



von Heusinger, Klaus & Weeber, Frederike & Hoek, Jet & Brocher, Andreas. 2024. Informativity, information status and the accessibility of indefinite noun phrases. *Glossa: a journal of general linguistics* 9(1). pp. 1–20. DOI: <https://doi.org/10.16995/glossa.8888>

OH Open Library of Humanities

Informativity, information status and the accessibility of indefinite noun phrases

Klaus von Heusinger, Department of German Linguistics, University of Cologne, Cologne, DE, klaus.vonheusinger@uni-koeln.de

Frederike Weeber, Department of German Linguistics, University of Cologne, Cologne, DE

Jet Hoek, Department of Language and Communication, Radboud University, Nijmegen, NL, jet.hoek@ru.nl

Andreas Brocher, Department of German Linguistics, University of Cologne, Cologne, DE

In discourse processing, speakers collaborate toward a shared mental model by introducing discourse referents and picking them up with the adequate linguistic forms. Discourse referents compete with each other with respect to their prominence and their accessibility for pronouns. This study focuses on transitive sentences with proper names as subjects and indefinite noun phrases as second arguments, typically direct objects. An ambiguous pronoun in the subsequent sentence may access either referent of the first sentence. Various factors have been shown to influence pronoun resolution, including informativity (how informative is the phrase in which the referent is introduced? E.g., *the waiter* vs. *the waiter at the entrance*) and information status (is the referent given or new in the context?). While both factors have been independently shown to increase referent accessibility, our visual-world eye-tracking experiment shows an original and quite unexpected effect: informativity and information status interact when it comes to the accessibility of indefinite noun phrases: a higher degree of informativity increases accessibility when a referent is brand-new, but surprisingly decreases accessibility when a referent is inferred. We discuss a potential explanation for this surprising pattern in terms of a mismatch between the denotational type of the indefinite and the type required by the modification. We conclude that indefinites strongly interact with additional semantic, contextual and communicative parameters in establishing their referents.

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by the Open Library of Humanities. © 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

OPEN ACCESS



1 Referent accessibility in ambiguous pronoun resolution

Two interlocutors that want to establish a discourse model of which they both have a similar understanding must engage in cooperative referent management. The speaker introduces new referents into the discourse by means of proper names, demonstratives, or definite or indefinite noun phrases (NPs). These discourse referents can be rementioned using anaphoric expressions such as pronouns or definite NPs. The choice of an anaphoric expression depends on the accessibility of the antecedent. More prominent antecedents tend to be picked up by attenuated expressions, such as personal pronouns. Upon hearing a pronoun, a listener will therefore try to resolve it to the most prominent referent.

In our study, we focus on sentences with a subject and a second argument, typically a direct object as in (1), and a second sentence that contains a subject pronoun that is ambiguous between the subject (*Peter*) and the object (*a waiter*) of the first sentence.

(1) Peter called a waiter. He....

We assume that the degree to which a subsequent ambiguous pronoun is resolved to one of the antecedents mirrors the prominence of that referent. We define prominence as a relation between two or more discourse referents (see von Heusinger & Schumacher 2019). Accessibility, on the other hand, is a static property of one discourse referent with respect to its potential anaphoric expressions in the sense of Ariel (1990) and Gundel et al. (1993), reflected by the lexical size of an anaphoric expression. In our experiment with two competing referents and one ambiguous pronoun, both notions are closely correlated, as the pronoun is linked to the more prominent or more accessible antecedent (Crawley & Stevenson 1990; Stevenson et al. 1994; Bosch et al. 2007; Bouma & Hopp 2007; Kaiser & Trueswell 2008; Arnold 2010, Kaiser 2011; Kurczek et al. 2013; Rohde & Kehler 2014, von Heusinger and Schumacher 2019).

There are several factors that determine the prominence of referents and, as a consequence, the processing and resolution of ambiguous pronouns (e.g. Chafe 1976; Hobbs 1979; Givón 1983; Ariel 1990, Gundel et al. 1993, Almor 1999, Kehler et al. 2008). Two of the main factors guiding the disambiguation of pronouns are grammatical role and the position of the argument in the sentence. Often these two factors align, as subjects and arguments that are first in the linear order of referents in a sentence are the most accessible (Givón 1983; Gernsbacher & Hargreaves 1988; Gordon et al. 1993; Gundel et al. 1993; Portele & Bader 2016). According to these factors, the pronoun in (1) would preferably refer to the subject *Peter*, which is also the first argument in that sentence. In addition to these more structural properties, the accessibility of referents is also influenced by semantic factors, including thematic role (Stevenson et al. 1994; Ferretti et al. 2009; Fukumura & van Gompel 2010; Schumacher et al. 2015), topichood (Kaiser 2011), focus (Kaiser 2011), aspect (Kehler et al. 2008; Kehler & Rohde 2013), verb semantics (Stevenson

et al. 1994; Arnold 2001; Kehler et al. 2008; Fukumura & van Gompel 2010; Hartshorne & Snedeker 2013), and prosodic prominence (Baumann & Schumacher 2020).

Two additional factors that might influence accessibility but that are rarely investigated are informativity¹ and information status. In a series of production and comprehension studies, Karimi, Fukumura, Ferreira and Pickering (2014), Karimi and Ferreira (2016) and Karimi, Diaz, and Ferreira (2019) investigated the effect of informativity, operationalized as the length of noun phrases, on accessibility. Karimi and Ferreira (2016) conducted a visual-world eye-tracking experiment, in which participants listened to two-sentence auditory stimuli, see **Table 1**, while being presented with a visual display that contained pictures of two human referents and one unrelated inanimate object. The first sentence of the auditory input had two human referents (NP1 and NP2); the second sentence started with an ambiguous pronoun. The informativity of the NP1 and the NP2 from sentence 1 was manipulated to create three conditions, see **Table 1** (a)–(c).

Sentence number	Manipulation	Sentence
1	short-short long-short short-long	(a) The wizard disagreed with the knight. (b) The wizard who was confused and depressed by the irreparable situation disagreed with the knight. (c) The wizard disagreed with the knight who was confused and depressed by the irreparable situation.
2	–	He suddenly came up with a good idea to solve the problem.

Table 1: Sample items (Karimi & Ferreira 2016).

