UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

The effect of meaning-related cues on pronoun resolution in Dutch

Permalink

https://escholarship.org/uc/item/8fz4c880

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Hoek, Jet

Wilke, Hans

Publication Date

2024

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

The effect of meaning-related cues on pronoun resolution in Dutch

Jet Hoek (jet.hoek@ru.nl)
Centre for Language Studies
Radboud University Nijmegen

Hans A. Wilke (h.a.wilke@rug.nl)
Center for Language and Cognition
University of Groningen

Abstract

Pronoun interpretation seems to be driven by structural factors, but also by factors related to meaning. In a forced-choice pronoun interpretation experiment, we compare the impact of the next-mention bias associated with transfer-of-possession-verbs on the interpretation of three Dutch pronominal forms that differ in the strength of their structural biases: reduced personal pronoun *ze* 'she$_{reduced}$', full personal pronoun *zij* 'she$_{full}$', and demonstrative pronoun *die* 'that'. In addition to replicating the common Goal-bias associated with transfer-of-possession verbs, results show significant differences in the proportion of pronoun resolved to the preceding subject between all three pronominal forms. However, the effect of the next-mention manipulation did not differ between pronominal forms. These findings are in line with a model of pronoun interpretation that combines structural and meaning-related factors, and present particularly strong evidence against models that posit that pronoun interpretation is the mirror image of pronoun production.

Keywords: pronoun interpretation; referring expressions; Dutch; next-mention biases

Introduction

Pronoun interpretation seems to be driven by structural factors (e.g., subjecthood, parallelism), but also by factors related to meaning. An early study that shows the interplay of these two types of factors is the story continuation experiment by Stevenson, Crawley, and Kleinman (1994), in which participants were presented with prompts such as the ones in (1) and (2).

(1) John$_{source}$ passed the comic to Bill$_{goal}$. (He) ...

(2) John$_{goal}$ seized the comic from Bill$_{source}$. (He) ...

Results show an effect of thematic role on pronoun interpretation: in the Source-Goal condition (1), where the subject is the Source of the transfer-of-possession verb and the Goal appears as a prepositional object, the pronoun was interpreted as referring to the preceding subject (*John* in (1-2)) less often than in the Goal-Source condition (2), where the subject is the Goal. The next-mention bias of transfer-of-possession verbs toward the Goal referent has been replicated many times since (e.g., Arnold, 2001; Rohde, Kehler, & Elman, 2006; Rosa & Arnold, 2017). Comparing the prompts with and without a pronoun, however, shows that the free prompt condition (i.e., without a pronoun) results in more rementions of the Goal as the focus of the continuation than the pronoun prompt condition. This suggests that the next-mention bias does not completely override the subject bias, but that the interpretation process combines both factors. In addition, it suggests that pronoun production and pronoun interpretation are not mirror images of each other.

Both these ideas are captured by the Bayesian model of pronoun interpretation (Kehler, Kertz, Rohde, & Elman, 2008), which posits that pronoun resolution is the result of comprehenders combining their expectations about a referent's remention rates with their estimation that that referent will be rementioned by means of a pronoun. The Bayesian model has been found to make good predictions for English (e.g., Cheng & Almor, 2019; Hoek, Kehler, & Rohde, 2021; Kehler & Rohde, 2019; Rohde & Kehler, 2014), but also for other languages, such as Mandarin Chinese (Zhan, Levy, and Kehler, 2020, though see Lam and Hwang, 2021) and Catalan (Mayol, 2018). Patterson, Schumacher, Nicenboim, Hagen, and Kehler (2022) tested the model on German, which has a more complex pronominal system than English. They found that the Bayesian model was not only able to accurately capture interpretation patterns of German personal pronouns, but also those of the demonstrative pronoun *dieser* 'this', which can be used to refer to human referents. Unlike the personal pronouns that were tested, *dieser* has an interpretation bias toward the non-subject; in addition, it allows for less interpretation variability (Patil, Bosch, & Hinterwimmer, 2020). Patterson et al.'s (2022) findings thus suggest that meaning-related cues can even sway fairly strict structural constraints.

