampling frequency of 44.1K. Native-speaker judgements confirmed that the *ze/zij* distinction was audible in the resulting sound files.

A total of 16 target items (i.e., scene–story pairs) were constructed. The scenes consisted of two easily identifiable human characters. The characters were positioned such that one was to the left side of the image, and the other was to the right. In the clip-art scenes, the positions of the referents were counterbalanced such that on half of the trials the subject was on the left and on half the trials it was on the right. The verbal story for each target item contained a sentence with two masculine or two feminine human arguments, followed by the critical sentence beginning with the anaphor *ze, zij, hij*, or *die*. (Some of the nouns were gender-marked, e.g., *leraar* "male teacher"/*lerares* "female teacher" in Example (6). In all cases, the scenes made it clear that both characters were either masculine or feminine.) Thus, there were four conditions: two masculine characters followed by the masculine pronoun *hij*, two masculine characters followed by the demonstrative *die*, two feminine characters followed by the weak feminine form *ze*, and two feminine characters followed by the strong feminine form *zij*. The average durations of the four anaphoric forms were as follows: *ze* 139 ms, *zij* 170 ms, *hij* 143 ms, and *die* 100 ms. Sample masculine and feminine scenes are shown in Figure 1. The stories for the scenes are shown in (6).

(6a) Sample item in the masculine condition

Het begon uit de hand te lopen in het klaslokaal.

"Things were beginning to get out of hand in the classroom".

De leerling stak de leraar speels met een scherp potlood.

"The student poked the teacher jokingly with a sharp pencil".



**Figure 1.** Sample pictures for masculine and feminine conditions of Experiment 1b.

**Hij/Die** was gekleed in een groene trui, omdat het buiten koud was.

"**Hij/Die** was wearing a green sweater, because it was cold outside".

Het lijkt erop dat ze naar de rector moeten.

"It looks like they will have to go see the principal".

(6b) Sample item in the feminine condition

Het begon uit de hand te lopen in het klaslokaal.

"Things were getting out of hand in the classroom".

De leerlinge stak de lerares speels met een scherp potlood.

"The student poked the teacher jokingly with a sharp pencil".

**Ze/Zij** was gekleed in een groene trui, omdat het buiten koud was.

"**Ze/Zij** was wearing a green sweater, because it was cold outside".

Het lijkt erop dat ze naar de rector moeten.

"It looks like they will have to go see the principal".

Each story began with an opening sentence, which was followed by a sentence introducing the two feminine or two masculine characters. The verbs used in this sentence were all agent-patient verbs (e.g., "poke") used with instrument phrases (e.g., "with a sharp pencil"). The instruments mentioned in the instrument phrase constituted the look-away object, i.e., an object located elsewhere in the scene whose aim was to encourage participants to look away from both mentioned characters at a neutral location.

The third sentence, which began with the anaphoric expression, is incorrect with respect to both referents in the picture (e.g., in this example, neither the student nor the teacher is wearing a green sweater, though both are wearing some other piece of clothing that is green), and thus the participant is expected to provide a verbal correction (see Kaiser & Trueswell, 2008, for a similar design). The average onset of the "error" word—i.e., the word that reveals that the sentence is incorrect with respect to both referents, e.g., "sweater" in Example (6)—was 1,054 ms after onset of the anaphor. As the stories were designed to minimise the possibility of construing any of the mentioned referents as contrastive, the story context

1604    KAISER

did not explicitly contrast the characters to each other, and the speaker used neutral intonation.

The third sentence was followed by a wrap-up sentence which, along with the opening sentence, was intended to make the stories sound natural and coherent.

Thirty-two fillers were also constructed, varying in the number of characters pictured and their locations in the scene. The characters used in the fillers and the critical items were all different from each other, such that each character was seen only once by a given participant. Twenty-four fillers were correct and eight contained mismatches. This was done to ensure that over the entire experiment, half of the trials were correct and half contained mismatches/mistakes.

Four presentation lists were constructed for this study by pseudo-randomly combining the 16 target stories with the 32 filler stories. Each target item was separated by at least one filler item from any other target item. Within a presentation list, eight of the target trials appeared with two feminine referents and eight appeared with two masculine referents. Each target item was rotated through the four conditions, generating four different presentation lists. Reverse order lists were also generated to control for trial order.

### Data analysis and coding

The videotapes of the participants' eye gaze were analysed as follows. A Sony DRS-30 digital VCR (which allows for frame-by-frame inspection of the video and audio components, at the frequency of 30 frames per second) was used to go through the audio portion of each videotape. The frame where the critical sentence begins (i.e., the onset of the anaphor) was located on the videotape by listening to it frame by frame. (Onset coding followed the procedure used by Snedeker, Thorpe, and Trueswell, 2001; Snedeker and Trueswell, 2004.) The video was then analysed frame by frame (with the sound turned off) for 10 seconds (300 frames), beginning one second (30 frames) before the onset of the critical sentence. Coders recorded, frame by frame, whether the participant was looking to the left, right, centre, or elsewhere. Since the sound was turned off, coders were blind to experimental condition. The eye-movement coding was used to establish which characters participants had looked at over time, relative to the onset of the anaphoric form. To determine the reliability of the coding, the video record of three participants was fully double coded. The two scorers were in agreement on over 96% of the video record.

Participants' verbal responses for each trial, including which referent each correction referred to, were also recorded and analysed.

## Results

Figure 2 shows the proportion of looks to the **subject** of the preceding sentence, in all four conditions, plotted as a function of time. In the masculine conditions, starting approximately 600 ms after anaphor onset, the pronoun *hij* triggers more looks to the preceding subject than the demonstrative *die*, as predicted. As the graph illustrates, this preference becomes increasingly pronounced over time. For the feminine conditions there is no clear distinction between the weak form *ze* and the strong form *zij* at any point in time. Generally speaking, both seem to prompt a high proportion of looks to the subject, comparable to the masculine pronoun *hij*. Figure 3 shows the proportion of looks to the **object** of the preceding sentence, in all four conditions. The demonstrative *die* triggers more looks to the object than the other three conditions, starting approximately 700 ms after anaphor onset. On the whole, the graphs show that pronouns *hij*, *ze*, and *zij* behave in the same way, and differ from the demonstrative *die*.