Fixation times were measured in three windows starting at the onset of the pronoun (W1: 0–1000 ms, W2: 1000–2000 ms, W3: 2000–3000 ms), and the fixation time of NP1 was compared to that of NP2. In the baseline short-short condition, there were significantly more looks to NP1 only in W3. In the long-short condition, there were significantly more looks to (long) NP1 in W2 and W3. Finally, in the short-long condition, there were significantly more looks to (short)

¹ A reviewer suggested that we should use *modification* instead of *informativity*. We agree that in all the discussed experiments the NPs are modified either by a relative clause or by a prepositional phrase (PP) – and we also use ‘modification’ to describe such constructions. Modification is not a very well-defined concept in discourse pragmatics; it primarily refers to a syntactic operation, which is semantically interpreted as providing more information to identify the referent. We think that for building a discourse model, it is about the information we add to discourse referents in the model, and not the operation of modification in the syntactic representation. Therefore, we think the notion of informativity is broader and more appropriate for our contribution.

NP1 in W1, but more looks to (long) NP2 in W3. Karimi and Ferreira (2016) suggest that the higher proportion of looks to NP1 in W1 in the short-long condition might be caused by the complexity of the sentences: People confronted with such a complex sentence might initially rely on structural factors, i.e. on referring to the subject or the first referent. Overall, however, their experiment shows that informativity contributes to the prominence and thus to a higher accessibility of discourse referents. This result is in line with findings from Karimi et al. (2014) and Karimi et al. (2019): More informative, i.e., longer and more complex, noun phrases appear more accessible to pronouns.

The second factor investigated in this paper is information status. We operationalize the familiarity of noun phrases in three categories, namely *given*, *inferred*, and *brand-new*, cf. (Prince 1981b: 237). This three-way distinction, as opposed to a two-way distinction between *given* and *new*, helps accommodate the difference between coreferential and bridged definites (*a car ... the car* vs. *a car ... the steering wheel*) on the one hand and inferred indefinites and brand-new indefinites (*a bar ... a waiter* vs. *an exhibition hall... a waiter*) on the other hand. Strube & Hahn (1999) expand Centering Theory (Grosz et al. 1995) with this 3-way information status categorization in such a way that given referents outrank inferred ones and inferred ones outrank brand-new ones. Using an English and German corpus, they show that their enriched Centering algorithm performs slightly better than the original algorithm.

In a series of eye-tracking experiments, Brocher and von Heusinger (2018) and von Heusinger and Brocher (2019) investigated the effect of information status on the accessibility of indefinite noun phrases in object position in transitive sentences, comparing inferred and brand-new indefinites.² Von Heusinger and Brocher (2019) compared the accessibility of regular indefinite objects to the accessibility of indefinite demonstrative objects in a visual-world experiment very similar to the design of Karimi and Ferreira (2016).³ Auditory input consisted of German short stories of three sentences each, see **Table 2**. The first sentence provided an anchor expression (*bar*) that activated a frame. The second sentence was a transitive sentence with a proper name in subject position and the critical referent (a regular indefinite or an indefinite demonstrative noun phrase) as second argument. The third sentence contained an ambiguous pronoun. The first sentence manipulated the information status of the critical item by using an anchor expression that either evoked a frame such that the critical item was part of that frame and thus pre-activated

² Since given indefinites are highly marked, Brocher and von Heusinger (2018) and von Heusinger and Brocher (2019) compare only inferred with brand-new noun phrases.

³ Indefinite demonstratives are demonstrative noun phrases that are not used deictically or anaphorically. They introduce new discourse referents and can be used in existential sentences (e.g., *There was this young guy standing next to the window*). They are typically used to introduce new discourse items of certain noteworthiness (Prince 1981a; Ionin 2006; von Heusinger 2011).

that item (inferred: *bar* → *waiter*), or a frame that was unrelated to the critical item and thus did not pre-activate that item (brand-new: *exhibition hall* → *waiter*).⁴

Sentence number	Manipulation	Sentence
1	Information status: inferred	In der Bar war es mal wieder rappellvoll. <i>The bar was again very crowded.</i>
	Information status: brand-new	In der Messehalle war es mal wieder rappellvoll. <i>The exhibition hall was again very crowded.</i>
2	Type of indefinite: regular indefinite	(a) Peter rief einen Kellner. <i>Peter called a waiter.</i>
	Type of indefinite: indefinite demonstrative	(b) Peter rief diesen Kellner. <i>Peter called this waiter.</i>
3	–	Als ein Glas vom Tisch fiel, drehte er sich verwundert um und schüttelte den Kopf. <i>When a glass fell off the table, he turned around and shook his head.</i>

Table 2: Sample item from von Heusinger and Brocher (2019).

While participants listened to the short stories, they were presented with a visual display on a screen with pictures for the two arguments from Sentence 2 and two pictures of unrelated inanimate objects. Looks towards the picture of the first argument (NP1: *Peter*) were compared to looks to the picture of the second argument (NP2: *waiter*) from 100 ms before the onset of the ambiguous pronoun to 1,400 ms after pronoun onset, divided in time frames of 300 ms each. There was no difference between inferred and brand-new indefinite demonstratives, but there was a significant effect of information status for regular indefinites in the 1,100 to 1,400 ms bin, with more looks to the inferred indefinites than to the brand-new indefinites. In other words, indefinite demonstratives (*this waiter*) show the same level of accessibility whether brand-new (*exhibition hall ... this waiter*) or inferred (*bar ... this waiter*). But regular indefinites are more accessible when they are inferred than when they are not inferred. This result supports the assumption that information status contributes to the accessibility of discourse referents introduced by definite and indefinite noun phrases (see also Strube & Hahn 1999; Brocher et al. 2016).

⁴ A reviewer correctly noted that a waiter in the context of an exhibition hall is still a plausible referent, which we agree with. However, this seems to be by way of design: Both conditions – the inferred condition (*bar ... waiter*) as well as the brand-new condition (*exhibition hall... waiter*) allow for the continuation with the indefinite, so that both context sentences are able to form a coherent, plausible discourse together with the critical sentence (as opposed to, for example, a waiter showing up at someone's bathroom). Crucially, though, there is a clear difference in expecting a waiter to appear in the frame of a bar or in the frame of an exhibition hall, in that the waiter seems almost obligatory in a bar frame, and merely optional in an exhibition hall.