The current study tests how structural factors and meaning-related factors affect pronoun interpretation in an even more complex pronominal paradigm. Similar to German *dieser*, the demonstrative pronoun *die* 'that' in Dutch can be used to refer to human referents. In contexts with multiple referents, *die* is also biased away from the subject (or biased toward the most recent referent, which in the vast majority of cases is not the subject), but this bias appears to be even stronger than the non-subject bias of German *dieser*. In their German pronoun-selection experiment, Patil et al. (2020) find that participants opted to use *dieser* to refer to the preceding subject 19.2% of the time (when the alternative options were either a personal pronoun or 'neither'). In their story continuation experiment, Patterson et al. (2022) found that in the pronoun prompt condition, *dieser* was resolved to the preceding subject on average about 10% of the time. In addition, they found that when the next-mention bias was directed at the subject, subject resolution rates of *dieser* went up to around 30%; when the next-

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

mention bias was shifted away from the subject, subject resolution rates of *dieser* dropped to around zero. By contrast, in a Dutch story continuation experiment using free prompts (i.e., without a pronoun) following sentences containing two human referents (the non-subject occurring later in the sentence than the subject), Vogels (2019) finds (virtually) no uses of *die* to refer to the subject of the preceding sentence, even when the next-mention bias was directed at the subject. Similarly, in a Dutch story continuation experiment using pronoun prompts, Kaiser (2011) finds only a single case (out of 160 cases) of *die* being used to refer to the subject of preceding sentences that contained a sentence-early subject and another human referent mentioned later in the sentence. By studying Dutch *die*, we aim to investigate whether meaning-related factors can 'override' even very strong structural biases.

In addition, many Dutch personal pronouns have a full, or 'strong', form (e.g., *zij* 'She$_{full}$') and a reduced, or 'weak', form (e.g., *ze* 'She$_{reduced}$), which appear to differ in their structural bias (Kaiser, 2011). In her story continuation experiment, Kaiser (2011) found a subject bias for *ze* (found to refer to the subject in 63.1% of the cases), but no subject bias for *zij* (found to refer to the subject in 50% of the cases).[1] The follow-up visual world eye-tracking study, in which *zij* was not phonologically stressed, showed a subject bias for both *ze* and *zij*. This suggests that the difference in interpretation bias between *ze* and *zij* found in the written story continuation experiment is due to the potential contrastive reading of *zij*, which results in a non-subject interpretation – *ze*, being a reduced form, cannot be emphasized and (therefore) cannot be used contrastively. By also comparing full and reduced Dutch pronominal forms, we test whether meaning-related factors have a bigger effect on pronoun resolution when the pronoun allows for more interpretation variation.

We thus compare the impact of meaning-related cues (specifically: the Goal-bias associated with transfer-of-possession-verbs) on the interpretation of three Dutch pronominal forms that differ in the strength of their structural biases. Since the structural bias of the demonstrative pronoun *die* appears to be *very* strong, it might be expected that the next-mention bias impacts the interpretation of this pronominal form less than it impacts the interpretation of personal pronouns *ze* and *zij*. Similarly, since full *zij*, at least in the written mode, allows for more interpretation variability than reduced *ze*, the next-mention might be expected to have a larger effect on the interpretation of *zij* than on *ze*. On the other hand, Patterson et al.'s (2022) study on German shows that next-mention expectations can sway even fairly strict structural biases, which could suggest that there is no difference in the impact of the Goal-bias on resolution rates of *ze* and *zij*.

We test these predictions in a forced-choice pronoun interpretation experiment. Unlike in many other coreference studies manipulating next-mention biases, the critical manipulation in our experimental items is embedded in a larger context. Several recent pronoun studies have shown that the inclusion of context in experimental prompts can impact results (production: Demberg, Kravtchenko, & Loy, 2023; Rosa & Arnold, 2017; interpretation: Hoek et al., 2021), with meaning-related factors having smaller, inconsistent, or insignificant effects in studies using decontextualized prompts.

Method

In a forced-choice experiment, we measured the rate at which comprehenders interpreted an ambiguous pronoun as referring to the subject of the preceding sentence. We manipulated the next-mention bias (Source-Goal vs. Goal-Source) and the type of pronoun (*Ze* 'She$_{reduced}$' vs. *Zij* 'She$_{full}$', vs. *Die* 'That'). We thus contrast six (3x2) conditions. If the effect of next-mention bias differs depending on the strength of a pronoun's structural bias, we should find a significant interaction effect between Pronoun and Next-mention. The experiment was approved by the Ethics Committee of Radboud University (2021-1680).

All experimental materials and data, as well as the preregistered hypotheses and analysis plan are available on the Open Science Framework page: `http://doi.org/10.17605/OSF.IO/Z2MSU`.

