Processing Overt and Null Subject Pronouns in Italian: a Cognitive Model

Margreet Vogelzang (margreet.vogelzang@rug.nl)
Center for Language and Cognition Groningen, University of Groningen
P.O. Box 716, 9700 AS Groningen, the Netherlands

Petra Hendriks (p.hendriks@rug.nl)
Center for Language and Cognition Groningen, University of Groningen
P.O. Box 716, 9700 AS Groningen, the Netherlands

Hedderik van Rijn (hedderik@van-rijn.org)
Departments of Experimental Psychology & Psychometrics and Statistics, University of Groningen
Grote Kruisstraat 2/1, 9712 TS, the Netherlands

Abstract
In this paper, we present a cognitive model that simulates the processing of subject pronouns in Italian. The model is implemented in the cognitive architecture ACT-R and uses hierarchically ranked constraints to select the most likely referent of a pronoun. When this model is combined with a measure of accessibility in discourse and a processing time limit imposed by the speed of natural language production, the model accounts for novel empirical data of the interpretation of null as well as overt subject pronouns in Italian. The model generates concrete predictions on the basis of variations in cognitive capacities, which can be tested in subsequent experiments.

Keywords: pronoun interpretation; cognitive modeling; null subjects; constraint-based modeling

Introduction
Anaphoric pronouns such as he and she are commonly used to refer to entities that were previously mentioned in the discourse. Such anaphoric expressions can range from multiple-word full noun phrases (NPs) to anaphoric pronouns without an overt form (null subjects). Anaphoric pronouns are potentially ambiguous and have to be resolved by the listener in order to be interpreted correctly.

Italian, like Spanish, is a null subject language. In null subject languages, a subject pronoun can be expressed overtly (e.g., "he runs"), or it can be omitted (e.g., "runs"), resulting in a null subject. Compared to non-null subject languages such as English, Italian thus has an additional subject form. The availability of this additional subject form may influence the processing of an overt pronoun in these languages. When processing a sentence, the form of the subject provides information about what the intended referent of the subject is. For example, short forms tend to refer to referents that are highly salient in the discourse, whereas longer forms tend to refer to less salient referents. Since Italian allows for two different types of pronominal form in subject position, the use of one pronoun type over the other may inform a listener about the intended referent. Indeed, Carminati (2002) has found that null subjects generally refer to the discourse topic and overt subject pronouns generally refer to a non-topical antecedent.

However, speakers are not always consistent in their interpretation of null and overt pronouns, and show a substantial amount of variation. A satisfactory explanation for this variation in the interpretation of null and overt subject pronouns is still lacking, but interpretations have been shown to vary on the basis of several discourse factors, such as pragmatic plausibility (Carminati, 2002), implicit causality of the verb (Fedele & Kaiser, 2014), and recency of competitor antecedents (Sorace & Filiaci, 2006). So, the interpretation of Italian subject pronouns is influenced by discourse factors, but how these discourse factors interact and in what way they influence processing is not clear yet.

Most of the studies mentioned above investigated the effects of a single discourse factor on pronoun interpretation. In this study, we aim to provide an account of the interaction between different factors in pronoun processing. In addition, we wish to account for the observed variation in the interpretation of overt and null pronouns in Italian. We will do so through a combination of experimental investigation and cognitive modeling. In a cognitive model, predictions and assumptions about which factors influence pronoun processing are implemented computationally. This allows for the explicit testing of hypotheses and for the development of predictions that can be tested in future experiments.

In this paper, we present a cognitive model of the processing and interpretation of subject pronouns in Italian. The model builds on the earlier work by Hendriks, Van Rijn and Valkenier (2007) and Van Rij, Van Rijn and Hendriks, (2010), and is implemented in the cognitive architecture ACT-R (Anderson, 2007). The model uses linguistically motivated constraints to select the most likely referent of a pronoun. Different sources of variation in the interpretation of null as well as overt subject pronouns in Italian will be investigated using the model. The model is validated based on empirical data, which we will first discuss.