The two experiments reported above suggest that informativity and information status each contribute to a higher accessibility of referents. However, these two factors may not do so additively. Brocher and von Heusinger (2018) investigated the accessibility of modified (i.e., more informative) definite and indefinite NPs in inferred and brand-new conditions in a visual-world paradigm similar to that of von Heusinger and Brocher (2019). The NPs were modified by post-nominal locative prepositional phrases (PPs) such as *am Fenster* ('next to the window'), see **Table 3**.

Sentence number	Manipulation	Sentence
1	–	Das Fitness-Studio war wieder rappellvoll. <i>The gym was very crowded, as always.</i>
2	inferred + indefinite NP	(a) Philip beobachtete einen Trainer am Fenster. <i>Philip stared at a trainer next to the window.</i>
	inferred + definite NP	(b) Philip beobachtete den Trainer am Fenster. <i>Philip stared at the trainer next to the window.</i>
	brand-new + indefinite NP	(c) Philip beobachtete einen Schönling am Fenster. <i>Philip stared at a pretty boy next to the window.</i>
	brand-new + definite NP	(d) Philip beobachtete den Schönling am Fenster. <i>Philip stared at the pretty boy next to the window.</i>
3	–	Als das Licht ausging, hat er sein Handy als Taschenlampe benutzt./ <i>When the lights went out, he used his cell phone as a flashlight.</i>

Table 3: Sample item from Brocher and von Heusinger (2018).

Unlike in von Heusinger and Brocher (2019), there was no significant positive difference in the proportion of looks toward the object between the brand-new indefinite and the inferred indefinite condition. Numerically, the effect was even reversed.

In this paper, we further investigate influences on the accessibility of indefinite noun phrases building on these experimental insights. Since the main difference between the Brocher and von Heusinger (2018) and von Heusinger and Brocher (2019) experiments is the modification on the NPs, it appears that informativity interacts with information status when it comes to the accessibility of referents introduced by indefinite NPs. The current experiment tests this hypothesized interaction.

2 Current Experiment

We conducted a visual world eye-tracking study, comparing the accessibility, as measured by the probability of looks toward a referent upon hearing an ambiguous pronoun, of unmodified

brand-new indefinites, unmodified inferred indefinites, informative brand-new indefinites, and informative inferred indefinites. This experiment was different from Brocher and von Heusinger (2018) as it did not manipulate the determiner (regular indefinite article vs. indefinite demonstrative), but used regular indefinites in all items and manipulated information status and informativity. If informativity and information status interact, we expect informative brand-new indefinites to be more accessible than less informative brand-new indefinites, but either a reversal of the informativity effect for inferred indefinites or no difference in accessibility between the two inferred indefinite conditions. If the effect of the two factors on the accessibility of referents is additive, we expect informative indefinites to be more accessible than unmodified indefinites, and inferred indefinites to be more accessible than brand-new referents.

2.1 Method

2.1.1 Participants

Thirty-four native speakers of German participated in the experiment for either course credit or monetary compensation (EUR 8/hour). All participants had normal or corrected-to-normal vision.

2.1.2 Materials

We constructed 40 German short stories consisting of three sentences each, see **Table 4**. The first sentence contained context information that did or did not provide an anchor expression of a frame related to the target noun in object position of sentence 2. The second sentence introduced two human referents, one in subject position and one in direct object position. The object referent was either modified by a post-modifying PP, such as *am Eingang* (*at the entrance*) or unmodified.⁵ The PPs that were included in the respective conditions were always descriptions of locations.⁶

⁵ There were several reasons to use PPs to manipulate the informativeness of the NPs: First, we wanted to follow the design of Brocher & von Heusinger 2018; second, we worried about using relative clauses as in the studies by Karimi et al., since relative clauses introduce their own subject referent into the discourse representation. The additional use of that referent could be a confound for informativity. Third, we did not use modification via adjectives like *big* or *enthusiastic* since they would add properties that should be mirrored in the pictures.

⁶ A reviewer observed that depending on the test item the PP can be either interpreted as modifying the NP or the VP or both, even though we designed the items with the intention of the PP only modifying the NP. As a post-hoc analysis, two persons (one of the authors and one independent person) annotated all sentences with respect to the PP clearly modifying the NP or the VP or both. Both annotators judged that in 21 of the 40 test items the PP only modified the NP, where the judgments overlapped in about 62% of the cases. We can therefore split the items in three different ways: 1) NP-modifying versus not NP-modifying as judged by person A; 2) NP-modifying versus not NP-modifying as judged by person B; 3) and finally those items where both coders agreed the PP modified only the NP versus the rest. Regardless of which subset we analyze, the crucial effect reported in the Results section goes in the same direction as the effect based on the full data (effects are either significant or marginally significant; statistical power is of course diminished in these additional explorative analyses). This leads us to believe that this potential ambiguity in our experimental items has not impacted our results.

The locations were not represented in the accompanying visual display, which only showed the referents in isolation (see **Figure 1**). The final sentence of the auditory input always started with a temporal subordinate clause,⁷ followed by a main sentence including an ambiguous pronoun that referred to either the subject or the object of the second sentence. The pronoun was always the German masculine singular *er* ('he') and ambiguous between two male referents, because the feminine singular *sie* ('she' or 'they') is also compatible with a plural interpretation, which would then no longer be ambiguous between two singular antecedents.

Sentence number	Manipulation	Sentence
1	Information status: inferred	In der Bar war es mal wieder rappellvoll. <i>The bar was again very crowded.</i>
	Information status: brand-new	In der Messehalle war es mal wieder rappellvoll. <i>The exhibition hall was again very crowded.</i>
2	Informativity: modified	Peter rief einen Kellner am Eingang. <i>Peter called a waiter at the entrance.</i>
	Informativity: unmodified	Peter rief einen Kellner. <i>Peter called a waiter.</i>
3	–	Als ein Glas vom Tisch fiel, drehte er sich verwundert um und schüttelte den Kopf. <i>When a glass fell off the table, he turned around and shook his head.</i>

Table 4: Sample stimulus (English translations in italics).

In addition to the 40 experimental short stories we also included 80 filler stories in the experiment, which were similar in length and structure to the experimental items. To make them as similar as possible to the experimental stories, the fillers also only introduced male characters. None of the filler items involved any referential ambiguity.