Participants

We recruited 87 participants through the Radboud University SONA online recruiting platform. We removed those who did not meet the preregistered language inclusion criteria (speaking Dutch growing up as a majority household language and Dutch being the majority language in current daily life – 16 participants) and those whose accuracy on the catch fillers, see Fillers, was not significantly above chance (12 participants), leaving 59 participants for analysis. These were self-reported native Dutch speakers between the ages of 17-25 (mean=19.2, SD=1.8) living in the Netherlands. Participants provided informed consent and were compensated with course credits.

Materials

Target items The experiment had 48 target items in six (3x2) conditions. The target items consisted of five sentences that together formed a short narrative passage. The first three sentences set up a context and introduced the two referents involved in the scenario. The key manipulations appeared in the fourth sentence (Next-mention bias) and in the fifth sentence (Pronoun). A full example item is given in Table (1).

The first manipulation – in sentence four – is the Next-mention bias. The manipulation used transfer-of-possession verbs, in which an object is transferred between two human referents, from the Source to the Goal. As was discussed in the Introduction, these verbs have been shown to bias next-mention expectations toward the Goal. The verbs appeared in either a Source-Goal configuration (e.g., *geven* 'to give'), in

[1] Note that even though Kaiser (2011) investigated pronoun resolution rates of Dutch *ze*, *zij*, and *die*, she did not directly compare *die* to *ze* and *zij*. Instead, she compared *ze* to *zij*, and *die* to the full masculine pronoun *hij* 'he'. In addition, there was no manipulation of meaning-related factors in these experiments.

Table 1: Example item

CONTEXT	*De koningin en de barones zijn al jaren vriendinnen. De koningin vierde vorige week haar verjaardag. De barones had een bijzonder cadeau meegenomen.* 'The queen and the baroness have been friends for years. The queen celebrated her birthday last week. The baroness had brought a special gift.'
THEMATIC ROLE ORDER	
Source-Goal	*De barones gaf een bronzen kandelaar aan de koningin.* 'The baroness gave a bronze candelabra to the queen.'
Goal-Source	*De koningin kreeg een bronzen kandelaar van de barones.* 'The queen received a bronze candelabra from the baroness.'
PRONOUN *red. / full / dem.*	[*Ze / Zij / Die*] *heeft vervolgens...* ['She$_{reduced}$ / She$_{full}$ / That] then went and...'
QUESTION	*Wie is Ze / Zij / Die?* 'Who is She$_{reduced}$ / She$_{full}$ / That?'
FORCED CHOICE	*De koningin / De barones* 'The queen / The baroness'

which case the subject takes on the thematic role of Source, or Goal-Source (e.g., *krijgen* 'to receive'), in which case the subject is the Goal. Within each item, we used a Source-Goal verb and a Goal-Source verb that closely resembled each other, such that the crucial difference between conditions reflects the order of thematic roles while keeping the resulting meaning/unfolding of events constant between conditions (e.g., *geven* 'to give' and *krijgen* 'to receive'). We constructed six verb pairs in total, all of which appeared in eight items. For each item, the thematic role was held constant for each specific referent: in the item in Table 1, the queen was for instance the Goal in both Next-mention conditions. Which referent takes on which role was counterbalanced across items such that both the referent mentioned first in the context and the referent mentioned second in the context were the Source 50% of the time and the Goal 50% of the time.

The second manipulation – in sentence five – is the type of Pronoun. This (unfinished) sentence was headed by one of three pronouns – *Ze* 'She$_{reduced}$','*Zij* 'She$_{full}$', or *Die* 'That' – followed by a predicate that contained a singular auxiliary verb (specifying that the pronoun had to be interpreted as singular – all tested pronominal forms can also be used to refer to plural referents) and a temporal connective (reinforcing the next-mention bias associated with tranfer-of-possession verbs, see e.g., Rohde et al., 2006), but was otherwise kept open so as not to provide any potentially disambiguating or biasing information, for example: *heeft vervolgens...* 'then went and...' (lit.: 'has then...'). Participants were then asked to answer the question which referent (forced choice between the two referents, i.e., *De koningin* 'the queen' and *De barones* 'the baroness' for the item in Table 1) they believe to be the antecedent for the pronoun in the final sentence: *Wie is Ze/Zij/Die?* 'Who is She$_{reduced}$/She$_{full}$/That?'.

For the pronoun manipulation, we only used feminine personal pronouns. While Dutch has a reduced form of the masculine personal pronoun (*'ie*), it has several big restrictions (see also Kaiser, 2011): it cannot be used sentence-initially, but only as a clitic; in addition, it mainly occurs in spoken language. The weak form of the Dutch feminine pronoun (*ze*) does not have these restrictions (although there are some other syntactic restrictions irrelevant to the current study, see Kaiser, 2011).

