The Role of Discourse in Italian Pronoun Interpretation: Investigating Variations in Experimental Results with Cognitive Modeling

Margreet Vogelzang (margreet.vogelzang@uni-oldenburg.de)
Institute of Dutch Studies and Cluster of Excellence "Hearing4all"
University of Oldenburg, Germany

Abstract

This paper investigates to what extent the variations in experimental results on the interpretation of Italian subject pronouns can be explained by the different discourses used in the experimental studies. A cognitive model implemented in ACT-R was used to simulate pronoun processing and interpretation in discourse, which is influenced by the various contexts used in empirical experiments. Our simulations show that the discourse contexts used in the experiments strongly influence the interpretation of Italian subject pronouns, but not to the extent that all data in different experiments can be explained by it. We therefore conclude with suggestions for further research both on the influence of discourse context and the influence of task on the interpretation of Italian pronouns and (linguistic) experiments in general.

Keywords: pronoun interpretation; Italian; cognitive modeling; null subjects; discourse context

Introduction

Referring expressions such as pronouns (*he*, *she*) occur frequently in daily life. It is therefore essential for successful communication that such expressions are understood correctly. Nevertheless, experimental studies on subject pronoun interpretation in Italian show that healthy native adults sometimes have interpretation preferences of around 50% (Tsimpli, Sorace, Heycock, & Filiaci, 2004), suggesting that they are not sure what the meaning is supposed to be. Notably, experimental findings on Italian subject pronoun interpretation show a substantial amount of variation (compare, e.g., Carminati, 2002; Tsimpli et al., 2004; Vogelzang, Foppolo, Guasti, Van Rijn, & Hendriks, 2019). For example, Carminati (2002) and Vogelzang et al. (2019) found a strong preference for null pronouns to refer back to the subject antecedent, whereas in the experiment of Tsimpli et al. (2004) participants only selected the subject antecedent for a null pronoun around half of the time.

One explanation for the varying results could be that these experiments have used different experimental stimuli and different tasks and methodologies. It is generally known that the interpretation of referring expressions can be influenced by the surrounding discourse context. Previous studies have identified several discourse factors that can influence the interpretation of referring expressions. For example, the prominence of discourse referents determines whether they are likely antecedents of a referring expression. In general,

less informative referring expressions such as pronouns refer to entities that are highly prominent in the discourse, and more informative referring expressions such as full noun phrases (NPs) refer to entities that are less prominent in the discourse (cf. a.o. Ariel, 1990; Givón, 1983; Gundel, Hedberg, & Zacharski, 1993; for an overview see Arnold, 1998). Additional attention has been given in the literature to coherence relation (Kehler, Kertz, Rohde, & Elman, 2008) and pragmatic plausibility (Carminati, 2002) as influencing the interpretation of referring expressions. In the case of object pronoun interpretation in Dutch, it has been shown that a change in discourse, making the most prominent referent (i.e. the discourse topic) more clear, can eliminate interpretation difficulties that children typically show in other studies (Spenader, Smits, & Hendriks, 2009).

Thus, it is known that discourse context influences the processing and interpretation of referring expressions. It is also known that experimental findings on the interpretation of Italian subject pronouns have not been consistent. It is therefore a logical next step to investigate to what extent differences in discourse can explain variations in experimental results on Italian pronoun interpretation. In this paper, we will focus solely on the influence of the specific experimental stimuli used, putting aside any differences in task, participant sample, and so on. To this end, we will more extensively examine the discourse contexts and experimental results of Tsimpli et al. (2004) and Vogelzang et al. (2019).

The influence of these experimental discourses will be investigated using a cognitive model developed within the cognitive architecture ACT-R. The model will be used to simulate existing empirical data, which will be discussed in the next section. This way, our model simulations will examine to what extent the observed variations in experimental findings can be explained merely by the discourses used.