To analyse these eye-movement patterns in more detail, I conducted one-way ANOVAs, followed by Bonferroni-corrected contrasts, to compare the four conditions. These analyses were run on the proportion of looks to the subject and the proportion of looks to the object on eight 400-ms time slices, starting 800 ms before the onset of the anaphoric expression and continuing for 3,200 ms. For each time slice, participant and items means of the proportion of looks to the subject and the object of the preceding sentence were analysed. (Analyses were conducted both on raw proportion data and on arcsine-transformed data, to help compensate for the fact that proportion data are bounded between 1 and 0. The statistics are reported for the analyses of the raw data. The transformed data yielded the same significance
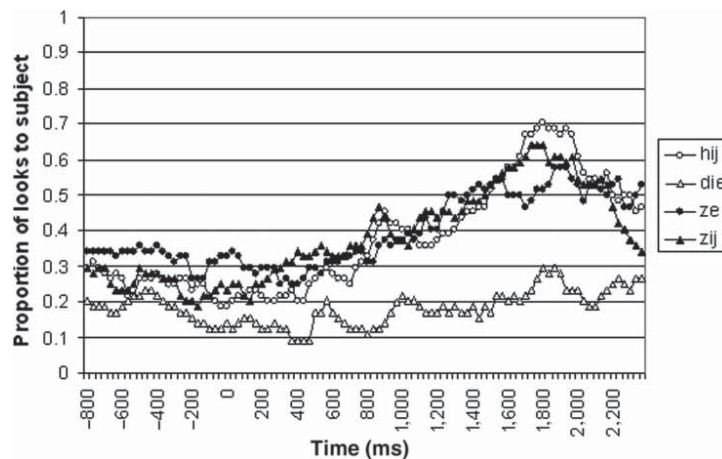


**Figure 2.** Probability of fixating the subject of the preceding sentence as a function of time. (Onset of the anaphoric expression is at 0 ms.)
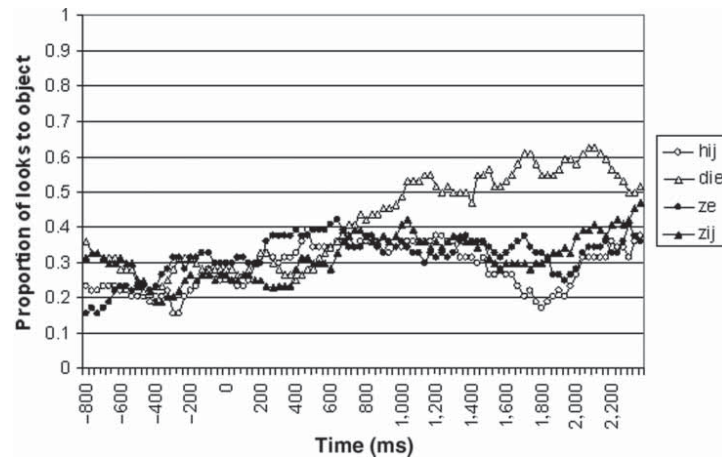
**Figure 3.**    Probability of fixating the object of the preceding sentence as a function of time.

patterns, except where stated.) When looking at the eye-movement patterns, it is important to keep in mind that it takes about 200 ms to programme and execute an eye movement (see Hallett, 1986), and thus the earliest point at which we can expect to see anaphor-driven eye movements is about 200 ms after recognition of the anaphor.

As can be seen in Table 2, during the first three time slices (from –800 to –400, –400 to 0, and 0 to 400 ms), the one-way ANOVAs reveal no significant differences between the conditions either in the proportion of looks to subject or proportion of looks to object. (The third time slice is the only one where the statistical significance patterns for the arcsine-transformed data differ from the raw proportion data. The one-way ANOVA on the transformed data—but not the raw data—reveals significant effects of anaphor type for the proportion of subject looks, $F_1(3, 45) = 3.104$, $p < .05$, $F_2(3, 45) = 2.974$, $p < .05$; hinting at the patterns in the next time slice.)

Starting with the fourth time slice (400–800 ms), we see significant effects of anaphor type on the proportion of *subject looks* for all remaining time slices. Significant effects of anaphor type on the proportion of *object looks* start to emerge in the sixth time slice (1,200–1,600 ms) and persist for all subsequent time slices.

Bonferroni-corrected contrasts were used to further investigate the differences between the four conditions in the time slices where the ANOVA revealed significant effects. Let us first consider the patterns for the masculine conditions. Starting in the fourth time slice (400–800 ms), *hij* triggered a significantly higher proportion of looks to the <u>subject</u> than *die* [*fourth time slice*: mean difference = 0.125, 95% CI for difference = ± 0.121 by subjects

TABLE 2
Results of one-way ANOVAs for Experiment 1b, visual-world eye tracking

| Time slice (ms) | Looks to subject (by subjects) | Looks to subject (by item) | Looks to object (by subjects) | Looks to object (by items) |
|---|---|---|---|---|
| –800 to –400 | $F_1(3, 45) = 1.49$, $p = .252$ | $F_2(3, 45) = 1.178$, $p = .329$ | $F_1(3, 45) = 0.795$, $p = .503$ | $F_2(3, 45) = 1.126$, $p = .349$ |
| –400 to 0 | $F_1(3, 45) = 1.51$, $p = .225$ | $F_2(3, 45) = 1.911$, $p = .141$ | $F_1(3, 45) = 0.541$, $p = .657$ | $F_2(3, 45) = 0.639$, $p = .594$ |
| 0–400 | $F_1(3, 45) = 2.479$, $p = .073$ | $F_2(3, 45) = 1.917$, $p = .14$ | $F_1(3, 45) = 0.369$, $p = .776$ | $F_2(3, 45) = 0.585$, $p = .628$ |
| 400–800 | $F_1(3, 45) = 4.598$, $p < .05$ | $F_2(3, 45) = 4.092$, $p < .02$ | $F_1(3, 45) = 0.173$, $p = .914$ | $F_2(3, 45) = 0.166$, $p = .919$ |
| 800–1,200 | $F_1(3, 45) = 5.203$, $p < .005$ | $F_2(3, 45) = 4.1$, $p < .02$ | $F_1(3, 45) = 1.652$, $p = .191$ | $F_2(3, 45) = 0.961$, $p = .419$ |
| 1,200–1,600 | $F_1(3, 45) = 11.42$, $p < .001$ | $F_2(3, 45) = 7.354$, $p < .001$ | $F_1(3, 45) = 3.6$, $p < .05$ | $F_2(3, 45) = 2.79$, $p = .051$ |
| 1,600–2,000 | $F_1(3, 45) = 16.233$, $p < .001$ | $F_2(3, 45) = 13.291$, $p < .001$ | $F_1(3, 45) = 14.181$, $p < .001$ | $F_2(3, 45) = 11.053$, $p < .001$ |
| 2,000–2,400 | $F_1(3, 45) = 7.622$, $p < .001$ | $F_2(3, 45) = 9.175$, $p < .001$ | $F_1(3, 45) = 5.129$, $p < .005$ | $F_2(3, 45) = 5.053$, $p < .005$ |

($\pm 0.21$ by items, ns by items); *fifth time slice*: mean difference = 0.223, 95% CI = $\pm 0.179$ ($\pm 0.221$); *sixth time slice*: mean difference = 0.271, 95% CI = $\pm 0.197$ ($\pm 0.242$); *seventh time slice*: mean difference = 0.411, 95% CI = $\pm 0.195$ ($\pm 0.228$); and *eighth time slice*: mean difference = 0.289, 95% CI = $\pm 0.256$ ($\pm 0.143$)].

Conversely, *die* triggered a significantly higher proportion of looks to the object than *hij*, but this effect did not reach significance until the seventh time slice, mean difference = 0.354, 95% CI = $\pm 0.181$ ($\pm 0.211$). It remains significant in the by-items analyses in the eighth time slice, mean difference = 0.240, 95% CI = $\pm 0.258$ ($\pm 0.161$).