All items in this experiment (target and filler) were designed with a fantasy world in mind inhabited by women only. We chose to construct this fantasy world to create a narrative that sounded as natural as possible with cohesion throughout, rather than presenting participants with seemingly unrelated or isolated sentences. Because of the pronoun manipulation, we opted for an all-female cast, with characters explicitly marked as feminine (e.g., *prinses* 'princess', *raadsvrouw* 'councilwoman'). While *ze* and *zij* are specified for gender, the demonstrative *die* is not. By only including women as referents, we eliminated any potential effect of gender on the interpretation of the pronouns. In total, 14 unique referents were included in the items. This was the minimum amount required to ensure that every combination of referents happened exactly once and no more than that.

The first three sentences of each item were the same across conditions. The function of these sentences was to provide

context and to introduce the two referents, while holding the topicality of the two referents (which can influence pronoun interpretation, see e.g., Cowles & Ferreira, 2012) as constant as possible. The first context sentence introduces both referents as the conjoined subject of this sentence. The two referents each individually appeared as the subject of the second or third context sentence.

The target items were distributed across six lists in a Latin Square design such that all participants saw eight target items in each condition and each participant saw each item only once. The distribution of the target items and fillers and the order in which participants saw these was fully randomized.

Fillers The experiment also had 48 fillers. 24 fillers were items from a different experiment probing the accessibility of different possible referents – which were presented in a sentence containing an appositive relative clause – for the pronoun *het* 'It'. For this experiment, the forced-choice question that accompanied the items was *Wat is Het?* 'What is It?'.

In addition, there were 24 other fillers that can be subdivided into three subcategories, all containing 8 items. Fillers in the first subcategory were catch fillers that had only one possible correct answer to the question *Wie is Ze / Zij?* 'Who is She$_{reduced}$ / She$_{full}$?', as in (3). The final pre-question sentence, as well as the preceding context, provided only one possible referent for the pronoun *Ze* or *Zij*. The incorrect answer options were different referents not mentioned in the item (though they were mentioned in other items and were compatible with the overarching story world). Fillers in the second subcategory were catch fillers that had only one possible correct answer to the question *Wat is Het?* 'What is It?', as in (4). The final pre-question sentence contained a transfer-of-possession verb. The only possible referent for *Het* 'It' was the object of transfer. The incorrect answer option was a random referent that was not mentioned in the item. Fillers in the third subcategory were items that had only one likely answer to the question *Wie is Ze / Zij?* 'Who is She$_{reduced}$ / She$_{full}$?', as in (5). The final pre-question sentence provided only one possible referent for the pronoun *Ze* or *Zij*; this referent was introduced in the first sentence and rementioned, either by full noun phrase or by a pronoun, in every other sentence in the prompt. However, in the very first context sentence, which contained an appositive relative clause, another referent was mentioned.

Answers to the first two subcategories were used as an exclusion criterion. Participants whose performance on these was not significantly above chance (at least 75% of questions answered correctly) were excluded from the analysis. After exclusion, average performance on these questions was 97% answered correctly.

(3) *De prinses organiseerde vorige zomer een groot atletiektoernooi. Tientallen onderdelen kwamen aan bod. De prinses heeft zelf ook meegedaan. Ze won de zilveren medaille op de hink-stap-sprong. Ze heeft toen...*
Q: *Wie is Zij?* A: *De prinses / De boerin*
'The princess organized a major athletics tournament last summer. Dozens of components were featured. The princess herself also participated. She$_{reduced}$ won the silver medal in the triple jump. She$_{reduced}$ then...'
Q:'Who is She$_{reduced}$?' A: 'The princess / The farmer$_{fem}$'

(4) *De jageres werd vanochtend opgeroepen door de prinses. De jageres had geen idee waarom. De prinses vertelde dat ze last had van herten in de tuin. De jageres leende een kruisboog aan de prinses. Het was een...*
Q: *Wat is Het?* A: *een kruisboog / een schroevendraaier*
'The huntress was summoned by the princess this morning. The huntress had no idea why. The princess said that she was bothered by deer in her garden. The huntress lent a crossbow to the princess. It was a...'
Q: 'What is It?' A: 'A crossbow / A screwdriver'