Experimental Findings on Italian Pronouns

Italian, like Spanish, Catalan, Romanian, Arabic and many other languages, is a language that knows null pronouns. This means that in many cases, a grammatical subject does not have to be realized but can be omitted, creating a null pronoun or null subject (e.g., *corre* 'he/she/it runs'). In addition, Italian has overt pronouns such as *lui* 'he', through which a grammatical subject is explicitly realized. Null

pronouns are generally used to refer to the most prominent referent in a discourse, whereas overt pronouns are generally used to refer to a different referent (Carminati, 2002).

Importantly, the influence of discourse prominence on the processing and interpretation of null and overt pronouns in Italian indicates that differences between the findings of different experimental studies (Tsimpli et al., 2004; Vogelzang et al., 2019) may be the result of the different contexts being used. The results of Tsimpli et al. (Figure 1), obtained using a picture selection task, show a much weaker preference for subject antecedents for null pronouns than would be expected based on Carminati's (2002) classical findings.
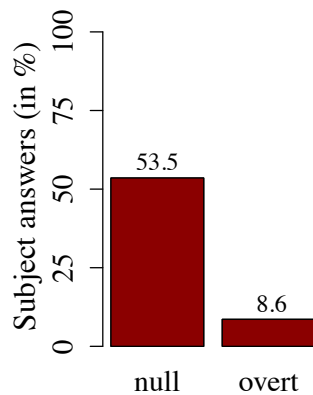

Figure 1. Experimental data from Tsimpli et al. (2004) on the interpretation of null and overt pronouns in Italian.[1]

As stated, this may in part be due to the specific discourses used in the experiment. More specifically, Tsimpli et al. (2004) used short discourse contexts with two clauses such as in (1):

(1) La mamma dà un bacio alla figlia,
    mentre lei/Ø si mette il cappotto.

   "*The mother kisses the daughter,*
   while she/Ø puts on her coat."

In the example context in (1), the prominence of the two referents in terms of frequency is the same, although their grammatical roles differ. The grammatical subject of the first clause is only mentioned once and therefore it is conceivable that this character has not been clearly established as the most prominent referent, which is generally also the discourse topic.

Vogelzang et al. (2019), in contrast, found, using a referent selection task, a strong subject preference for null pronouns (Figure 2).
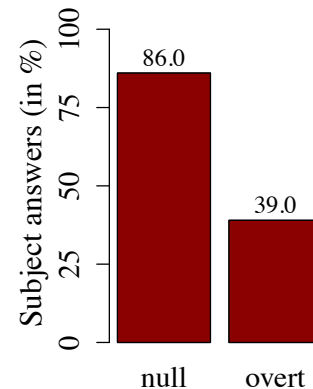

Figure 2. Experimental data from Vogelzang et al. (2019) on the interpretation of null and overt pronouns in Italian.[2]

The longer discourse contexts with three clauses used by Vogelzang and colleagues as in (2) make one referent, which occurs as the grammatical subject twice, much more prominent than the other referent:

(2) Il riccio compra della moquette per il soggiorno.
    Ieri il riccio ha raccontato al topo una storia,
    mentre lui/Ø si annoiava davanti alla tv.

   "*The hedgehog is buying some carpet for the living*
   room. Yesterday the hedgehog has told the mouse a
   story, while he/Ø was bored in front of the TV."

Thus, discourse prominence in terms of recency, frequency, and grammatical role differs between these two experiments, as do the interpretations of participants. In the following sections, we will examine the processing of these discourses and of the pronouns within these discourses more closely with a computational cognitive model.

---

[1] The experiments of Tsimpli et al. (2004) were mainly focused on language attrition, but this paper will only look at their data on monolingual adults. Additionally, Tsimpli et al. (2004) looked at both forward and backward anaphora, but only backward anaphora will be taken into account in this analysis. Finally, the picture selection task of Tsimpli et al. (2004), in which pictures with two characters and the mentioned action were shown, contained a third answer option, namely an 'other' character, not mentioned in the discourse. To allow for a better comparison between the two studies, this option was not taken into account in the current description of the data, and the percentages of answers were adjusted accordingly.