Crucially, Bonferroni-corrected contrasts showed that unlike *hij* and *die*, the feminine weak and strong forms *ze* and *zij* do not differ significantly from each other during any time slice, either in the proportion of subject looks (*p*s > .46) or the proportion of object looks (*p*s > .57). They also do not differ significantly from the masculine pronoun *hij* during any time slice (subject looks *p*s > .297 and object looks *p*s > .413).

Participants' **verbal corrections**, which provide an indication of their off-line referential judgements, matched the eye-movement patterns. For example, in Example (6b) a participant might say, "No, that is not right, she is wearing a red sweater", which would indicate that the pronoun had been interpreted as referring to the student (who is wearing a red sweater)

and not the teacher (who is wearing a purple sweater).[5] In conditions with the masculine pronoun *hij*, participants interpreted *hij* as referring to the preceding subject in 78% of the cases, 95% CI $= \pm 0.11$ ($\pm 0.156$). In contrast, the demonstrative *die* was interpreted as referring to the preceding object in 80% of the cases, 95% CI $= \pm 0.203$ ($\pm 0.159$). The feminine strong pronoun *zij* was treated as referring to the subject in 66% of the responses, 95% CI $= \pm 0.121$ ($\pm 0.186$), and the weak form *ze* was interpreted as referring to the subject in 65.3% of the responses, 95% CI $= \pm 0.139$ ($\pm 0.24$).

## Discussion

Overall, the results of this eye-tracking study confirm the finding from Experiment 1a that **the pronoun *hij* "he" and the demonstrative *die* "that"** differ in their referential properties; *hij* is significantly more likely than *die* to be interpreted as referring to the subject of the preceding sentence, whereas *die* triggers significantly more looks to the object. In contrast to the striking differences exhibited by *hij* and *die*, the eye movements triggered by **the strong and weak forms of the feminine pronoun**, ***ze* and *zij***, are very similar. Both indicate an equally strong preference to be interpreted as referring to the subject of the preceding sentence. Thus, in a context where the possibility of contrastive interpretations is minimised, we do not see any hints of a difference in *ze* and *zij*'s subject-preference strength.

In summary, the eye-movement patterns suggest that the masculine pronoun *hij* and the weak and strong feminine forms *ze* and *zij* all exhibit a clear preference for the preceding subject, whereas the demonstrative *die* mainly triggers looks to the object. In addition, the eye-movement patterns for *hij* and *die* provide further support for the differences in the strength of referential biases that we observed in Experiment 1a. Recall that participants' sentence completions indicated that *die* has a very strong object preference (only 0.6% subject interpretations) whereas *hij*'s subject preference is not as strong (14.4% object interpretations). We also find indications of this asymmetrical "preference strength" in the eye-movement patterns. More specifically, although the proportions of looks triggered to the <u>subject</u> by *hij* and *die* start to differ already during the 400–800 ms time slice, there are no significant differences in looks to the <u>object</u> until the 1,600–1,800 ms time slice. A closer look at the graphs suggests that this is probably due to the asymmetry in the strength of the referential preferences of *hij* vs. *die*, something that we already observed in Experiment 1a and that Comrie

---

[5] Some corrections were ambiguous (e.g., "No, no one is wearing a green sweater") and could not be used to distinguish the two referents. Also, sometimes by the end of the story participants apparently had forgotten the details and simply responded "correct". These kinds of responses are excluded from the percentages reported here. However, the basic pattern of subject vs. object preferences does not change even if these responses are included.

(1997) also mentioned. The pronominal form *hij* (as well as *ze* and *zij*) primarily triggers looks to the subject, and it also results in *some* looks to the dispreferred antecedent, the object, especially in the first few time slices after anaphor onset. This contrasts with *die*, which results in very few looks to its dispreferred antecedent (the subject) at any point in time. The eye-movement patterns suggest that, due to *die* exhibiting a clearer referential bias than *hij*, the distinction in the proportion of object looks triggered by *die* and *hij* takes longer to emerge than the distinction in the proportion of subject looks.

Participants' spoken corrections fit with the eye-tracking results (as well as the sentence-completion results) in showing equal levels of subject preference for *ze* and *zij*, as well as a subject preference in the *hij* conditions and an object preference in the *die* conditions.

## EXPERIMENT 2: INVESTIGATING EFFECTS OF CONTRAST

Experiment 2 investigated whether the presence/absence of contrast influences the interpretation of strong and weak pronouns in Dutch. Thus, this experiment focused only on *ze* and *zij*, and tested their interpretation in contrastive and noncontrastive contexts. Although the role of contrast in triggering strong pronoun use in Dutch has been commented on briefly in existing work (e.g., the descriptive grammar of Haeseryn et al., 1997), as far as I know there is no published psycholinguistic work investigating this issue in depth. Experiment 2 aims to further our understanding of the role that contrast plays in the choice of *ze* vs. *zij* by testing (1) whether contrastive topics are more likely to be referred to with *zij* or *ze*, and (2) whether choice of *ze* vs. *zij* is sensitive to the information structure of the rest of the sentence, in particular the size of the set of alternatives.

The term "contrast" has been used in different ways by different researchers (e.g., É. Kiss, 1998; Rooth, 1992; Vallduví & Vilkuna, 1998; Zimmermann, 2007; see Molnár, 2006, for an overview), but it is generally agreed that, at least on an intuitive level, contrast has to do with contextually salient alternatives. For the purposes of the contrast account, the relevant configuration is one where *zij* is used to refer to an entity that is salient in the discourse model and the discourse model contains at least one salient alternative to that entity. The weak form *ze* can then be described as referring to an entity that is salient in the discourse model in a situation where there are no salient alternatives to that entity. To formalise the notion of alternatives more explicitly, I build on work in Alternative Semantics (Rooth, 1992), as well as research on contrastive topics by Büring (2003) and Jackendoff (1972). Consider Example (7), discussed by Jackendoff and Büring:

(7a) Well, what about FRED? What did HE eat?

(7b) [FRED]$_{CT}$ ate the [BEANS]$_F$.

The question in (7a) can felicitously be uttered in a context where we are discussing what Fred ate, what Tina ate, what Mary ate, and so on. Thus, in the answer in (7b), Fred is contrastive in that he is a member of a salient set under discussion (the set of people who ate something) and he contrasts with the other members of that set in terms of what he ate. Thus, the entities that are allowed to count as alternatives to Fred are the members of this set. Büring (2003) calls FRED the contrastive topic (subscript CT) since he was also mentioned in the question and is thus old information. Büring refers to the object BEANS as the focus (subscript F), the new information that answers the question for each person. Of course, which parts of a sentence are construed as the CT and the F (and hence what accents they are realised with) depends on the context in which the sentence is uttered. If someone asks the question "Well, what about the BEANS? Who ate THEM?", in addition to other questions about who ate various other foods, then the answer could be "FRED ate the BEANS", with FRED being the focus, and BEANS being the contrastive topic (Büring, 2003; Jackendoff, 1972, p. 261).