Experimental data
With a referent selection task, we examined how native Italian adults (n=40) interpret anaphoric expressions in...
In a lab-based setting, participants were presented with 48 short auditory stories of three clauses each. The final clause of each story contained one of three anaphoric subject forms: A full NP such as the dog as an unambiguous baseline condition, a null subject, and the overt subject pronouns lui (‘he’) or lei (‘she’). Consider the following sample story:

1. Il cane va a fare un viaggio in Germania.
   *The dog is going on a trip to Germany.*
2. Ieri sera il cane ha invitato il gatto a viaggiare insieme,
   *Last night the dog has invited the cat to travel together,*
3. mentre Ø/lui/il cane si lavava prima della partenza.
   *while Ø/he/the dog washed himself before the departure.*

Participants’ interpretations of the anaphoric subject in the final sentence of the story were determined on the basis of referent selection questions such as ‘Who washed himself?’ Importantly, each story featured two characters that participants could select as the referent of the encountered anaphor. The first character, the dog, is the sentential subject in the first two sentences, is the first-mentioned character, and holds the same grammatical role as the anaphoric subject form in the final sentence. Therefore, this character is the most prominent character in the discourse and the discourse topic. The second character in the story, the cat, is introduced in the second clause of the story as a direct or indirect object, and is therefore less prominent than the dog (Ariel, 1990). This second character will be referred to as the non-topical antecedent.

The results of the experiment are plotted in Figure 1. We coded the results as the percentage of topic continuations, which is the percentage of answers in which the discourse topic was selected as the referent of the subject. The baseline condition with a full NP was answered correctly as referring to the topic in 99% of the cases. This is not surprising, as the NPs unambiguously referred to the topic (e.g., the dog in the sample story). Null subjects were often, but not always, interpreted as referring to the topic (86% of the cases). Overt subject pronouns were not interpreted consistently as referring to either the topic or the non-topical referent. Specifically, they were interpreted as referring to the topical referent in 39% of the cases, and to the non-topical referent in 61% of the cases.

Since full NPs were unambiguous in the experiment and were nearly always interpreted correctly, we will focus on pronouns in the cognitive model. Thus, the model will simulate the processing and interpretation of null pronouns and overt pronouns in subject position.

**A constraint-based approach**

A promising approach to study the interaction of various factors in pronoun interpretation is a constraint-based approach. The constraint-based linguistic framework Optimality Theory (OT; Prince & Smolensky (2004), see for earlier models based on this approach Hendriks et al. (2007) Misker & Anderson (2003), and Van Rij et al., (2010)) accounts for the interaction between linguistic factors in production and interpretation through its mechanism of optimization over potential forms or meanings. In addition, OT is able to account for speaker-listener coordination in language use by bidirectional optimization (Blutner, 2000), which can be seen as a formalization of the process of perspective-taking (Van Rij, Van Rijn, & Hendriks, 2013).

In OT, the grammar consists of a set of hierarchically ranked constraints, with each constraint being either violated or not. In production, an input meaning is mapped onto potential forms for expressing that meaning. On the basis of the constraints, from the set of potential forms the optimal output form is determined that satisfies the constraints of the grammar best. Likewise, in interpretation the optimal output meaning for a given input form is the meaning that satisfies the constraints best. In the case of pronoun interpretation, a listener encountering a pronoun will generate potential interpretations for this pronoun. By applying constraints on pronoun interpretation, the listener will be able to determine the optimal interpretation of the pronoun. Crucially, constraints in OT may conflict. In case of a conflict, the higher-ranked constraint has priority over the lower-ranked constraint.

Observing that English-speaking children’s production of object pronouns seems to be ahead of their interpretation of the same pronouns, Hendriks and Spenader (2005/2006) argue that mapping the input form onto the optimal meaning by uni-directional optimization does not suffice for the correct interpretation of object pronouns. Rather, they argue that listeners must also consider the perspective of the speaker. In OT, this process of perspective-taking can be modeled as bi-directional optimization. Bi-directional optimization is thus conceptually related to Gricean reasoning, according to which language users are cooperative and assume the other to be cooperative as well (Grice, 1975). In a bi-directional approach to pronoun interpretation, listeners start with uni-directional optimization from their own perspective, determining the optimal interpretation for the encountered pronoun. In some cases, however, the constraints still allow for several meanings. In these cases a second, bi-directional...
optimization step is necessary. In this step, the listener randomly selects one of the potential meanings and then takes the perspective of the speaker in order to determine if a speaker would indeed have used the encountered pronoun for the selected meaning. If the optimal form from the speaker’s perspective is identical to the encountered form, the selected meaning is bi-directionally optimal. If, on the other hand, the optimal form from the speaker’s perspective is different from the encountered form, the selected meaning is discarded and another meaning is selected.

A number of models have successfully implemented this constraint-based bi-directional approach to simulate, for example, the interpretation of subject pronouns in discourse (Van Rij et al., 2013) and the acquisition of object pronouns and reflexives (Hendriks, Van Rijn, & Valkenier, 2007; Van Rij et al., 2010).