(5) *De alchemiste had een afspraak met de koningin, die handelt in chemische stoffen. De koningin had net haar laatstje flesje ethanol verkocht. Zij had nog wel ammoniak en azijnzuur. Zij heeft toen...*
Q: *Wie is Ze?* A: *De naaister / De zangeres*
'The alchemist had an appointment with the queen, who deals in chemicals. The queen had just sold her last bottle of ethanol. She still had ammonia and acetic acid. She$_{full}$ then..'
Q: 'Who is She$_{full}$?' A: 'The queen / The alchemist$_{fem}$'

Procedure

The experiment was deployed on the PennController for Internet Based Experiments (PCIbex) platform (Zehr & Schwarz, 2018). Participants carried out the experiment remotely on their own computers via a link distributed through the Radboud University SONA online recruiting platform. They were instructed to read the short passages and to select (by mouse-click) the referent they thought the ambiguous pronoun most likely referred to. The passage, the forced-choice question and the two answer options were all visible at the same time. When an answer was selected, the next item would appear after a short delay. It was not possible to return to previous passages. Participants were instructed not to overthink their answers. In addition, it was specified that most items would not have a clear answer, but that there would be a couple of attention check items that did have a correct answer, which would be easy to get right when taking the task seriously. After completing all passages, participants filled out a short demographic questionnaire, after which they were thanked for their participation and sent back to SONA. The experiment lasted approximately 30 minutes.

Results

Figure 1 shows the distribution of pronoun interpretations for each pronominal form, per Next-mention condition. The figure shows that *ze* was most often resolved to the subject and *die* least often. In addition, it shows a clear effect of the next-mention manipulation, with fewer pronouns resolved to the

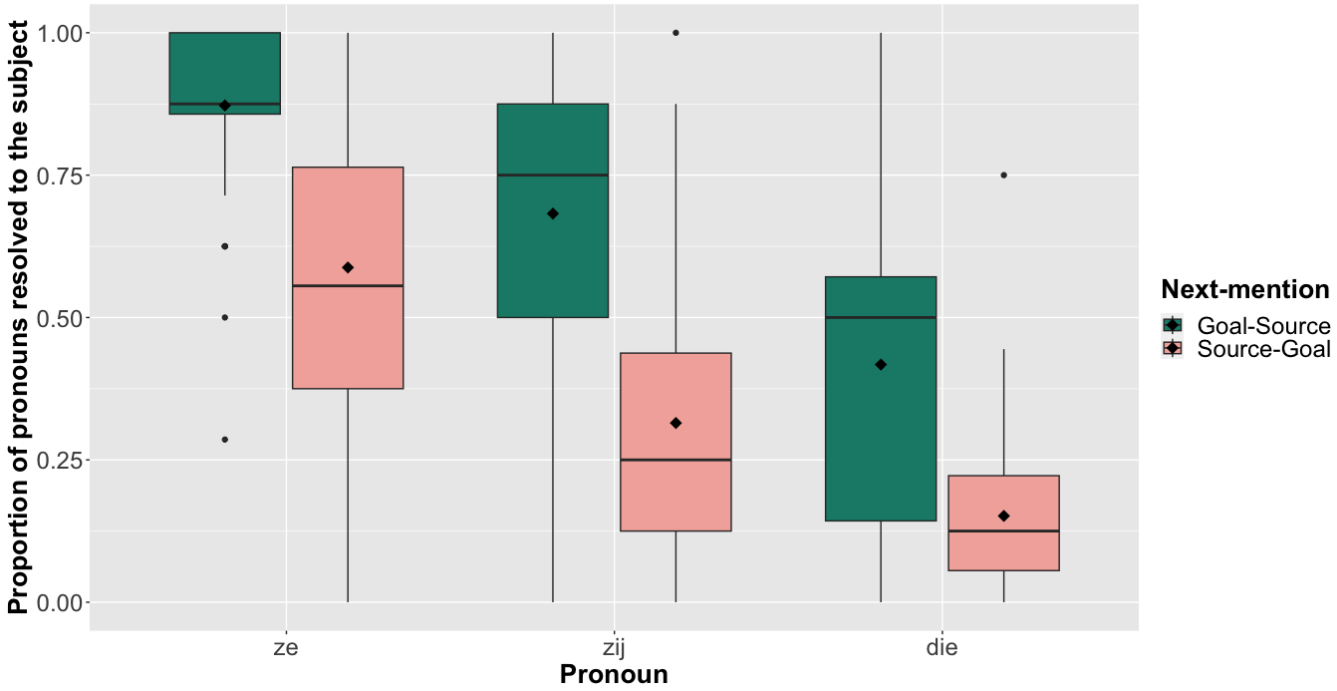

Figure 1: The proportion of ambiguous pronouns resolved to the preceding subject, per pronominal form, per next-mention condition. The tilted squares indicate the means per condition.

preceding subject in the Source-Goal condition than in the Goal-Source condition, across all three pronominal forms.