[2] Vogelzang et al. (2019) included a third subject condition in their experiment, namely a full noun phrase. To allow for a better comparison between the two studies, this option was not taken into account in the current description of the data. In the referent selection task that Vogelzang et al. (2019) used, pictures of the two mentioned characters were shown without any action.

Cognitive Model

The cognitive model we will use to examine the influence of different discourses on Italian pronoun processing is implemented in the cognitive architecture ACT-R (Anderson, 2007; Anderson et al., 2004), which pre-specifies constraints on human cognition and processing. The current model[3] builds on the pronoun interpretation models of Hendriks et al. (2007), Van Rij et al. (2010) and Vogelzang (2017), the latter of which has previously been used to fit the data of Vogelzang et al. (2019) on Italian pronoun interpretation. The most relevant and important aspects of the model will be discussed here; for a more elaborate discussion of the mechanisms used in the model we refer you to Van Rij et al. (2010) and Vogelzang (2017).

The model uses a constraint-based bi-directional approach to pronoun processing, in which a listener reasons both from his/her own perspective and from the perspective of the speaker (cf. Blutner, 2000; Hendriks & Spenader, 2006; previously implemented in a computational cognitive model in Hendriks et al., 2007; Van Rij et al., 2010). The idea that the interpretation of ambiguous referring expressions requires listeners to reason about alternative forms from the perspective of the speaker was originally proposed by Hendriks and Spenader (2006) for Dutch. They formulated their perspective-taking account within the constraint-based linguistic framework Optimality Theory (OT; Prince & Smolensky, 2004), in which hierarchically ranked constraints are used to determine the optimal meaning for an input form (in interpretation) or the optimal output form for a meaning (in production).

In order to expand this approach to Italian, the model discussed here incorporates a constraint regulating the interpretation of null pronouns, stating that null pronouns refer to the discourse topic (similarly to overt pronouns in non-null subject languages, cf. Beaver, 2004; Grosz, Joshi, & Weinstein, 1995; Hendriks, Englert, Wubs, & Hoeks, 2008; Van Rij et al., 2013). As a consequence, a listener would reason that if a speaker would have wanted to refer to the discourse topic, they would have used a null pronoun. So, if the speaker used an overt pronoun instead, they likely wanted to refer to something other than the discourse topic. The other constraints incorporated in the model are based on referential economy (Burzio, 1998), and reflect the idea that speakers prefer to be as efficient as possible and therefore prefer shorter linguistic expressions such as null pronouns over longer linguistic expressions such as overt pronouns. In interpretation, these constraints referring to referential economy will not be relevant. However, because the model additionally reasons about alternative forms from the perspective of the speaker, they will be used to reason about which form a speaker would have most likely used for reference to the topic (null pronoun) or reference to a non-topical referent (overt pronoun).

---

[3] full model code is available at
https://sites.google.com/view/margreetvogelzang/experiment-files

In addition to constraints, the discourse also affects pronoun processing. More specifically, the model uses a discourse processing component based on the model of Van Rij et al. (2013), in which the prominence of a referent in discourse is determined by the standard ACT-R mechanisms of activation (which is based on its frequency and recency in the discourse) as well as an additional 'boost' that represents additional activation for referents associated with the grammatical subject (set to 1.0). This will most likely make the previous grammatical subject the referent with the highest activation, which can thus be considered the discourse topic. In Van Rij et al. (2013), this grammatical subject boost is argued to represent working memory (WM) capacity, as differences in this activation boost to associated information can account for individual differences in WM capacity (Daily, Lovett, & Reder, 2001).

Results

In this section, different simulations will be described for the processing and interpretation of Italian subject pronouns in the studies of Vogelzang et al. (2019) and Tsimpli et al. (2004). Importantly, every round of simulations uses the same model to simulate the data from both studies, only varying the input (the discourses) presented to the model. Every simulation will differ slightly due to pre-defined mechanisms of the cognitive architecture, such as varying latencies when retrieving information from memory.