A cognitive model of subject pronoun processing in Italian

We present a model of Italian subject pronoun processing implemented in the cognitive architecture ACT-R (Anderson, 2007). This architecture is a useful tool to model and explain human behavior and cognition. It increases the psychological plausibility of models built in this framework as it is constructed to reflect assumptions about human cognition. Because it is a cognitive architecture, ACT-R allows for the investigation of specific cognitive capacities. On the basis of individual and situation-dependent variations in cognitive capacities, the model can generate predictions about performance in other tasks and by other populations.

The presented model is adapted from the object pronoun resolution model of Van Rij et al. (2010). Importantly, the underlying mechanisms of the current model are identical to the constraint-based approach and perspective-taking mechanism of Van Rij et al. (2010). Moreover, all constraints used in the model are based on previous research.

The task of the model is to determine the referent of a pronoun presented to the model. The model uses the following, hierarchically ordered, linguistic constraints:

1. Null subjects refer to the topic
2. Avoid overt pronouns

The first constraint restricts the use and interpretation of null subjects, stating that they must refer to the current discourse topic (adapted from Van Rij et al., 2013). No comparable constraint is used for overt pronouns, as we will assume that their interpretation is derived from the interpretation of null pronouns (see below). The second constraint is adopted from Hendriks and Spenader (2005/2006) and Van Rij et al. (2010) and is only relevant in production, when taking the perspective of the speaker. This constraint states that overt pronouns should be avoided. The constraint is based on referential economy principles (Burzio, 1998), and indicates that a speaker, when given the choice, would rather use a null subject than an overt subject pronoun, because a null subject is more economical to produce.

The processing of pronouns proceeds in four steps in the model, the last three of which are shown in Figure 2:

(a) Determining the discourse topic
(b) Interpretation (uni-directional optimization)
(c) Perspective-taking (bi-directional optimization)
(d) Evaluation

Having heard the first two clauses of a story, the listener has already encountered the two different characters. To simplify the model, we did not model the processing of the first two clauses of the story. Rather, the model starts with two referents being available in memory, each with a set activation level. The activation of a character in memory represents the accessibility or saliency of that character in the discourse (Ariel, 1990; Givón, 1983). When determining the discourse topic, the character with the highest activation in memory is retrieved. This will generally be the topic the

---

Figure 2: After the discourse topic is determined (a, not shown), the processing steps are shown of (b) interpreting the pronoun (uni-directional optimization), (c) perspective-taking (bi-directional optimization), and (d) evaluating the output by comparing the optimal output form in (c) with the input form in (b). Null subjects are interpreted correctly as referring to the topic already after interpretation step (b), whereas overt subject pronouns are still ambiguous at this point. To interpret overt subject pronouns, the perspective-taking step in (c) is therefore required in order to arrive at the correct interpretation: reference to a non-topic rather than the topic. In the diagram, the possible candidate forms or meanings are shown in boxes. A grey box indicates that the candidate violates the current constraint, and therefore it is not considered as an optimal candidate further.

---

The model code can be found at
let.webhosting.rug.nl/~vogelzang/experiments.html
dog in the sample story, since this character occurred more often than the other character and was also the sentential subject. However, when a character is retrieved from memory its activation is mediated by noise (see Anderson (2007) for more details and formulas), and so errors can occur. Specifically, the model retrieves the correct discourse topic about 90% of the time, and identifies the non-topical referent as the discourse topic about 10% of the time.

Once the discourse topic has been determined, the model will start with interpreting the pronoun from the perspective of the listener (corresponding to step b in Figure 2). Either a null pronoun or an overt pronoun has been encountered and needs to be interpreted. In order to do this, both possible interpretations of the pronoun (in the model: the topic referent and the non-topic referent) are retrieved as possible meanings. These meanings are evaluated on the basis of the constraints listed above. If the input is a null subject, constraint [1] requires a null subject to refer to the topical antecedent, and thus the topic is selected as the optimal meaning. When a meaning has been selected, the model continues with the bi-directional optimization step in the model. Alternatively, if the input is an overt pronoun, the first constraint does not restrict its interpretation. In that case, the second constraint is used. This constraint, ‘Avoid overt pronouns’, does not distinguish between potential meanings for an overt pronoun. As a consequence, the overt pronoun remains ambiguous between reference to the topic and reference to a non-topical referent. The model will now randomly select one of these two meanings and continue with the next step of the resolution process.