All materials were recorded by a trained female speaker and were pre-processed before the experiment. All pre-processing procedures were conducted in Praat (Boersma 2001). For all experimental stories, the three sentences (context sentence, referent sentence, ambiguous pronoun sentence) were first separated and then re-arranged to keep everything except the manipulated sentence identical between conditions (e.g., all four conditions of an item had the

⁷ The temporal clause served as a neutral region before the ambiguous pronouns was introduced. We tried to use either no referential arguments ("When the sun set") or unrelated referents. However, one reviewer noted that some of the referents might be associated with the indefinite in the critical sentence (*monk ... church bells*). We only found 4 out of 40 test items with such a possible association; but even in these cases the association to a referent in the context sentence is much stronger (*monastery ... church bells*).

same final sentence, both stories in the brand-new condition used the same recording of the first sentence, etc.).

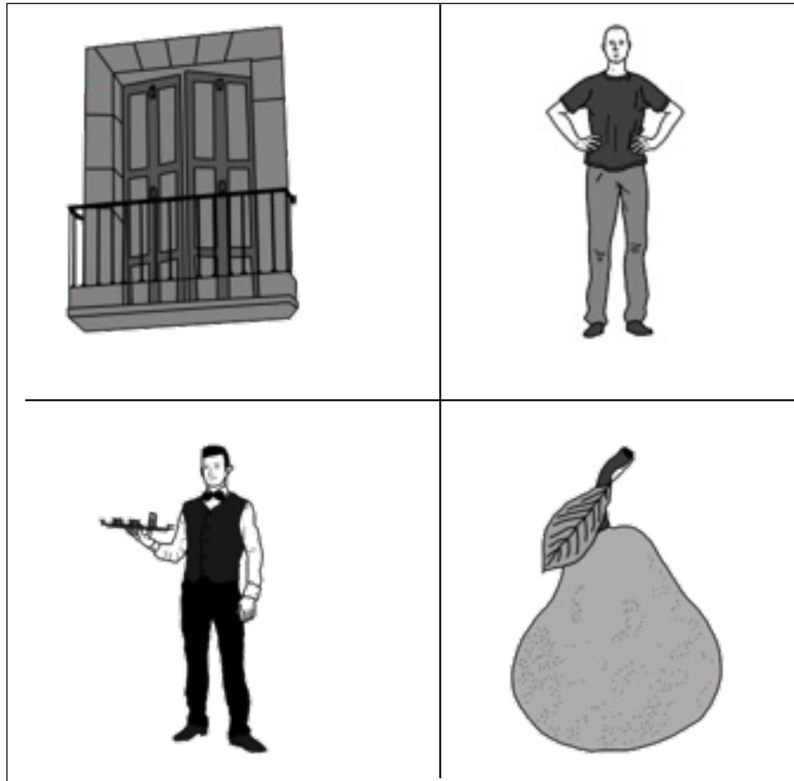


Figure 1: Example display. The referent in subject position (see Table 4), *Peter*, is presented in the upper right corner, the critical referent in object position, *waiter*, in the lower left corner.

Each auditory stimulus had a corresponding visual display with four pictures: one for the referent in subject position, one for the referent in object position, and two unrelated filler pictures showing inanimate objects. **Figure 1** shows the display for the item in **Table 4**. We partly selected our pictures from a database by Duñabeitia et al. (2018). The remaining pictures for test items, including all pictures of the human referents, were provided by a professional illustrator. The four pictures were arranged in the same way for each condition of an item. Between items, subject and object pictures appeared equally often in each corner of the screen.

All stories were distributed across four presentation lists. No item appeared more than once per list, and all conditions appeared equally often between lists.

2.1.3 Procedure

Participants sat in front of an EyeLink 1000 configured with a tower mount at a distance of approximately 65 cm from the screen. We recorded participants' eye movements at a sampling

rate of 250 Hz. At the beginning of a trial, we presented participants with the picture of the subject referent that was about to appear in the following story. This picture was accompanied by a phrase that introduced the referent, such as *Hello, my name is Peter*. This was done because, unlike for the object referents, subject referents could not be depicted in an identifiable way (e.g., there is no picture of ‘Peter’). After familiarization, participants clicked their mouse to start the next screen, which showed the entire visual display, see **Figure 1**.

After 500 ms, the audio started. Participants were asked to carefully follow the stories to be able to correctly answer a comprehension question in case one appeared. 1,000 ms after the end of the auditory stimulus, participants saw either a comprehension question or the question *Weiter?* (‘Continue?’).⁸ In case of a comprehension question, participants pressed the right mouse button to respond ‘yes’ and the left button to respond ‘no.’ 2,000 ms after the response, the next trial was initiated.

2.1.4 Data analysis

For pre-processing of the data and inferential statistics, we used R (version 3.5.0) with the *lme4* (version 1.1.19) and the *tidyverse* package (version 1.2.1). We analyzed eye fixations on the object and subject pictures for a total of 1,500 ms, from 100 ms before until 1,400 ms after onset of the ambiguous pronoun. We divided the 1500 ms into five windows of 300 ms each. For each time window and picture individually (i.e. object and subject picture), we fitted a linear mixed effects model with the proportion of looks to the target picture as a dependent variable and information status (inferred vs. brand-new) and informativity (modified vs. unmodified) as predictors.

All models contained between-participants and between-items intercepts as well as random slopes for information status, informativity, and their interaction. We assume absolute *t*-values of 2 or higher as indicating statistical robustness (Baayen 2008).

2.2 Results

Accuracy of responses to the comprehension questions was 90.1%. 12.9% of the data were not included in the analyses, either because they reflected blinks or because no picture on the screen had been fixated.⁹

⁸ Of the 80 fillers, 30 were followed by a “yes”/”no” comprehension question, as in (i) (English translation):

(i) A new picture for the ID card was due. Jonas went to a photographer in the neighboring town. When suddenly the lens of the camera fell down, the two of them were startled and looked to the ground.
control question: Did Jonas need a new picture for his ID card? (correct answer: yes).

⁹ Non-fixations were equally distributed between conditions: brand-new/with-PP = 3.3%, brand-new/without-PP = 3.3%, inferred/with-PP = 3.2%, and inferred/without-PP = 3.1%. Looks to the two unrelated filler pictures were also equally distributed between conditions: brand-new/with-PP = 1.8%, brand-new/without-PP = 1.7%, inferred/with-PP = 2.1%, and inferred/without-PP = 1.9%.