In line with the original experiments, the model was run on 32 discourses (items) for 40 simulations (participants) to simulate the experiment of Vogelzang et al. (2019). The model was run on 10 discourses (items) for 20 simulations (participants) to simulate the experiment of Tsimpli et al. (2004). Half of the discourses contained a null pronoun and half contained an overt pronoun.

Simulation 1

For the first round of simulations, the activation boost given to grammatical subjects, representing WM capacity, and the number of practice items presented to the model were kept the same as in Vogelzang (2017); the activation boost was set to 1.0 and the number of practice items to 2000. The results of the simulation for the discourses of Vogelzang et al. (2019) and Tsimpli et al. (2004) are presented in Figures 3 and 4, respectively.

As can be seen in Figure 3, the model data shows very similar interpretational preferences to the experiment data of Vogelzang et al. (2019). The model shows different interpretations, however, compared to the experiment data of Tsimpli et al. (2004), although the tendency of null pronouns referring to the subject and overt pronouns referring to the non-subject is present in both the model data and the experiment data; this can be seen in Figure 4. Specifically, the model shows a higher percentage of subject interpretations for both null pronouns and overt pronouns than the participants in the experiment.
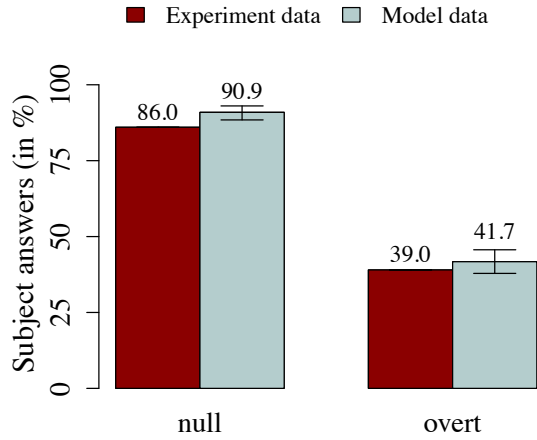
Figure 3. Experimental data from Vogelzang et al. (2019) and model output on Italian pronoun interpretation.
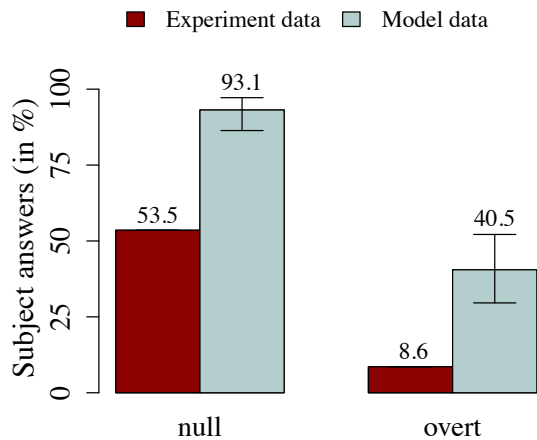

Figure 4. Experimental data from Tsimpli et al. (2004) and model output on Italian pronoun interpretation.

One possible explanation for this could be that the activation boost that is associated with the grammatical subject is constant, i.e. is equally strong no matter how many times a referent is mentioned in the grammatical subject position. Since making the topic of a discourse more clear aids pronoun interpretation (Spenader et al., 2009), however, it is more likely that this activation boost is gradually increasing with each mention rather than a consistently large boost. This possibility will be explored in the next simulation.