Experiment 2 aimed to test whether this notion of contrastive topic influences the interpretation of the strong form *zij*. In the experiment, I manipulated (1) whether the antecedent of the pronoun is a contrastive topic or not, and (2) the size/specificity of the set of alternatives to that contrastive topic, to see whether these factors influence how likely participants are to choose *zij* over *ze* for that referent. The contrastive topics used in Experiment 2 are also "switched topics" (i.e., they occur after mention of another topical entity, due to the nature of contrast), as discussed below. Given that these contrastive topics by nature involve switching, the aim of this experiment is not to argue for the role of contrast separately from topic-switching, but rather to take a first step towards assessing the validity of existing claims regarding the role of contrast. (We will also consider how the behaviour of *zij* compares with the demonstrative *die*, in light of claims regarding use of *die* in situations, that could be defined as topic-switches.)

The focus here is on contrastive topics, rather than other kinds of contrastive elements (such as contrastive foci), because we are focusing on the information-structural properties of pronouns, which by definition refer to already-mentioned entities.

The labels "contrastive topic" and "focus" are taken from Büring (2003), who uses them as labels for certain kinds of pitch accents (see also Jackendoff, 1972). However, in the present discussion I do not address the role of accent/intonation since I am dealing with written stimuli and thus in this respect I depart from Büring's use of the terms.

## Method

### Participants

Twenty-four adult native Dutch speakers volunteered for participation in the experiment. They were not paid for their participation.

### Materials and procedure

A forced-choice task was used. Participants read brief three-sentence dialogues, and the last sentence of each dialogue involved a choice task between two possible words. In all the target items, the two choices were the weak feminine pronoun *ze* and the strong feminine pronoun *zij* (Example 8). Participants were instructed to choose the word that sounded more natural. Participation took place over the internet, via a webpage where participants could enter their responses.

For the target items, 15 mini-dialogues were constructed. Each dialogue consisted of a context sentence, a question and the critical sentence which answered the question. The critical sentence was the same in all three conditions (e.g., *{Zij/Ze} heeft Claartje gekrabd*, "{Zij/Ze} scratched Claartje"). The context sentence and question were used to manipulate the focus-presupposition structure of the critical sentence.

(8a) Narrow focus:

Ik weet dat Emma Willemijn gekrabd heeft. Maar wie heeft Marjolein gekrabd?

{Zij/Ze} heeft Claartje gekrabd.

*"I know that Emma scratched Willemijn. But who did Marjolein scratch?"*

*"She scratched Claartje".*

(8b) Verb-phrase (VP) focus:

Ik weet dat Emma Willemijn gekieteld heeft. Maar wat heeft Marjolein gedaan?

{Zij/Ze} heeft Claartje gekrabd.

*"I know that Emma tickled Willemijn. But what did Marjolein do?"*

*"She scratched Claartje".*

(8c) Baseline:

Ik weet dat Emma Willemijn gekieteld heeft. En wat nog meer?

{Zij/Ze} heeft Claartje gekrabd.

*"I know that Emma tickled Willemijn. And what else?"*

*"She scratched Claartje".*

In the narrow-focus condition shown in (8a), the question results in the predicate "Marjolein scratched x" being presupposed, and the critical sentence fills in the missing information (she scratched [Claartje]$_F$).[6] In the VP-focus condition shown in (8b), the question triggers the presupposition that Marjolein did something, and the target sentence provides the information regarding what she did (she [scratched Claartje]$_F$). (See e.g., Steedman's (2000) work on intonation for a discussion of similar narrow vs. VP-focus question–answer pairs in English, but without pronouns.) In both (8a) and (8b), the subject of the answer sentence refers to the contrastive topic: in (8a) Marjolein is a member of the set of people who tickled someone, and contrasts with the other members of that set in terms of who she tickled. In (8b) Marjolein is a member of the set of people who did something, and she contrasts with the other members in terms of what she did. Thus, the relevant set is more specific in (8a) than (8b), since in the (8a) answer only the object is new/focused information, whereas in the (8b) answer the entire VP is focused/new information.

The third condition, shown in (8c), was designed as a baseline; the question was constructed not to create any particular focus-presupposition structure for the final sentence, except perhaps the expectation that something else also happened (which could be regarded as "sentence focus", to set it apart from narrow (object) focus and VP focus). In this case, the subject of the answer sentence also does not refer to a contrastive topic; rather, it refers to a highly salient referent. (See also following discussion regarding topic-switching and topic-continuity.)

The characters mentioned in the 15 target items were all female. Each name was only used in one target item, such that no participant saw the same name in more than one item. All verbs were agent-patient verbs (Stevenson et al.,

---

[6] It is worth noting that *wie* in (8a) is case-ambiguous in that the question could be interpreted as an object-wh question or a subject-wh question. The answer, however, disambiguates and signals that the question is asking "Who did Marjolein scratch?" (object-wh question). Thus, when participants read the answer, the intended narrow-focus structure is clear. An interesting question for future work would be to investigate whether languages/constructions where the grammatical role of wh words can be ambiguous (as in Example (8a)) differ from languages/constructions where the grammatical role of wh words is unambiguous, e.g., with respect to the effects of narrow focus.

1994), e.g., *tickle*, *scratch*, *help*, and *push*. The order in which *ze* and *zij* were presented was controlled such that on eight targets, *ze* linearly preceded *zij*, and on seven targets, *zij* linearly preceded *ze*.

Thirty filler items were also constructed. They had the same three-sentence structure as targets, but the choice tasks in the fillers did not contain pronouns. Some fillers involved choosing between pairs of words where there was a (fairly) clear correct answer (e.g., gender marking), but others contained word pairs with no clear answer (e.g., global attachment ambiguities). Three presentation lists were constructed by pseudo-randomly combining the 15 target items with the 30 filler items. Each target item was separated by at least one filler item. Within a presentation list, five targets were in the narrow-focus condition, five were in the VP-focus condition, and five in the baseline condition. Each target item was then rotated through these three conditions, generating three different presentation lists. To control for trial order, reverse order lists were also generated.

## Predictions

If the choice of *zij* over *ze* is sensitive to contrast (such that *zij* is more likely to be used for contrastive topics than *ze*), we predict more *zij* choices than *ze* choices in the narrow-focus and the VP-focus conditions, and more *ze* choices than *zij* choices in the baseline condition.

Moreover, a comparison of the narrow-focus condition with the VP-focus condition will show whether use of *zij* is influenced by the information structure of the rest of the sentence; specifically, whether *zij* prefers narrow-focus constructions in which the set of alternative referents is more constrained (e.g., set of people who scratched someone), or whether it can also be used in contexts where the set of alternatives is less constrained (e.g., set of people who did something). Less constrained sets of alternatives are potentially larger than more constrained sets. At any point in a particular discourse, the number of potentially relevant people who did something is logically at least as big as (or larger than) the number of people who did a specific action.

Theoretical work on the concepts of contrast and focus often makes a distinction between closed sets (with a limited number of members) and sets with a potentially unlimited number of members (for discussion, see Chafe, 1976; Molnár, 2006; Rooth, 1992, and many others), with many researchers regarding contrast as something that only arises with closed sets of alternatives. If a contrastive antecedent is defined as one alternative chosen out of a closed set, then the narrow-focus condition may be perceived as a more canonical example of contrast than the VP-focus condition because the closed nature of set may be easier to perceive/process in the narrow-focus condition, simply by virtue of its being more constrained. Thus, if use of *zij* is

made more likely by the presence of a limited set of alternatives, more *zij* choices are expected in the narrow-focus condition than in the VP-focus condition.