The time course of looks to the picture of the object referent is plotted in **Figure 2**. The plot shows that, as early as the onset of the pronoun, referents of unmodified NPs received overall more looks when they were inferred than when they were brand-new. This part of the results is in line with the findings from von Heusinger and Brocher (2019). For modified, i.e., more informative, NPs, the opposite pattern emerges somewhat later in the time course: Referents of brand-new noun phrases were fixated more than referents of inferred ones. This pattern is in line with the results of Brocher and von Heusinger (2018). **Table 5** summarizes the results of the inferential statistics.

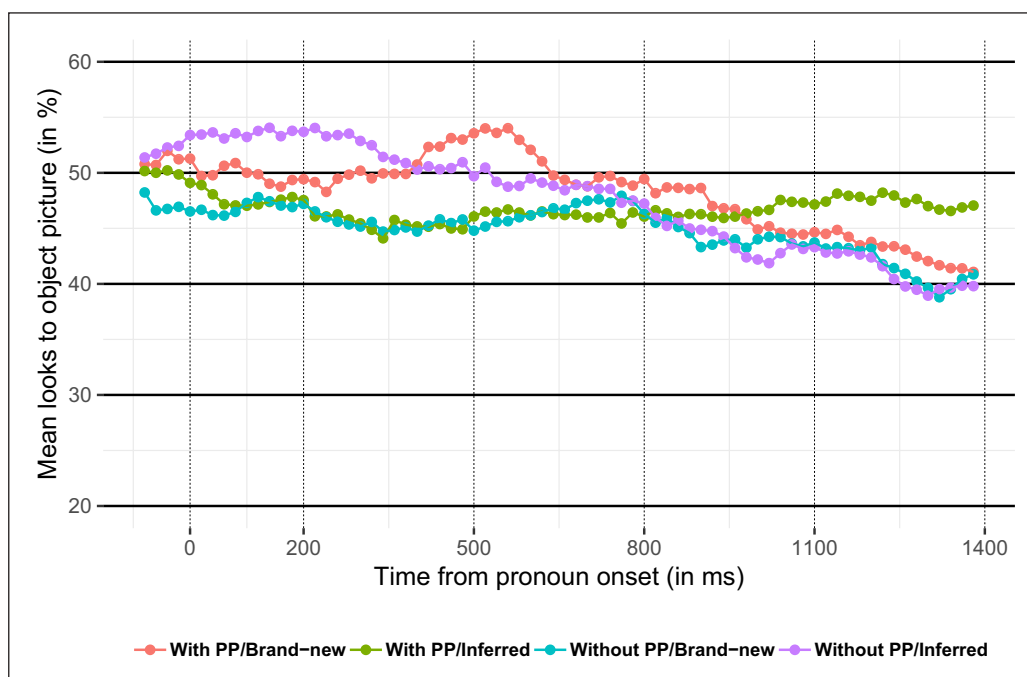


Figure 2: Time course of looks (in %) to the picture of the referent in object position; Zero marks the onset of the ambiguous pronoun; *Without PP* = less informative object NPs; *With PP* = informative object NPs; *Brand-new* = object NPs with brand-new concept; *Inferred* = object NPs with inferred concept.

The output of the regression model that included the fixation times for the window 200 ms – 500 ms¹⁰ post pronoun onset revealed a significant Information status x Informativity interaction.

¹⁰ A reviewer correctly observed that the significant difference in this study is in a very early bin (200 ms – 500 ms after pronoun onset), while in the study of von Heusinger & Brocher (2019) it is in the 1,100 to 1,400 ms bin and in the Karimi et al. studies even after 2,000 ms. Since in all other experiments indefinite NPs were compared to definite expressions (definite NPs, demonstrative NPs), we speculate that in absence of definiteness, potential effects of informativity and information status show up earlier.

No other effects reached statistical reliability.¹¹ To help interpret the interaction effect, we plotted the model estimated marginal means (obtained using the *emmeans* package, version 1.8.5), see **Figure 3**. As the plot shows, the effect of Informativity is different for brand-new referents than for inferred referents: while a PP makes inferred referents less accessible, it makes brand-new referents more accessible. This effect is in line with the main hypothesis of our study.

Time	Main Effect/Interaction	b	SE	t
–100 – 200	Intercept	0.440	0.031	14.00
	Information status	–0.016	0.026	–0.60
	Informativity	0.012	0.031	0.40
	Information status × Informativity	–0.079	0.050	–1.59
200 – 500	Intercept	0.428	0.027	16.12
	Information status	–0.016	0.026	–0.63
	Informativity	0.006	0.031	0.19
	Information status × Informativity	–0.111	0.051	–2.17
500 – 800	Intercept	0.421	0.027	15.87
	Information status	0.014	0.031	0.47
	Informativity	–0.015	0.027	–0.55
	Information status × Informativity	–0.060	0.054	–1.10
800 – 1100	Intercept	0.394	0.026	14.88
	Information status	0.004	0.030	0.13
	Informativity	–0.027	0.028	–0.95
	Information status × Informativity	0.004	0.048	0.09
1100 – 1400	Intercept	0.372	0.024	15.27
	Information status	–0.016	0.033	–0.49
	Informativity	–0.033	0.026	–1.29
	Information status × Informativity	0.048	0.047	1.02

Table 5: Model output for looks to the picture of the object referent.

Notes. Time in milliseconds relative to the onset of the ambiguous pronoun. *Information status* = information status of the object noun phrase (inferred or brand-new); *Informativity* = Informativity of the object noun phrase (modified by a PP or not). Significant effects are highlighted in bold.

¹¹ We also see the reverse effect on the subject. There is a statistical trend for the Information status × Informativity interaction in the 500 ms – 800 ms window, as well as a reliable main effect of complexity in the 1100 ms – 1400 ms window.