Simulation 2

When examining the discourses used in the experiments (see (1) and (2)) more closely, it can be seen that the same character is mentioned in the grammatical subject position once in the discourses used by Tsimpli et al. (2004) compared to twice in the discourses used by Vogelzang et al. (2019). We will now assume, following findings of Spenader et al. (2009), that the discourse topic becomes more clear the more consistent a discourse is, so the more often a certain referent occurs in the grammatical subject

position. Figure 5 shows a proposed stepwise activation boost according to the following function:

$$boost = n^2/10$$

in which n is the number of consecutive occurrences of a referent in the grammatical subject position within a discourse.
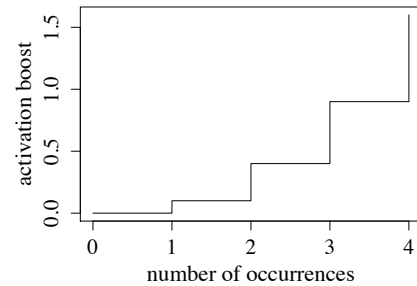

Figure 5. Effect of the number of occurrences of a referent in the grammatical subject position on the activation boost.

Using this function to calculate the boost in activation given to grammatical subjects, new simulations were run for the experiments. The results show that the model can still account for the interpretational preferences of participants in the experiment of Vogelzang et al. (2019) with longer discourses (Figure 6) and that the same model can now also account for the interpretation of null pronouns found in the experiment of Tsimpli et al. (2004) (Figure 7). Notably, the predictions of the model differ considerably based on the discourse used. However, although the interpretational preference of overt pronouns referring to the non-subject can be seen in both the model data and the experiment data, the actual interpretation of overt pronouns in the data of Tsimpli and colleagues is not reproduced by the model. Possible explanations for this are discussed in the next section.
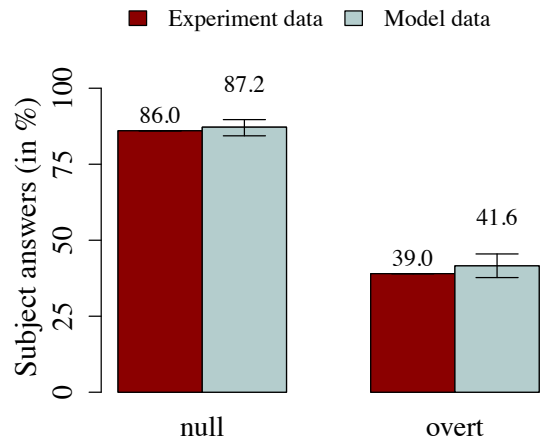

Figure 6. Experimental data from Vogelzang et al. (2019) and model output on Italian pronoun interpretation.
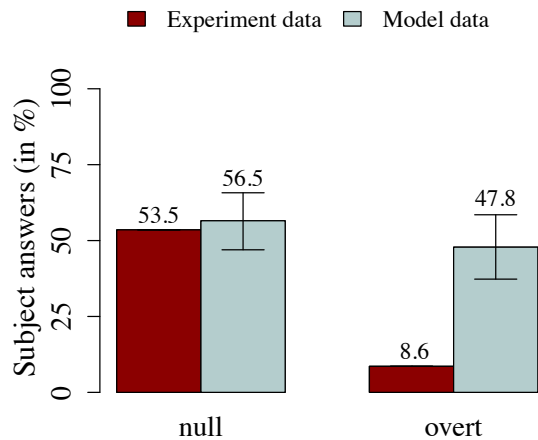
Figure 7. Experimental data from Tsimpli et al. (2004) and model output on Italian pronoun interpretation.

Discussion

In this paper, we investigated with the help of a cognitive model to what extent differences in discourse context can explain variations in experimental results on Italian pronoun interpretation. We examined the discourses used in experiments by Tsimpli et al. (2004) and Vogelzang et al. (2019), which differ in the number of clauses used and the number of times that the referents are mentioned. We simulated pronoun processing in these discourses using an ACT-R model which built on previous cognitive models of pronoun processing (Hendriks et al., 2007; Van Rij et al., 2010; Vogelzang, 2017). The model used an activation boost to keep referents associated with the grammatical subject of the previous sentences active in memory as the discourse topic.