Although this experiment focuses on the effects of contrast, a finding that all three conditions in Experiment 2 show a clear majority of *ze* choices would be evidence in favour of the salience account. In all conditions, the pronoun refers to the preceding subject, so if the choice of *ze* vs. *zij* is sensitive to grammatical role, *ze* should be preferred in all conditions.

## Results

Participants' choices reveal a clear preference for *zij* over *ze* in both the narrow-focus condition (8a) and the VP-focus condition (8b), as well as a clear *ze* preference in the baseline condition (8c). As Figure 4 illustrates, the narrow- and VP-focus conditions pattern very similarly, showing 75% *zij* choices (narrow focus) and 75.8% *zij* choices (VP focus), as compared to 25% *ze* choices (narrow focus), and 24.2% *ze* choices (VP focus). The baseline condition ("sentence focus") shows almost the opposite pattern, 82.5% *ze* choices and 17.5% *zij* choices. A one-sample *t*-test on the proportion of *zij* choices reveals that both the narrow- and VP-focus conditions exhibit a higher-than-chance rate of *zij* choices [narrow focus: $t_1(23) = 3.498$, $p < .005$, $t_2(14) = 6.831$, $p < .001$; VP focus: $t_1(23) = 3.795$, $p = .001$, $t_2(14) = 7.278$, $p < .001$], whereas the baseline condition has a higher-than-chance rate of *ze* choices, $t_1(23) = -4.677$, $p < .001$, $t_2(14) = -6.5$, $p < .001$. Thus, when the subject of the sentence is a contrastive topic, there is a significant preference for the strong form, but when no contrast is present, the weak form is preferred. (As for the other experiments, analyses were conducted on raw as well as arcsine-transformed data. The statistics are reported for the analyses on the raw data; the transformed data resulted in the same significance patterns.)
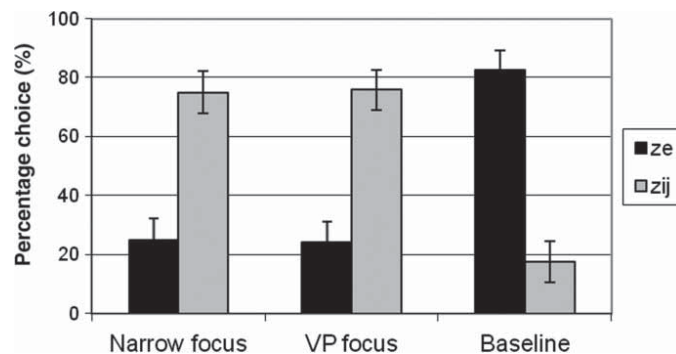


**Figure 4.**   Percentage of *ze* vs. *zij* choices in the forced-choice task in Experiment 2 (error bars show ±1 *SE*).

One-way ANOVAs revealed significant main effects of the focus manipulation on the choice of anaphoric form, $F_1(2, 46) = 39.7$, $p < .001$, $F_2(2, 28) = 64.475$, $p < .001$. Bonferroni-corrected contrasts were used to further investigate the differences between the three conditions. Participants' responses in the narrow- and VP-focus conditions do not differ significantly from each other [proportion of *zij* choices: mean difference = –0.008, 95% CI of difference = $\pm 0.085$ by subjects ($\pm 0.147$ by items)]. In contrast, the baseline condition differs from both the narrow-focus condition and the VP-focus condition [baseline vs. narrow focus, proportion of *zij* choices: mean difference = –0.575, 95% CI = $\pm 0.233$ ($\pm 0.196$); baseline vs. VP focus, proportion of *zij* choices: mean difference = –0.583, 95% CI = $\pm 0.226$ ($\pm 0.132$)]. In other words, the narrow- and VP-focus conditions result in a significantly higher proportion of *zij* choices than the baseline condition, but whether the focus is narrow focus or VP focus does not have an effect.

## Discussion

These results fit well with the idea that use of *zij* is sensitive to the presence of salient alternatives. *Zij* is preferred over *ze* for referring to the contrastive topic in both the narrow-focus and the VP-focus conditions, and *ze* is preferred over *zij* in the baseline condition. However, the size or specificity of the alternative set, at least as manipulated here, has no effect on the likelihood of *zij* choices, as shown by the similarity of the results for the narrow-focus and the VP-focus condition. As a whole, these findings are compatible with the contrast account, and also improve our understanding of what kind of contrast is relevant; what matters is that the referent be a contrastive topic, i.e., stands in a contrastive relation to a salient set of alternatives. The size or specificity of the alternative set does not seem to matter.

When considering the results of Experiment 2, in particular the idea that *zij* is sensitive to the presence of contrasting alternatives, we also need to connect the notion of contrastive topics to topic-switching and topic-continuity. Both of the conditions that involve contrast also involve a switch to a new topic: in the VP- and narrow-focus conditions (Example (8)), after talking about Emma, we've now switched to talking about Marjolein (and end up mentioning a total of four characters). However, in the baseline condition, we continue talking about Emma (and end up mentioning a total of three characters). The topic-switch situation can be regarded as a property closely connected to the notion of contrastive topic. In order for an entity to be a contrastive topic (in the way we have defined it here, building on Büring, 2003), another alternative entity must also exist. Thus, the establishment of a contrastive topic is likely to frequently involve topic-switching, and use of *zij* may well be sensitive both to topic-switches and contrastive topics.

There are at least two reasons why it seems reasonable to assume that contrast—possibly in addition to topic-switching—is relevant. First, native-speaker impressions (e.g., as reflected by the Haeseryn et al. descriptive grammar) suggest that contrast can play a role in guiding the use and interpretation of *zij*. One should also keep in mind that in fact it is the demonstrative *die* that has been claimed to be connected to topic-switching (e.g., Rullmann, 2001).

Second, if reference to a new topic/switched topic is more likely to be accomplished with *zij* than *ze* and contrast does not matter, the results of Experiment 1a are unexpected. Given that each trial was only one sentence long, every entity in subject position was essentially a new/switched topic (there was no topic continuity beyond the test sentence and the continuation provided by the participant). Thus, we might have expected to see *zij* exhibiting a stronger subject preference than *ze* in Experiment 1a. However, a direct comparison of these two forms in that study found no significant differences. In fact, the finding that *ze* (but not *zij*) has a higher-than-chance rate of subject continuations points in the opposite direction. Thus, it does seem that contrast is also playing some role, i.e., that contrastive topics are more likely to be referred to with *zij* than *ze*. Of course, this does not mean that contrast is the *only* factor that guides the interpretation of *zij*, and in fact corpus data suggest that other factors may also play a role. Kaiser (2003) reports a small-scale preliminary corpus study looking at occurrences of *zij* in a novel by Renate Dorrestein called *Het Hemelse Gerecht* (Dorrestein, 1990). The results suggest that *zij* is used contrastively in many cases, but there are also cases where its use seems related to topicality, as indicated by grammatical role and amount of intervening discourse between *zij* and its antecedent. (Though, clearly, more work—both corpus-based and experimental—is needed on this issue.) As a whole the results suggest that *zij* is indeed sensitive to contrastive topics but that contrast is probably not the only factor that matters, an issue we return to in the general discussion.