We modeled the binary outcome of whether the ambiguous pronoun was resolved to the preceding subject or not in a generalized mixed effects model (GLMM: Jaeger (2008)) in R (R Core Team, 2023), using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015); *p*-values were obtained using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017). The three-level independent variable Pronoun was reverse Helmert coded such that the model compared (1) *ze* to *zij* and (2) *die* to the mean of *ze* and *zij*. The two-level independent variable Next-mention bias (N-M) was deviation coded. We included by-item and by-participant random effects in the maximum random effects structure permitted by the model (cf. Barr, Levy, Scheepers, & Tily, 2013). The summary of the model is shown in Table 2.

Table 2: Model summary: Coefficient estimates, standard errors of those estimates, z-values, and *p*-values.

	β	SE	*z*	*p*
ze vs. zij	−0.62	0.06	−10.83	**<.001**
ze&zij vs. die	−0.59	0.03	−16.93	**<.001**
N-M	−1.67	0.17	−9.91	**<.001**
[ze vs. zij]×N-M	−0.02	0.11	−0.13	.90
[ze&zij vs. die]×N-M	−0.06	0.07	0.83	.41

There was a statistically significant difference in the proportion of pronouns resolved to the preceding subject between *ze* and *zij*: *ze* was more often resolved to the subject than *zij*. In addition, *die* was significantly less often interpreted as referring to the subject than *ze* and *zij*. There was also a significant main effect of Next-mention bias: pronouns were more often resolved to the subject in the Goal-Source condition than in the Source-Goal condition. There was no interaction between Pronoun and Next-mention bias.

Discussion

This study investigated how meaning-related and structural cues affect the interpretation of Dutch pronouns, testing three pronominal forms that previous studies had revealed to differ in the strength of their structural bias: reduced personal pronoun *ze*, full personal pronoun *zij*, and demonstrative pronoun *die* . The results were partly as predicted.

For the pronoun manipulation, we found significant differences in the proportion of subject interpretations between all three pronominal forms. As expected, *ze* was most often resolved to the preceding subject, and *die* least often. In addition, we found a main effect of next-mention bias, replicating findings from previous pronoun studies using transfer-of-possession verbs (e.g., Arnold, 2001; Rohde et al., 2006; Rosa & Arnold, 2017; Stevenson et al., 1994). We did not find that the effect of the next-mention manipulation differed between the pronominal forms. The fact that we find no difference between *ze* and *zij* is in line with findings from Patterson et al. (2022), who find that meaning-related cues even impact the interpretation of referring expressions with a fairly strong structural bias. However, our results are not in line with our

prediction that the effect of the next-mention manipulation would be smaller, or even absent, for *die*.

The basis for this prediction was that previous studies have shown a very strong, almost categorical non-subject bias for *die* in contexts with multiple referents (both in production and interpretation). This is not, however, what our results show: on average, *die* was resolved to the preceding subject in 28.2% of all items. When the next-mention expectation was biased toward the subject (Goal-Source condition), the proportion of subject interpretations for *die* was 41.7%, and even when the next-mention expectation was biased toward the non-subject (Source-Goal condition), *die* was interpreted as referring to the subject in 15.3% of the items.

A possible explanation for the, compared to previous studies, relatively weak anti-subject bias we find for *die* might be that we embedded our experimental prompts in a larger context. Kaiser (2011) tested isolated prompts without any meaning-related manipulation, although the verbs used in the prompts, agent-patient / agent-evocator verbs, have been associated with a next-mention bias toward the subject, (e.g., Ferstl, Garnham, & Manouilidou, 2011). While Vogels (2019) did explicitly manipulate next-mention biases, he too used experimental prompts without any context. As discussed in the Introduction, decontextualized experimental prompts have yielded inconsistent results for pronoun studies, especially when it comes to the effect of meaning-related factors. However, even if the addition of context has led to an increased proportion of subject interpretations of *die* in our study, it is still surprising that in the condition where both structural and meaning-related factors bias away from the subject, *die* was interpreted as referring to the subject in 15.3% of the items.