In the next step of the model (step c in Figure 2), the model takes the perspective of the speaker. The optimal meaning of the previous step is the input in this step. Now, the model will determine the optimal form for this input meaning. In other words, the model determines whether the meaning selected by the listener would indeed be referred to by a speaker with the form encountered. The same constraints are used as in the first step and thus the first constraint to be retrieved is again constraint [1]. If the input is a topic, then this constraint does not restrict the use of a null or overt subject pronoun. In that case, constraint [2] is retrieved. This constraint states that a speaker prefers the use of a null pronoun to an overt pronoun. Thus, the optimal form to refer to the topic is a null pronoun. Alternatively, if the input is a non-topic, the first constraint disallows the use of a null pronoun. In that case, the optimal form for referring to a non-topic is an overt pronoun.

So far, the uni-directional optimization step of pronoun interpretation provided the model with the optimal meaning for the input, and the bi-directional optimization step of pronoun production provided the model with the optimal form for the selected meaning. In the final step of the model (step d in Figure 2), the optimal form is compared to the original input. If the evaluation shows that the optimal form is equal to the original input, the model assumes that the selected optimal meaning is indeed the correct one. This is the case for null pronouns, where the optimal meaning is the topic and the optimal form is again a null pronoun. The model will then yield the discourse topic as the output. Note that which referent is the discourse topic is not determined by the processes of uni-directional and bi-directional optimization, but is determined in the first step. If, on the other hand, the evaluation shows that the optimal form is not equal to the original input, the selected meaning is revised.

In the case of an overt subject pronoun, the optimal meaning was selected randomly after the uni-directional optimization step. If a non-topic was selected, this would be the correct choice, as the optimal form for expressing a non-topic is an overt subject pronoun. On the other hand, if a topic is selected, the bi-directional optimization step will result in a null subject as the optimal form for referring to a topical referent. Since the original input was an overt pronoun, the evaluation will show that the original input differs from the optimal form. When this happens, the optimal meaning of the pronoun must be revised from the referent that is the discourse topic to the non-topical referent. After this final step has been completed, the model will yield as its output the referent that is selected as the optimal meaning in this final step.

If the model would have unlimited time and resources, it would interpret a null pronoun as referring to the discourse topic, and an overt pronoun as referring to the non-topical referent. However, due to the speed of natural language production the model only has limited processing time, and therefore errors can occur (processing speed limit based on Van Rij et al., 2010). If the model does not have sufficient time to run all four processing steps within the available time, it will select the optimal meaning at that point in the process. If no optimal meaning has been found so far, the model will guess the answer. So if the model only had time to complete the uni-directional optimization step but not the bi-directional optimization step, the optimal meaning selected for a null pronoun will be correct, whereas the optimal meaning for an overt pronoun is incorrect roughly half of the time due to the model’s random selection of a meaning at the end of the uni-directional step.

Results of the model

The model was run for 40 simulations, simulating 40 participants. No explicit individual differences were modeled, but small differences can occur between simulations. In each simulation, the model was presented with 1000 items (half presenting a null pronoun and half presenting an overt pronoun), of which the first 968 were used as practice items and the last 32 were used for the test phase. The practice items were included to provide the model with prior experience resolving pronouns.

By means of this experience, the model develops to perform the resolution process more efficiently. The processing speed of a simulated participant can increase through experience because of two mechanisms: activation and production compilation. When facts are retrieved from memory more often, their activation increases and this makes it easier and faster to retrieve these facts. Production
compilation is a mechanism that allows the model to combine multiple production rules (small decision steps in the model) into one, so that fewer steps have to be taken when performing a task that has been performed frequently before. Since every production rule takes a fixed amount of time in ACT-R, production compilation also speeds up the model.

In all practice and test items, participants were given 0.6 seconds to process and interpret the encountered pronoun. This time limit was given to simulate reasonably slow speech of 100 words per minute (Wong, 2015). When the time limit was reached, the model was not allowed to execute any more processing steps.

The answers to the test items, and the comparison to the experimental data, are shown in Figure 3. The figure shows the mean percentage of responses indicating a discourse topic interpretation for null and overt subject pronouns as provided by the model, and as taken from the experimental data with real participants. As Figure 3 shows (given the limited number of data points, we have refrained from reporting explicit model comparison measures), the model accounts for the general trends in the data that were described earlier: null pronouns are often (but not always) interpreted as referring to the discourse topic, and overt pronouns vary in their interpretation, but are most often interpreted as referring to the non-topical referent.

![Figure 3: The experimental data and the model output of the interpretations of null pronouns and overt pronouns in subject position.](image)

**Sources of variation**

In our explanation of the cognitive model, we mentioned several features that influence model’s responses. Here we will discuss these features and explain how they contribute to the variation in interpretations shown by the model.