One might also wonder whether the demonstrative *die* could be used for contrastive referents. A follow-up study to Experiment 2 (a pen-and-paper study using the same test sentences as Experiment 2, with 30 native Dutch speakers, mainly students at the University of Nijmegen) showed that even when participants are given a three-way choice between *zij*, *ze*, and *die*, there is a significant preference ($ps < .005$) for *zij* over *ze* and *die* in both the narrow- and VP-focus conditions. Although *die* is occasionally chosen in both the narrow- and VP-focus conditions (narrow focus 25.3% and VP focus 27.3%), *zij* is clearly preferred (narrow focus 67.3% and VP focus 63.3%). *Ze* is rarely chosen (narrow focus 7.3% and VP focus 9.3%). These results suggest that although *die* seems to be better suited for contrastive/switched topics than *ze*—perhaps due to its stressable/accentable nature—*zij* is clearly the preferred choice for contrastive switched topics. (Unsurprisingly, the

baseline condition, without contrast, revealed a statistically significant preference for *ze* (71.3%), occasional *zij* choices (27.3%), and very few *die* responses (1.3%).)

# GENERAL DISCUSSION

The three experiments presented in this paper aim to contribute to our understanding of the nature of the form-function mapping in anaphoric paradigms, in particular, the referential properties of weak and strong pronominal forms. In three experiments, I probed the interpretation of the Dutch weak feminine pronoun *ze* and the strong feminine pronoun *zij*, as well as the masculine pronoun *hij* and the demonstrative *die*.

Experiment 1a used sentence completion to investigate whether the four forms under investigation differ in their preferences to refer to the subject or object of the preceding sentence. The results showed that while the masculine pronoun *hij* and the feminine weak form *ze* preferred subjects and the demonstrative *die* preferred objects (and was hardly ever interpreted as referring to subjects), the feminine strong form *zij* showed only hints of a possible subject preference. This result left open at least two possible interpretations: (1) the strong form *zij* is used to refer to slightly less salient referents than *ze*, and/or (2) *zij* is used to refer to antecedents (subjects or objects) that can be construed as contrastive. Experiment 1b, an eye-tracking study, took a closer look at effects of grammatical role by minimising the availability of contrastive interpretations. The results showed that, similar to Experiment 1a, *ze* and *hij* preferred subjects and *die* preferred objects. In fact, the results confirm the asymmetry in strength of referential bias that we noticed in Experiment 1a; the object preference of the demonstrative *die* is stronger than the subject preference of the pronoun *hij*. In addition, *zij* patterned like *ze* and exhibited a clear subject preference. In other words, when contrastive interpretations are minimised, *zij* is as likely as *ze* to refer to the preceding subject.

To test whether use of the strong form *zij* is sensitive to contrast, Experiment 2 used a forced-choice task. The results showed that *zij* is strongly preferred over *ze* for referring to contrastive, switched topics. The results are not intended to demonstrate that contrast (and the associated topic-switching) is the *only* factor influencing the interpretation of *zij*, but they do support the idea that referring to a contrastive topic is more likely to be done with the strong from *zij* than the weak form *ze*.

The results of Experiment 2 also help to resolve a possible objection to Experiment 1b. As mentioned above, the sentences in Experiment 1b were spoken with neutral intonation, and the results show that the strong form *zij* patterns very much like the short form *ze* in exhibiting a subject preference.

This raises the question of whether the absence of a contrastive pitch accent/ stress on *zij* is related to the subject preference exhibited by *zij*. Perhaps *zij* cannot refer properly when it is unstressed and thus defaults to the preceding subject, rather than being able to exhibit its actual preference for the object? In other words, the claim would be that stress is what makes it possible for *zij* to refer to lower-salience referents, and the fact that the stimuli were auditorily presented without stress prevented *zij* from attaining its preferred (lower-salience) interpretation. However, the results of Experiment 2 suggest that this is not a valid concern. Even though Experiment 2 used written stimuli, which allow participants to impose their own intonational contours (see also Fodor, 2002), and participants had the option of choosing *ze* over *zij*, the results show that the strong form *zij* can be used to refer to the preceding, highly salient subject, at least in a situation where that subject is a contrastive topic.

Generally speaking, the finding that the strong form *zij* exhibits a sensitivity to the presence of contrastive, switched topics while the weak form *ze*, the masculine pronoun *hij*, and the demonstrative pronoun *die* seem to be more sensitive to referent salience in the absence of contrast, is compatible with the claims of the form-specific approach (e.g., Kaiser & Trueswell, 2008). The observation that the object bias of the demonstrative *die* is stronger than the subject bias of regular pronouns also fits with this approach.

According to the form-specific multiple-constraints framework, anaphoric forms can differ in how sensitive they are to different antecedent properties. The form-specific approach resembles existing multiple-constraint approaches to reference resolution (e.g., Ariel, 1990; Arnold, 1998) in assuming that anaphor resolution is not determined by a single constraint but rather is the result of the interaction of multiple constraints. Furthermore, this approach allows for the multiple constraints that play a role in the interpretation of referential forms to not necessarily carry the same weight for all referential forms. The approach was originally formulated on the basis of data from Finnish showing that pronouns and demonstratives, both of which can be used to refer to human antecedents mentioned in a preceding clause, do not show the same level of sensitivity to the antecedent's syntactic role and linear position (Kaiser, 2003, 2005a). The asymmetrical sensitivities indicate that not all anaphoric forms are equally sensitive to the same antecedent properties. Further evidence for form-specific effects comes from cross-clausal data from Estonian (Kaiser & Hiietam, 2004; Kaiser & Vihman, 2006) and English (Brown-Schmidt et al., 2005), as well as the interpretation of pronouns and reflexives in picture-NP constructions in English (Kaiser, Runner, Sussman, & Tanenhaus, 2009).

Whereas these previous studies looked at the interpretation of referential forms that are lexically clearly distinct (pronouns vs. demonstratives and

pronouns vs. reflexives), the research reported here suggests that form-specific effects also occur with two related forms such as a weak pronoun (*ze*) and a strong pronoun (*zij*). In addition, the asymmetrical referential biases of *hij* vs. *die* suggest that even if two forms seem to be sensitive to the same factor, their level of sensitivity is not always equal. *Die* has a stronger preference for objects (lower salience), whereas *hij*'s preference for subjects is not as strong (it is occasionally used to refer to objects as well).

In addition to contributing to our understanding of the form-specific effects present in anaphoric paradigms, this research has implications for the nature of the mapping between particular referential forms and their discourse functions. There are at least two ways in which one could think about the connection between the strong form and contrast. The first would be to view *zij* as intrinsically related to the notion of contrastiveness—for example, one could say that use of *zij* is only licensed when the antecedent is marked as being [+ contrastive]. Another way of thinking about the connection between *zij* and contrast is a more Gricean approach. The general idea that Gricean implicatures guide anaphor resolution is not new, and has been suggested by researchers such as Huang (1991, 2000) and Levinson (1987), much of whose work focused on Binding Theory and the interpretation of pronouns and reflexives. For example, Huang (2000, p. 221) argues that use of a pronoun in a sentence such as "Mozart admired him" triggers a noncoreferential interpretation (i.e., him ≠ Mozart) via implicature. Because the reflexive "himself" could also occur in this syntactic position ("Mozart admired himself" is grammatical) but the speaker did not use a reflexive form, this gives rise to the implicature that with the pronoun coreference is not intended.