Another explanation could be that the anti-subject bias of *die* in contexts with multiple referents, i.e., a bias toward the most recent referent, is acquired mainly through exposure to written language. Especially in spoken language, it is very common for *die* to refer to the preceding subject (e.g., Wets, Suijkerbuijk, den Hartog, & de Hoop, 2023), but crucially, in these contexts, the preceding subject is always the most recent referent (e.g., *Peter die gaat ervoor!* 'Peter THAT is giving it his all!'; *Gerry is er niet? Nee, die is ziek thuis.* 'Gerry is not here? No, THAT is sick at home'). While we did not ask our participants any questions related to print exposure, the population we tested was very young (M = 19.2 years). Research on other discourse-level elements, connectives (e.g., *because, but*), shows that mastery of connectives mainly confined to the written mode happens only very late: Zufferey and Gygax (2020a) for instance find that many teenagers have a low ability to use these connectives and Zufferey and Gygax (2020b) find that even some adults have trouble correctly understanding them, but that performance increases as a person's amount of print exposure goes up. If in spoken language, *die* is often used in contexts where the preceding subject is also the most recent referent, lots of exposure to written language would be needed to acquire the interpretation bias for *die* that the current experiment aimed to test. To find out whether this could indeed explain our results, a follow-up study should also measure participants' print exposure.

A final, but probably less likely, explanation for the high proportion of subject interpretations for *die* we found might be that *die* is one of the two options that been proposed as a non-binary pronoun in Dutch (the other one being *hen* ∼ 'they'). When used as a non-binary pronoun, *die* is used to refer to non-binary subject referents, even in contexts with other referents, which is bound to lead to a reduction in *die*'s anti-subject bias. If this is the case, it may have influenced our results, even though all referents in our experiment were explicitly feminine. Since this development is fairly recent, this explanation could also account for the discrepancy between our findings and those of Kaiser (2011) and Vogels (2019): even though the use of non-binary *die* is not very frequent (yet), we tested a very young population (M = 19.2 years). Further research would be needed to investigate whether the non-binary use of *die* is (already) affecting the structural bias of *die*.

Overall, our results show that the impact of meaning-related biases on the interpretation of referring expressions does not change depending on the strength of the structural bias. This constitutes especially strong evidence against a model of pronoun interpretation (or language in general) that posits that comprehension and production are mirror images of each other (e.g., Ariel, 1990; Givón, 1983; Gundel, Hedberg, & Zacharski, 1993). Comprehenders are apparently willing to partly disregard the fact that they rarely encounter a pronominal form used to refer to a specific referent in favor of an interpretation that makes the discourse most coherent. While the results are much more in line with the Bayesian model of pronoun interpretation, the surprising resolution rates for *die*, in combination with our lack of production data from the same set of participants, make it hard to estimate how well the model captures our data. A future study should combine interpretation and production data to formally test how accurately the Bayesian model can capture pronoun resolution patterns in Dutch.

In addition, as Figure 1 shows, there appears to be quite some variation between participants. Individual differences in the effect of meaning-related versus structural factors on pronoun interpretation in English have for instance been linked to language users' print exposure (Arnold, Strangmann, Hwang, Zerkle, & Nappa, 2018; Johnson & Arnold, 2021; Langlois & Arnold, 2020); future research should investigate whether this variable has the same explanatory power in more complex pronominal systems, where pronominal forms differ in the strength of their structural bias, and pronominal forms may differ in the frequency with which they are used in written and spoken language.

Acknowledgments

This project was made possible by CLS grant RC2024-03.

References

Ariel, M. (1990). *Accessing noun-phrase antecedents*. Routledge.

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse processes*, *31*(2), 137–162.

Arnold, J. E., Strangmann, I. M., Hwang, H., Zerkle, S., & Nappa, R. (2018). Linguistic experience affects pronoun interpretation. *Journal of Memory and Language*, *102*, 41–54. doi: 10.1016/j.jml.2018.05.002

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. doi: 10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Cheng, W., & Almor, A. (2019). A Bayesian approach to establishing coreference in second language discourse: Evidence from implicit causality and consequentiality verbs. *Bilingualism: Language and Cognition*, *22*, 456–475. doi: 10.1017/S136672891800055X

Cowles, H. W., & Ferreira, V. S. (2012). The influence of topic status on written and spoken sentence production. *Discourse Processes*, *49*(1), 1–28. doi: 10.1080/0163853X.2011.635989

Demberg, V., Kravtchenko, E., & Loy, J. E. (2023). A systematic evaluation of factors affecting referring expression choice in passage completion tasks. *Journal of Memory and Language*, *130*, 104413. doi: 10.1016/j.jml.2023.104413