First of all, due to noise, the model makes mistakes when identifying the discourse topic. These topicality mistakes are also expected to occur in real life, where distractions or insufficient working memory capacity influence how well the discourse is recalled. In the model, most of the answers indicating reference of a null pronoun to the non-topical referent are made because of mistakes in identifying the topic. Such mistakes also occur when interpreting overt pronouns, but these are less visible because overt pronouns vary more in their interpretation. However, these types of mistakes do not account for the difference in the percentages of topical interpretations between null and overt pronouns.

A second source of variation, one that applies differently to null and overt pronouns, has to do with time constraints on processing: if the model has not finished all processing steps within the allotted time, a meaning that has been found to be optimal in an intermediate processing step is taken as the optimal meaning. If no such meaning is available, the meaning is guessed by the model. This guessing behavior is not implausible, as the human participants in the experiment also had to choose between the two characters, even if they did not know the answer. The time constraint accounts for the difference in variation between null and overt subject pronouns: null pronouns do not require the bi-directional optimization step to arrive at the optimal meaning.

**Predictions of the model**

Based on the properties of the cognitive model a number of predictions can be formulated, which can be tested in subsequent experiments. The predictions of the model have to do with identifying the correct discourse topic and with speed of processing, which were argued to be sources of variation. A very general prediction concerns children, whose linguistic experience in pronoun interpretation (affecting processing speed) and working memory capacity (affecting discourse processing) are lower than that of adults. Therefore, children are expected to show more variation than adults when interpreting pronouns in discourse. However, this is not a very specific prediction.

More specific and testable predictions can be made by looking at the interaction between the sources of variation in the model. One source of variation was the incorrect identification of the discourse topic. If the discourse topic is clearer (that is, much more prominent than the other referent), fewer mistakes will be made in the interpretation of null pronouns. However, because retrieval of the non-topical referent is necessary for the interpretation of overt pronouns, and this retrieval becomes more difficult if the activation of that referent is lower, it is predicted that overt pronouns will also be interpreted more often as referring to the discourse topic when the activation of this topic is much higher than the activation of the non-topical referent.

Additionally, the model predicts that when working memory is insufficient while processing speed is sufficient (for example, when carrying out a parallel, secondary task that places competing demands on working memory), more mistakes will be made when identifying the discourse topic. This would lead to a worse performance on null pronouns, whereas overt pronouns would be less affected.

Related to processing speed, the model predicts that a lower processing speed will influence the interpretation of overt pronouns more than the interpretation of null pronouns. This is because null pronouns can be resolved already after the uni-directional processing step, whereas the interpretation of overt pronouns requires completion of the next step of bi-directional optimization. If this bi-directional
experiments. generates pronoun interpretation in Italian processing, the model can plausibly simulate subject cognitive architecture ACT mechanism to select the most likely interpretation of an constraint processing of We have implemented a cognitive model that simulates the to be taken into consideration language such as Italian or Spanish, complete model of pronoun processing in a null subject pronouns may well continue after the average time it takes to say a word, but the resolution of pronoun. On the other hand, the time limit was based on the have less time to process a null subject than an overt subject deduced from the dis overtly subject pronoun. However, since null subjects are not therefore simulation use pronoun expressions into account (null prominence (accessibility). Moreover, the model only takes two anaphoric expressions into account (null subjects and overt subject pronouns). In natural speech, a speaker could also choose to use, for example, an NP instead of a pronoun. In the simulation however, this option was not provided and therefore the model is a simplification of natural language use. Finally, it was assumed that participants get just as much time to process a null subject as they get to process an overt subject pronoun. However, since null subjects are not overtly expressed in language and their occurrence has to be deduced from the discourse, it may be the case that listeners have less time to process a null subject than an overt subject pronoun. On the other hand, the time limit was based on the average time it takes to say a word, but the resolution of pronouns may well continue after this moment. For a full, complete model of pronoun processing in a null subject language such as Italian or Spanish, these aspects will have to be taken into consideration.

Conclusion

We have implemented a cognitive model that simulates the processing of subject pronouns in Italian. The model uses a constraint-based approach and a perspective-taking mechanism to select the most likely interpretation of an overt or null pronoun. Combining constraints from the cognitive architecture ACT-R and constraints on language processing, the model can plausibly simulate subject pronoun interpretation in Italian, and, most importantly, generates concrete predictions that can be tested in future experiments.

References

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