Applying a broadly Gricean approach to the weak vs. strong distinction, one could argue that using *zij* in a syntactic position where the morphologically more attenuated *ze* could also be used generates the implicature that *zij* is being used for some reason, such as to mark contrast.[7] However, there are situations where *ze* cannot be used due to purely syntactic reasons (such as coordination, see Footnote 2), and in such contexts use of *zij* does not trigger this implicature (because there is no possibility of choosing *ze*). (In his research on the interpretation of null and overt pronouns, Montalbetti

---

[7] Thanks to Jeanette Gundel for discussion regarding this. A reviewer asked whether this approach assumes that *ze* is the default. While such an assumption would not be incompatible with this, I do not think that it has to be made. To see why, let us consider scalar implicatures. Normally, if two forms can be arranged on a scale, e.g., <warm, hot >, use of the weaker term ("The weather is warm today") implies that the stronger term does not hold. Thus, use of one form can trigger an implicature (as long as the forms are on a scale) without the need for a "default" form.

(1984) also discusses the differences between situations where both forms are possible/available and situations where only one form is grammatically possible; see also Huang (2000) for relevant discussion.)

This Gricean approach has the advantage of locating the contrastive effect not in the morphological form *zij* itself, but rather in the choice of *zij* over *ze*. In other words, the form *zij* is not inherently connected to contrast; rather, it is the choice of *zij* in a context where *ze* could also be used that triggers the implicature. This appears to be a desirable result, given that (1) there are indeed certain syntactic configurations in which the weak form cannot occur (e.g., coordination) and where use of the strong form does not trigger contrast effects, and (2) the results of Experiment 2 show that the choice of *ze* vs. *zij* is not categorical. *Zij* is chosen occasionally (17.5%) in the baseline noncontrastive condition, a finding which seems more compatible with a Gricean pragmatic account than an intrinsic [+contrast]-marking account.

Another advantage of this approach is that it allows for the possibility that the implicature triggered by use of *zij* is underspecified. Let us assume that use of *zij* in a context where *ze* would also be possible carries some extra meaning/triggers an implicature. How specific is that extra information? One possibility is that use of *zij* signals specifically that the antecedent is contrastive, but another possibility is that the implicature be underspecified and context-dependent.

It seems reasonable to allow some level of variability or context-dependence, especially in light of corpus data suggesting that not all uses of *zij* are related to the presence of contrastive topics (e.g., earlier discussion of Kaiser's (2003) preliminary corpus study). The results of Experiment 1b also suggest that *zij* is not inherently connected specifically to contrast. Recall that the stimuli in Experiment 1b were designed to minimise any potentially contrastive interpretations. Thus, if the strong form *zij* requires a contrastive antecedent and none is available, one might expect that the sentence will be perceived as infelicitous. However, participants' comments did not indicate such infelicity. Thus, we find that *zij* is clearly preferred over *ze* for reference to contrastive topics, but this contrast use of *zij* is potentially a subcase of a more general, contextually specified "extra-meaning" implicature that is triggered by the use of *zij*.

As a whole, the results of the three experiments presented in this paper suggest that models of reference resolution should be flexible enough to incorporate effects of contrast and topic-switching as well as salience, and to allow for form-specific effects. Moreover, the behaviour of the strong form *zij* suggests that its referential properties may be attributable to implicatures triggered by use of *zij* in a context where *ze* could also be used. The referential behaviour of the four forms investigated here fits well with the claim of the form-specific multiple-constraints approach, according to which

referential forms can differ in what kinds of information they are most sensitive to.

# REFERENCES

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory, 21*, 435–483.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*, 247–264.

Ariel, M. (1990). *Accessing NP antecedents.* London, UK: Routledge/Croom Helm.

Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text representation, linguistic and psycholinguistic aspects* (pp. 29–87). Amsterdam/Philadelphia, PA: John Benjamins.

Arnold, J. (1998). *Reference form and discourse patterns* (Doctoral dissertation). Stanford University, Stanford, CA.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition, 76*, B13–B26.

Bosch, P., Rozario, T., & Zhao, Y. (2003). Demonstrative pronouns and personal pronouns: German *der* vs. *er*. In *Proceedings of the EACL2003. Workshop on the computational treatment of anaphora* (pp. 61–68). Budapest, Hungary: Association for Computational Linguistics.

Brennan, S., Friedman, M., & Pollard, C. (1987). A centering approach to pronouns. *ACL, 25*, 155–162.

Brown-Schmidt, S., Byron, D. K., & Tanenhaus, M. K. (2005). Beyond salience: Interpretation of personal and demonstrative pronouns. *Journal of Memory and Language, 53*, 292–313.

Büring, D. (2003). On D-trees, beans and B-accents. *Linguistics and Philosophy, 26*, 511–545.

Cardinaletti, A., & Starke, M. (1996). Deficient pronouns: A view from Germanic. In H. Thráinsson, S. D. Epstein, & S. Peter (Eds.), *Studies in comparative Germanic syntax II* (pp. 21–65). Amsterdam: Kluwer Academic Press.

Cardinaletti, A., & Starke, M. (1999). The typology of structural deficiency: A case study of the three classes of pronouns. In H. van Riemsdijk (Ed.), *Clitics in the languages of Europe* (pp. 145–233). Berlin/New York: Mouton de Gruyter.

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

Chambers, C. G., & Smyth, R. (1998). Structural parallelism and discourse coherence: A test of centering theory. *Journal of Memory and Language, 3*, 593–608.

Comrie, B. (1997). Pragmatic binding: Demonstratives as anaphors in Dutch. In M. L. Juge & J. L. Moxley (Eds.), *Proceedings of the twenty-third annual meeting of the Berkeley Linguistics Society* (pp. 50–61). Berkeley, CA: Berkeley Linguistics Society.

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

Crawley, R., & Stevenson, R. (1990). Reference in single sentences and in texts. *Journal of Psycholinguistic Research, 19*(3), 191–210.

de Hoop, H. (2003). On the interpretation of stressed pronouns. In M. Weisgerber (Ed.), *Proceedings of Sinn und Bedeutung 7* (pp. 159–172). Konstanz, Germany: University of Konstanz.

Donaldson, B. (1997). *Dutch – a comprehensive grammar*. New York: Routledge.

Dorrestein, R. (1990). *Het Hemelse Gerecht*. Amsterdam, The Netherlands: Pandora.

É. Kiss, K. (1998). Informational focus vs. identification focus. *Language*, 74, 245–273.

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. In M. Hirotani (Ed.), *Proceedings of the North East Linguistic Society (NELS 32)* (pp. 113–132). Amherst, MA: GSLA, University of Massachusetts.

Geerts, G., Haeseryn, W., de Rooij, J., & van den Toorn, M. C. (Eds.). (1984). *Algemene Nederlandse spraakkunst* [General Dutch Grammar]. Groningen: Wolters-Noordhoff.

Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274–307.

Haeseryn, W., Romijn, K., Geerts, G., De Rooij, J., & Van den Toorn, M. C. (Eds.), (1997). *Algemene Nederlandse Spraakkunst* [General Dutch Grammar] (Vol. 1). Groningen: Martinus Nijhoff.