The results from the first simulation showed that the data of Vogelzang et al. (2019) could be simulated accurately, but the data of Tsimpli et al. (2004) could not be accounted for. In the second model simulation, a function rather than a constant was used to determine the activation boost given to the referent associated with the grammatical subject of a sentence to reflect a gradual increase of certainty about the discourse topic, in line with experimental evidence from Spenader et al. (2009). The results showed that this simulation can account for the data of Vogelzang et al. (2019) and can partially account for of the data of Tsimpli et al. (2004). More specifically, null pronouns were accurately predicted to refer to the subject slightly more than half of the time. This indicates that a gradual increase in the activation of a recurring grammatical subject is a viable possibility, and something that should be seriously considered when modeling processes in which prominence in discourse plays an important role or when designing (linguistic) experiments. However, Tsimpli et al.'s (2004) findings for overt pronouns, which referred to the subject less than 10% of the time, were not replicated by the model. We will discuss three possible causes for this below.

First, Serratrice (2007) notes that Tsimpli et al.'s experimental results are not in line with the classical null-pronoun-refers-to-the-subject findings of, a.o., Carminati (2002). Additionally, the results from Tsimpli et al. (2004) show a stronger interpretation preference for overt pronouns than for null pronouns, which is also in contrast to Carminati (2002). Serratrice (2007, p. 233) suggests that this may be caused by the fact that "*In Carminati's experiment the subject and the object interpretation were presented in written form, while in this study the two alternatives were presented pictorially*". However, in Vogelzang et al.'s (2019) study the answers were also presented pictorially. Therefore, the pictorial presentation can not explain the differences in interpretation between Vogelzang et al.'s (2019) study, whose results are in line with Carminati (2002), on the one hand and the Tsimpli et al.'s (2004) study on the other hand. Nevertheless, there were some differences between the tasks that should be examined in more detail in future research, such as the type of pictures presented (pictures with actions, Tsimpli et al. vs. pictures without actions, Vogelzang et al.) and the number of answer possibilities (3 vs. 2, respectively).

A second possible explanation for the strong preference of overt pronouns to refer to the non-subject in the study of Tsimpli et al. (2004) could be related to aspects of the discourse that were not taken into account in the model. Discourse prominence was taken into account in terms of recency and frequency though standard ACT-R activation mechanisms, but it is possible that recency plays a bigger role than that, as it is known to influence the accessibility of a referent (Arnold, 1998; Givón, 1983). When examining the discourses in (1) and (2), we can see that Tsimpli et al. introduced the second referent at the end of the pre-critical clause, whereas Vogelzang et al. provided linguistic content (in (2) a direct object) in between the second referent and the end of the pre-critical clause. Thus, it may be possible that very recent referents hold a special status, which was not taken into account in the model.

A third possible explanation stems from the observation that for both null and overt pronouns participants showed fewer subject interpretations in the study of Tsimpli et al. (2004) compared to the study of Vogelzang et al. (2019). This might be related to the verbs used in the discourses, as verb bias or implicit causality (Garvey & Caramazza, 1974) can influence which referent (previous subject or non-subject) will likely be the actor in the continuation of the discourse. Similarly, an event-structure bias (Stevenson, Crawley, & Kleinman, 1994) of verbs could trigger the preference to continue the story with the end state of an action, which was the goal (non-subject) rather than the source (subject), of the verb in the pre-critical sentences. The model did not take any verb bias into account; potential effects of the verbs could be tested using a sentence completion task with the verbs used in both experiments.

Concluding, we investigated to what extent the variations in experimental results on the interpretation of Italian subject pronouns can be explained by the different discourses used in the experimental studies. Our simulations suggest that the discourse contexts used in the experiments

crucially influence the interpretation of Italian subject pronouns. Thus, discourse prominence in terms of recency, frequency, and grammatical role seem to play an important role in the processing and interpretation of pronouns, which has to be taken into account when interpreting experimental results. Nevertheless, the model was not able to account for all data, and further research, both on the processing of discourse and on the influence of the specific task, is needed to investigate variations in experimental results on Italian pronoun interpretation.