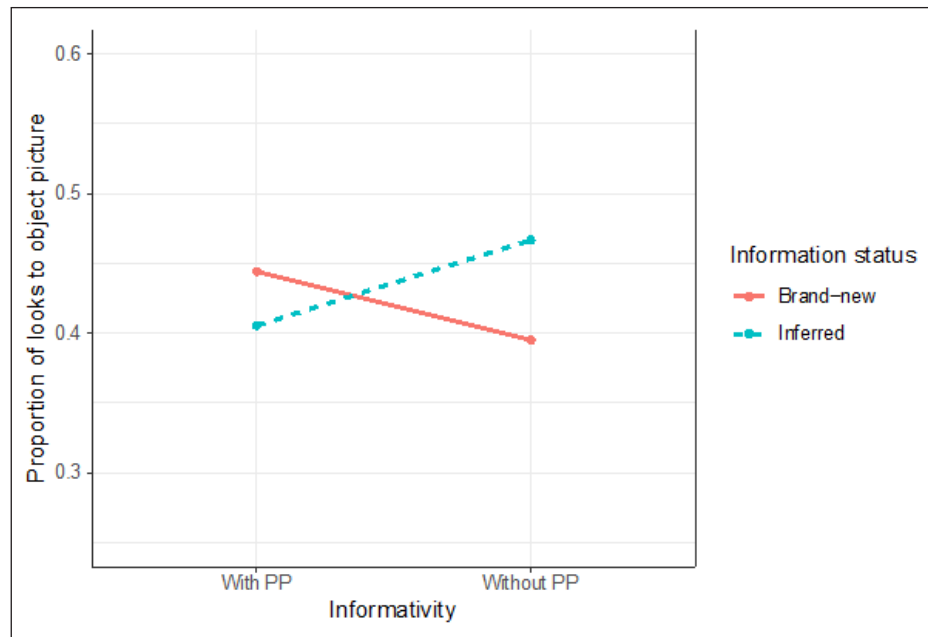


Figure 3: Model estimated marginal means of the proportion of looks to the object picture in the 200–500 ms time bin, by Information status (brand-new versus inferred) and Informativity (with PP versus without PP).

2.3 Discussion

Results from our experiment show a significant interaction between informativity (unmodified vs. modified realized by a postnominal PP) and information status (inferred vs. brand-new) with respect to the accessibility of indefinite direct objects. Specifically, unmodified indefinites in object position are more accessible when they are inferred than when they are brand-new, but for modified, i.e., more informative, indefinites, the pattern is reversed: Modified indefinites are more accessible when they are brand-new than when they are inferred.

3 General Discussion

Our results are in line with studies that show that informativity and information status affect the accessibility of referents. However, our study is the first to show that these two factors interact: Modification makes inferred indefinites less accessible, while it makes brand-new indefinites more accessible. Informativity boosting the accessibility of brand-new referents is in line with the results of von Heusinger & Brocher (2019) for indefinites, and of Karimi et al. for definites. However, the finding that informativity has a negative effect on the accessibility of inferred referents is more surprising, although a first hint of this effect was suggested by the experiment reported in Brocher & von Heusinger (2018) for indefinites. As this result is not predicted by any theoretical account, but seems robust, we are in need of an explanation of

why informativity impacts inferred indefinite NPs differently than brand-new ones. We discuss one potential explanation for this interaction, namely that the type or role reading triggered by inferred indefinites does not match the locative PP modification, which requires a token or individual reading.

Before we elaborate on this explanation, we have to briefly discuss two assumptions that have so far not been discussed in this paper: First, we assume that while individual referents are typically introduced by pronouns, demonstratives and definite NPs, the semantic contribution of indefinites is a predicate or a concept describing a set of elements with the same property (Heim & Kratzer 1998). Additional contextual and general pragmatic inferences allow type shifts from predicates to individual referents (tokens) or to roles (types) (Partee 1987). We assume that contextual information can trigger one or the other type shift. Second, we assume that in our test design the ambiguous pronoun in the third sentence refers to an individual referent, as required by the predicate in that sentence (e.g. *turned around*). We assume that the pronoun establishes the anaphoric relation to the indefinite direct object by two interacting processes: First, the referent of the pronoun must be an instance of the concept introduced by the indefinite, and second the pronoun refers to the same individual referent as the indefinite in the antecedent sentence. Both processes support each other and create the prominence level of the direct object that can compete with the prominence level of the subject.

A potential explanation could thus be developed along the following lines: we start with the assumption that indefinites introduce a predicate that describes a set of potential referents (say different waiters). During the process of the incremental interpretation this predicate can be shifted to an individual (or token) or to a role (or type) or the indefinite can be underspecified and both types are available. In the case of brand-new indefinites (*exhibition hall ... a waiter [PP]*), the indefinite is shifted to a token interpretation that can be picked up by the anaphoric pronoun in the next sentence without any type mismatch. The additional modification with a locative PP provides more information and makes the referent more prominent. This would account for the observation that for brand-new indefinites modification (more informativity) raises accessibility. The situation is different for inferred indefinites (*bar ... a waiter [PP]*). The frame of a bar triggers a role (or type) reading of the indefinite, i.e., *a waiter* refers to a stereotypical referent fulfilling the assumed properties of a waiter in a bar.¹² In the unmodified case, the personal pronoun in the

¹² There is no exhaustive linguistic definition of “role” as a semantic category, but there are many very similar categories, such as *type*, *stereotype*, *established (sub-)kind*, *socially accepted kind*, etc. While a full discussion of these notions is beyond the scope of this paper, we will very briefly sketch what we understand to fall under a “role”: A role is a stereotypical type that has various instantiations (tokens) with a set of characteristic properties or features. This notion is very close to the notion of an established (sub-)kind. The role of a waiter may include properties like *attentive* and *fast*, but not *having a red shirt* or *being tired*. A linguistic test for a role (or subkind) reading is the use of the anaphoric *so* ‘such’ in German (see König & Umbach 2018) and a psycholinguistic test is the acceptability of characteristic properties (*attentive*) vs. accidental properties (*having a red shirt*) as discussed in Prasada et al. (2013).

third sentence refers then to an instance of the role or type, and not yet to a specific individual referent. Since the locative PP clashes with the frame reading, the PP triggers a shift to a token reading. For both brand-new indefinites and inferred indefinites, the eventual interpretation is thus a token reading, but the frame reading is retained longer for inferred indefinites. Since the locative PP triggers the token reading, it does not have the prominence-boosting effect that it has on already established referents (token readings), resulting in an overall lower prominence level.