Hallett, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 10.1–10.112). New York: Wiley.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). The intermodal preferential looking paradigm: A window onto emerging language comprehension. In D. McDaniel & C. McKee (Eds.), *Methods for assessing children's syntax: Language, speech, and communication* (pp. 105–124). Cambridge: MIT Press.

Huang, Y. (1991). A neo-Gricean pragmatic theory of anaphora. *Journal of Linguistics*, 27, 301–335.

Huang, Y. (2000). *Anaphora – a cross-linguistic study*. Oxford, UK: Oxford University Press.

Jackendoff, R. (1972). *Semantics in generative grammar*. Cambridge, MA: MIT Press.

Kaiser, E. (2003). *The quest for a referent: A crosslinguistic look at reference resolution* (Doctoral dissertation). Philadelphia: University of Pennsylvania.

Kaiser, E. (2005a). When salience isn't enough: Pronouns, demonstratives and the quest for an antecedent. In R. Laury (Ed.), *Minimal reference in Finnic: The use and interpretation of pronouns and zero in Finnish and Estonian discourse* (pp. 135–162). Helsinki: Suomalaisen Kirjallisuuden Seura.

Kaiser, E. (2005b). What the referential properties of the Estonian pronoun *tema* can tell us about the role of contrast in pronoun interpretation. In *Proceedings of the 28th Penn Linguistics Colloquium. Penn Working Papers in Linguistics* (Vol. 11, pp. 113–127). Philadelphia, PA: University of Pennsylvania.

Kaiser, E. (2010). Effects of contrast on referential form: Investigating the distinction between strong and weak pronouns. *Discourse Processes, 47*, 480–509.

Kaiser, E., & Hiietam, K. (2004). A typological comparison of third person pronouns in Finnish and Estonian. In A. Dahl, K. Bentzen, & P. Svenonius (Eds.), *Proceedings of the workshop on Generative Approaches to Finnic Languages* [Working papers series, *Nordlyd, 31*(4), 654–667]. Tromso, Norway.

Kaiser, E., & Trueswell, J. C. (2004a). The referential properties of Dutch pronouns and demonstratives: Is salience enough? In C. Meier & M. Weisgerber (Eds.), *Proceedings of the conference sub8 Sinn und Bedeutung* (Working paper No. 177, pp. 137–149). Constance: FB Sprachwissenschaft, University of Constance.

Kaiser, E., & Trueswell, J. C. (2004b). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94(2), 113–147.

Kaiser, E., & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference. *Language and Cognitive Processes*, *23*(5), 709–748.

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*, 55–80.

Kaiser, E., & Vihman, V. -A. (2006). *On the referential properties of Estonian pronouns and demonstratives.* Paper presented at the 22nd Scandinavian Conference of Linguistics, Aalborg University, Denmark.

Kameyama, M. (1999). Stressed and unstressed pronouns: Complementary preferences. In P. Bosch & R. van der Sandt (Eds.), *Focus: Linguistic, cognitive, and computational perspectives* (pp. 306–321). Cambridge, UK: Cambridge University Press.

Kehler, A. (2002). *Coherence, reference, and the theory of grammar.* Stanford, CA: CSLI.

Kehler, A., Kertz, L., Rohde, H., & Elman, J. (2008). Coherence and coreference revisited. *Journal of Semantics*, *25*, 1–44.

Kibrik, A. (1996). Anaphora in Russian narrative discourse: A cognitive calculative account. In B. Fox (Ed.), *Studies in anaphora* (pp. 255–304). Amsterdam: Benjamins.

Krasavina, O., & Chiarcos, C. (2007). *Aspects of topicality in the use of demonstratives expressions in German and Russian.* Paper presented at the 29th annual meeting of the Linguistics Association of Germany, Siegen, Germany.

Levinson, S. C. (1987). Pragmatics and the grammar of anaphora. *Journal of Linguistics*, *23*, 379–434.

Molnár, V. (2006). On different kinds of contrast. In V. Molnar & S. Winkler (Eds.), *The architecture of focus* (pp. 197–233). Berlin/New York: Mouton de Gruyter.

Montalbetti, M. (1984). *After binding* (Doctoral dissertation). MIT, Cambridge, MA.

Nakatani, C. (1993). Accenting on pronouns and proper names in spontaneous narrative. In D. House & P. Touati (Eds.), *Proceedings of the European speech communication association workshop on prosody* (pp. 164–167). Lund: ESCA.

Pajusalu, R. (1995). Pronominit *see, tema* ja *ta* viron puhekielessä. *Sananjalka*, *37*, 81–93.

Pajusalu, R. (1997). Eesti pronoomeneid I. Ühiskeele *see, too* ja *tema/ta*. *Keel ja Kirjandus* [Estonian Pronouns I. *See, too, tema/ta* in Common Estonian], pp. 24–30, 106–115.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology.* Cambridge, MA: Bradford Books, MIT Press.

Rigau, G. (1986). Some remarks on the nature of strong pronouns in null subject languages. In I. Bordelois, H. Contreras, & K. Zagona (Eds.), *Generative studies in Spanish syntax* (pp. 143–163). Dordrecht: Foris.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, *1*, 75–116.

Rullmann, H. (2001). *Dutch demonstrative pronouns in discourse.* Paper presented at the 7th Germanic Linguistics Annual Conference, Banff, AB.

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, *23*, 197–229.

Snedeker, J., Thorpe, K., & Trueswell, J. (2001). On choosing the parse with the scene: The role of visual context and verb bias in ambiguity resolution. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 964–969). Mahwah, NJ: Erlbaum.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, *49*, 238–299.

Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language*, *58*, 574–608.

Steedman, M. (2000). Information structure and the syntax-phonology interface. *Linguistic Inquiry*, *31*, 649–689.

Stevenson, R. J., Crawley, R. J., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, *9*, 519–548.

Tanenhaus, M., & Trueswell, J. (2006). Eye movements and spoken language comprehension. In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 863–900). New York: Academic Press.

Tanenhaus, M. K., Spivey-Knowlton, M. K., Eberhard, K. M., & Sedivy, J. E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 632–634.

Vallduví, E., & Vilkuna, M. (1998). On rheme and kontrast. In P. Culicover & L. McNally (Eds.), *The limits of syntax, syntax and semantics* (Vol. 29, pp. 79–108)). New York: Academic Press.

Yang, N., & van Bergen, G. (2007). Scrambled objects and case marking in Mandarin Chinese. *Lingua*, *117*, 1617–1635.

Zimmermann, M. (2007). Contrastive focus. In C. Féry, G. Fanselow, & M. Krifka (Eds.), *Interdisciplinary studies on information structure: Vol. 6. The notion of information structure* (pp. 147–159) [Working Papers of the SFB 632, University of Potsdam]. Potsdam, Germany: University of Potsdam.

Zwart, J. W. (1993). Notes on clitics in Dutch. In L. Hellan (Ed.), *Clitics in Germanic and Slavic* (Vol. 4, pp. 119–155). [Working Papers from the ESF Programme on Language Typology, Theme Group 8]. Tillburg University, Netherlands.