Ferstl, E. C., Garnham, A., & Manouilidou, C. (2011). Implicit causality bias in English: A corpus of 300 verbs. *Behavior Research Methods*, *43*, 124–135. doi: 10.3758/s13428-010-0023-2

Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study* (Vol. 3). John Benjamins Publishing.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 274–307. doi: 10.2307/416535

Hoek, J., Kehler, A., & Rohde, H. (2021). Pronominalization and expectations for re-mention: Modeling coreference in contexts with three referents. *Frontiers in Communication*, *6*, 674126. doi: 10.3389/fcomm.2021.674126

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Johnson, E., & Arnold, J. E. (2021). Individual differences in print exposure predict use of implicit causality in pronoun comprehension and referential prediction. *Frontiers in Psychology*, *12*, 672109. doi: 10.3389/fpsyg.2021.672109

Kaiser, E. (2011). Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives. *Language and Cognitive Processes*, *26*(10), 1587–1624. doi: 10.1080/01690965.2010.522915

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, *25*(1), 1–44. doi: 10.1093/jos/ffm018

Kehler, A., & Rohde, H. (2019). Prominence and coherence in a Bayesian theory of pronoun interpretation. *Journal of Pragmatics*, *154*, 63-78. doi: 10.1016/j.pragma.2018.04.006

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. doi: 10.18637/jss.v082.i13

Lam, S. Y., & Hwang, H. (2021). Interpretation of null pronouns in Mandarin Chinese does not follow a Bayesian model. *Paper presented at the 34th Annual CUNY Conference on Human Sentence Processing*.

Langlois, V. J., & Arnold, J. E. (2020). Print exposure explains individual differences in using syntactic but not semantic cues for pronoun comprehension. *Cognition*, *197*, 104155. doi: 10.1016/j.cognition.2019.104155

Mayol, L. (2018). Asymmetries between interpretation and production in Catalan pronouns. *Dialogue & Discourse*, *9(2)*, 1-34. doi: 10.5087/dad.2018.201

Patil, U., Bosch, P., & Hinterwimmer, S. (2020). Constraints on German *diese* demonstratives: Language formality and subject-avoidance. *Glossa: A Journal of General Linguistics*, *5*(1), 1–14. doi: 10.5334/gjgl.962

Patterson, C., Schumacher, P. B., Nicenboim, B., Hagen, J., & Kehler, A. (2022). A Bayesian approach to German personal and demonstrative pronouns. *Frontiers in Psychology*, *12*, 672927. doi: 10.3389/fpsyg.2021.672927

R Core Team. (2023). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/` (version 4.3.0)

Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, *29*(8), 912–927. doi: 10.1080/01690965.2013.854918

Rohde, H., Kehler, A., & Elman, J. L. (2006). Event structure and discourse coherence biases in pronoun interpretation. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 28). Retrieved from `https://escholarship.org/uc/item/9js9w79w`

Rosa, E. C., & Arnold, J. E. (2017). Predictability affects production: Thematic roles can affect reference form selection. *Journal of Memory and Language*, *94*, 43–60. doi: 10.1016/j.jml.2016.07.007

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, *9*(4), 519–548. doi: 10.1080/01690969408402130

Vogels, J. (2019). Both thematic role and next-mention biases affect pronoun use in Dutch. In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society* (pp. 3029–3035.).

Wets, I., Suijkerbuijk, M., den Hartog, M., & de Hoop, H. (2023). Argument doubling with proper nouns in spoken Dutch: A corpus study. *Linguistics in the Netherlands*, *40*(1), 248–262. doi: 10.1075/avt.00090.wet

Zehr, J., & Schwarz, F. (2018). *PennController for Internet Based Experiments (IBEX).* doi: 10.17605/OSF.IO/MD832

Zhan, M., Levy, R., & Kehler, A. (2020). Pronoun interpretation in Mandarin Chinese follows principles of Bayesian inference. *PLOS ONE*, *15(8)*, 1-42. doi: https://doi.org/10.1371/journal.pone.0237012

Zufferey, S., & Gygax, P. (2020a). Do teenagers know how to use connectives from the written mode? *Lingua*, *234*, 102779. doi: 10.1016/j.lingua.2019.102779

Zufferey, S., & Gygax, P. (2020b). "Roger broke his tooth. However, he went to the dentist": Why some readers struggle to evaluate wrong (and right) uses of connectives. *Discourse Processes*, *57*(2), 184–200. doi: 10.1080/0163853X.2019.1607446