We thus postulate that the interaction of informativity and information status is caused by the very underspecified semantics of indefinites. This would predict that for other NPs, such as definite descriptions (*the waiter*) or demonstratives (*this waiter*), such an interaction cannot be found, a prediction that is supported by Brocher and von Heusinger's (2018) results: In their study, inferred informative definites are more accessible than brand-new informative ones.

To sum up, we corroborated studies by Brocher and von Heusinger (2018) and von Heusinger and Brocher (2019) that find that informativity and information status both influence the accessibility of discourse referents, and we provided novel empirical evidence that these features interact when it comes to the accessibility of indefinite NPs. Future research should work toward uncovering the exact mechanism behind this interaction, for instance by using another type of modification, or modification that is compatible versus incompatible with a role reading (e.g., *the mathematics teacher* vs. *the big teacher*); this will contribute to a better understanding of the discourse-level function of indefinite NPs, the processes governing pronoun resolution, and, in general, discourse processing.

Data Availability

All experimental items as well as our data and data analysis script can be accessed at <https://doi.org/10.17605/OSF.IO/46REU>.

Ethics and consent

The experiment was approved by the Ethics committee of the German Linguistic Society before data collection commenced.

Funding information

The research for this paper was funded by the German Research Foundation (DFG) as part of the SFB 1252 “Prominence in Language” (Project-ID 281511265) in the project C04 “Conceptual and referential activation in discourse” at the University of Cologne, Department of German Language and Literature I, Linguistics.

Acknowledgements

Materials discussed in this paper were presented at the University of Southern California in April 2022, at the 2022 Seoul International Conference on Linguistics (SICOL-2022) Online Conference in August 2022 and at the Workshop on the Semantics of Indefinites, Syntax and Semantics in September 2023 in Paris. We are indebted to three anonymous reviewers for their constructive feedback as well as to the editor Lyn Tieu for the continuous support and Cornelia Ebert, Hans Kamp, Elsi Kaiser and Carla Umbach for comments, which have greatly improved this contribution. We would also like to thank Alexander Walter and Hae-Eun Cho for their help with the experiment and Timo Buchholz for support with the statistical analysis.

Competing interests

The authors have no competing interests to declare.

References

- Almor, Amit. 1999. Noun-phrase anaphora and focus. The informational load hypothesis. *Psychological Review* 106(4). 748–765. DOI: <https://doi.org/10.1037//0033-295X.106.4.748>
- Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London, New York: Routledge.
- Arnold, Jennifer E. 2001. The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes* 31(2). 137–162. DOI: https://doi.org/10.1207/S15326950DP3102_02

- Arnold, Jennifer E. 2010. How speakers refer: The role of accessibility. *Language and Linguistics Compass* 4(4). 187–203. DOI: <https://doi.org/10.1111/j.1749-818X.2010.00193.x>
- Baayen, R. H. 2008. *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511801686>
- Baumann, Stefan & Schumacher, Petra B. 2020. The incremental processing of focus, givenness and prosodic prominence. *Glossa: a journal of general linguistics* 5(1). 1–30. DOI: <https://doi.org/10.5334/gjgl.914>
- Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5(9/10). 341–347.
- Bosch, Peter & Katz, Graham & Umbach, Carla. 2007. The non-subject bias of German demonstrative pronouns. In Schwarz-Friesel, Monika & Consten, Manfred & Knees, Mareile (eds.), *Anaphors in text: Cognitive, formal and applied approaches to anaphoric reference*, 145–164. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/slcs.86.13bos>
- Bouma, Gerlof & Hopp, Holger. 2007. Coreference preferences for personal pronouns in German. In Bittner, Dagmar & Gargarina, Natalia (eds.), *Proceedings of the Conference on Intersentential Pronominal Reference in Child and Adult Language* (ZAS Papers in Linguistics 48), 53–74. Berlin: ZAS. DOI: <https://doi.org/10.21248/zaspil.48.2007.354>
- Brocher, Andreas & Chiriacescu, Sofiana & von Heusinger, Klaus. 2016. Effects of information status and uniqueness status on referent management in discourse comprehension and planning. *Discourse Processes* 55(4). 346–370. DOI: <https://doi.org/10.1080/0163853X.2016.1254990>
- Brocher, Andreas & von Heusinger, Klaus. 2018. A dual-process activation model: Processing definiteness and information status. *Glossa: a journal of general linguistics* 3(1). 1–34. DOI: <https://doi.org/10.5334/gjgl.457>
- Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, Charles N. (ed.), *Subject and topic*, 25–55. New York: Academic Press.
- Crawley, Rosalind A. & Stevenson, Rosemary J. 1990. Reference in single sentences and in texts. *Journal of Psycholinguistic Research* 19(3). 191–210. DOI: <https://doi.org/10.1007/BF01077416>
- Duñabeitia, Jon Andoni & Crepaldi, Davide & Meyer, Antje S. & New, Boris & Pliatsikas, Christos & Smolka, Eva & Brysbaert, Marc. 2018. MultiPic: a standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology* 71(4). 808–816. DOI: <https://doi.org/10.1080/17470218.2017.1310261>
- Ferretti, Todd R. & Rohde, Hannah & Kehler, Andrew & Crutchley, Melanie. 2009. Verbal aspect, event structure, and coreferential processing. *Journal of Memory and Language* 61(2). 191–205. DOI: <https://doi.org/10.1016/j.jml.2009.04.001>
- Fukumura, Kumiko & van Gompel, Roger P. G. 2010. Choosing anaphoric expressions: Do people take into account likelihood of reference? *Journal of Memory and Language* 62. 52–66. DOI: <https://doi.org/10.1016/j.jml.2009.09.001>
- Gernsbacher, Morton Ann & Hargreaves, David J. 1988. Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language* 27(6). 699–717. DOI: [https://doi.org/10.1016/0749-596X\(88\)90016-2](https://doi.org/10.1016/0749-596X(88)90016-2)

- Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In Givón, Talmy (ed.), *Topic continuity in discourse: A quantitative cross-language study*, 1–42. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/tsl.3.01giv>
- Gordon, Peter & Grosz, Barbara & Gilliom, Laura. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science* 17(3). 311–341. DOI: https://doi.org/10.1207/s15516709cog1703_1
- Grosz, Barbara & Joshi, Aravind & Weinstein, Scott. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21. 203–226. DOI: <https://doi.org/10.21236/ADA324949>
- Gundel, Jeanette & Hedberg, Nancy & Zacharski, Ron. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274–307. DOI: <https://doi.org/10.2307/416535>
- Hartshorne, Joshua H. & Snedeker, Jesse. 2013. Verb argument structure predicts implicit causality: The advantages of finer-grained semantics. *Language and Cognitive Processes* 28(10). 1474–1508. DOI: <https://doi.org/10.1080/01690965.2012.689305>
- Heim, Irene & Kratzer Angelika. 1998. *Semantics in generative grammar*. Malden, Massachusetts: Blackwell.
- Hobbs, Jerry R. 1979. Coherence and coreference. *Cognitive Science* 3(1). 67–90. DOI: https://doi.org/10.1207/s15516709cog0301_4
- Ionin, Tania. 2006. This is definitely specific: Specificity and definiteness in article systems. *Natural Language Semantics* 14. 175–234. DOI: <https://doi.org/10.1007/s11050-005-5255-9>
- Kaiser, Elsi. 2011. Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives. *Language and Cognitive Processes* 26(10). 1587–1624. DOI: <https://doi.org/10.1080/01690965.2010.522915>
- Kaiser, Elsi & Trueswell, John C. 2008. Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes* 23(5). 709–748. DOI: <https://doi.org/10.1080/01690960701771220>
- Karimi, Hossein & Diaz, Michele & Ferreira, Fernanda. 2019. “A cruel king” is not the same as “a king who is cruel”: Modifier position affects how words are encoded and retrieved from memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 45(11). 2010–2035. DOI: <https://doi.org/10.1037/xlm0000694>
- Karimi, Hossein & Ferreira, Fernanda. 2016. Informativity renders a referent more accessible: Evidence from eyetracking. *Psychonomic Bulletin and Review* 23(2). 507–525. DOI: <https://doi.org/10.1080/17470218.2015.1053951>
- Karimi, Hossein & Fukumura, Kumiko & Ferreira, Fernanda & Pickering, Martin J. 2014. The effect of noun phrase length on the form of referring expressions. *Memory & Cognition* 42. 993–1009. DOI: <https://doi.org/10.3758/s13421-014-0400-7>
- Kehler, Andrew & Kertz, Laura & Rohde, Hannah & Elman, Jeffrey L. 2008. Coherence and coreference revisited. *Journal of Semantics* 25. 1–44. DOI: <https://doi.org/10.1093/jos/ffm018>

- Kehler, Andrew & Rohde, Hannah. 2013. A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics* 39. 1–37. DOI: <https://doi.org/10.1515/tl-2013-0001>
- König, Ekkehard & Umbach, Carla. 2018. Demonstrative of manner, of quality and of degree. In Coniglio, Marco & Murphy, Andrew & Schlachter, Eva & Veenstra, Tonjes (eds.), *Atypical demonstratives. Syntax, semantics and pragmatics*, 256–285. Berlin: De Gruyter. DOI: <https://doi.org/10.1515/9783110560299-010>
- Kurczek, Jake & Brown-Schmidt, Sarah & Duff, Melissa C. 2013. Hippocampal contributions to language: Evidence of referential processing deficits in amnesia. *Journal of Experimental Psychology: General* 142(4). 1346–1354. DOI: <https://doi.org/10.1037/a0034026>
- Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In Groenendijk, Jeroen & de Jongh, Dick & Stokhof, Martin (eds.), *Studies in Discourse Representation Theory and the theory of generalized quantifiers*, 115–143. Dordrecht: Foris. DOI: <https://doi.org/10.1515/9783112420027-006>
- Portele, Yvonne & Bader, Markus. 2016. Accessibility and referential choice: Personal pronouns and d-pronouns in written German. *Discours* 18. 3–39. DOI: <https://doi.org/10.4000/discours.9188>
- Prasada, Sandeep & Khemlani, Sangeet & Leslie, Sarah-Jane & Glucksberg, Sam. 2013. Conceptual distinctions amongst generics. *Cognition* 126. 405–422. DOI: <https://doi.org/10.1016/j.cognition.2012.11.010>
- Prince, Ellen F. 1981a. On the inferencing of indefinite-*this* NPs. In Webber, Bonnie & Joshi, Aravind K. & Sag, Ivan (eds.), *Elements of discourse understanding*, 231–250. Cambridge: Cambridge University Press.
- Prince, Ellen F. 1981b. Toward a taxonomy of given-new information. In Cole, Peter (ed.), *Radical Pragmatics*, 223–255. New York: Academic Press.
- Rohde, Hanna & Kehler, Andrew. 2014. Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience* 29(8). 912–927. DOI: <https://doi.org/10.1080/01690965.2013.854918>
- Schumacher, Petra B. & Backhaus, Jana & Dangl, Manuel. 2015. Backward- and forward-looking potential of anaphors. *Frontiers in Psychology* 6. 1–14. DOI: <https://doi.org/10.3389/fpsyg.2015.01746>
- Stevenson, Rosemary & Crawley, Rosalind & Kleinman, David. 1994. Thematic roles, focus, and the representation of events. *Language and Cognitive Processes* 9(4). 519–548. DOI: <https://doi.org/10.1080/01690969408402130>
- Strube, Michael & Hahn, Udo. 1999. Functional centering: Grounding referential coherence in information structure. *Computational Linguistics* 25(3). 309–344.
- von Heusinger, Klaus. 2011. Specificity. In von Heusinger, Klaus & Maienborn, Claudia & Portner, Paul (eds.), *Semantics. An International Handbook of Natural Language Meaning*, vol. 2, 1025–1058. Berlin: De Gruyter. DOI: <https://doi.org/10.1515/9783110255072.1025>

von Heusinger, Klaus & Brocher, Andreas. 2019. Indefinite demonstratives, definiteness, and referentiality. Paper presented at *the Workshop “Sorting out the concepts behind definiteness” at the 41th Meeting of the German Linguistic Society*. Bremen, March, 6–8, 2019. <https://sites.google.com/view/sortingoutdefiniteness/program>.

von Heusinger, Klaus & Schumacher, Petra B. 2019. Discourse prominence: Definition and application. *Journal of Pragmatics* 154. 117–127. DOI: <https://doi.org/10.1016/j.pragma.2019.07.025>