References

Anderson, J. R. (2007). *How Can the Human Mind Occur in the Physical Universe?* New York: Oxford University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of mind. *Psychological Review*, *111*(4), 1036–1060. http://doi.org/10.1037/0033-295X.111.4.1036

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

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

Beaver, D. (2004). The optimization of discourse anaphora. *Linguistics and Philosophy*, *27*(1), 3–56. http://doi.org/10.1023/B:LING.0000010796.76522.7a

Blutner, R. (2000). Some aspects of optimality in natural language interpretation. *Journal of Semantics*, *17*(3), 189–216. http://doi.org/10.1093/jos/17.3.189

Burzio, L. (1998). Is the Best Good Enough? Optimality and Competition in Syntax. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, & D. Pesetsky (Eds.), *Anaphora and Soft Constraints*. Cambridge, MA: MIT Press.

Carminati, M. N. (2002). *The processing of Italian subject pronouns* (Doctoral dissertation). Amherst, MA: University of Massachusetts.

Daily, L. Z., Lovett, M. C., & Reder, L. M. (2001). Modeling individual differences in working memory performance: a source activation account. *Cognitive Science*, *25*(3), 315–353. http://doi.org/10.1016/S0364-0213(01)00039-8

Garvey, C., & Caramazza, A. (1974). Implicit causality in verbs. *Linguistic Inquiry*, *5*(3), 459–464.

Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic Continuity in Discourse: A Quantitative Cross-language Study*. Amsterdam: John Benjamins Publishing Company.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics*, *21*(2), 203–225.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, *69*(2), 274–307.

Hendriks, P., Englert, C., Wubs, E., & Hoeks, J. (2008). Age differences in adults' use of referring expressions. *Journal of Logic, Language and Information*, *17*(4), 443–466. http://doi.org/10.1007/s10849-008-9065-6

Hendriks, P., & Spenader, J. (2006). When Production Precedes Comprehension: An Optimization Approach to the Acquisition of Pronouns. *Language Acquisition*, *13*(4), 319–348. http://doi.org/10.1207/s15327817la1304_3

Hendriks, P., Van Rijn, H., & Valkenier, B. (2007). Learning to reason about speakers' alternatives in sentence comprehension: A computational account. *Lingua*, *117*(11), 1879–1896. http://doi.org/10.1016/j.lingua.2006.11.008

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

Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford, UK: Blackwell Publishing. http://doi.org/10.1002/9780470759400

Serratrice, L. (2007). Cross-linguistic influence in the interpretation of anaphoric and cataphoric pronouns in English–Italian bilingual children. *Bilingualism: Language and Cognition*, *10*(3), 225–238. http://doi.org/10.1017/S1366728907003045

Spenader, J., Smits, E.-J., & Hendriks, P. (2009). Coherent discourse solves the pronoun interpretation problem. *Journal of Child Language*, *36*(1), 23–52. http://doi.org/10.1017/S0305000908008854

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

Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism*, *8*(3), 257–277. http://doi.org/10.1177/13670069040080030601

Van Rij, J., Van Rijn, H., & Hendriks, P. (2010). Cognitive architectures and language acquisition: A case study in pronoun comprehension. *Journal of Child Language*, *37*(3), 731–766. http://doi.org/10.1017/S0305000909990560

Van Rij, J., Van Rijn, H., & Hendriks, P. (2013). How WM load influences linguistic processing in adults: a computational model of pronoun interpretation in discourse. *Topics in Cognitive Science*, *5*(3), 564–80. http://doi.org/10.1111/tops.12029

Vogelzang, M. (2017). *Reference and Cognition: Experimental and computational cognitive modeling studies on reference processing in Dutch and Italian* (Doctoral dissertation). Groningen: University of Groningen.

Vogelzang, M., Foppolo, F., Guasti, M. T., Van Rijn, H., & Hendriks, P. (2019). Reasoning about alternative forms is costly: Comparing the processing of null and overt pronouns in Italian using pupillary responses. *Discourse Processes*. http://doi.org/10.1080/0163853X.2019.1591127