Bulletin of the *Transilvania* University of Braşov
Series IV: Philology and Cultural Studies • Vol. 8 (57) No. 1 – 2015

Effects of animacy and gender on the choice of referring expression

Sofiana I. CHIRIACESCU[1]

In communication, speakers constantly choose between different types of referring expressions to refer to discourse entities. Previous research (Chafe, 1976; Ariel, 1988; Gundel, Hedberg and Zacharski, 1993, among many others) has suggested that this choice depends on the degree of accessibility associated with each referent. In a story completion experiment, the effects of animacy and gender on the production of definite noun phrases and pronouns are investigated. Results show that, overall, pronominalization rates are negatively affected by the presence of two human referents, irrespective of their gender. In other words, even when a pronoun would non-ambiguously point to the intended referent, the presence of another human referent increased the use of definite noun phrases. These effects suggest that the accessibility of referents is affected by competition in terms of animacy. On a more general level, this study deepens our understanding of the different linguistic factors that are at work in the process of referring.

Keywords: *animacy, gender, referring expressions, accessibility, language production.*

1. Introduction

In discourse processing, interlocutors collaborate towards coherence, negotiating for a mutually shared mental discourse model. Coherence is achieved by means of lexical knowledge and different grammatical processing cues, as well as by the need to establish prominence relations between discourse referents (Givón, 1983, 1992). Most models put forth so far have looked at how *referent tracking* and *reference production* are used to achieve coherence. The general consensus is that in a given discourse segment referents are not equally accessible and that speakers (more or less) explicitly signal a referent's degree of accessibility (Chafe, 1976; Ariel, 1988; Gundel, Hedberg, and Zacharski, 1993). One way to signal referential accessibility is by using a particular type of referring expression. For example, it was shown that more elaborated types of referring expression (e.g. *the famous modern sculptor*, *Constantin Brâncuşi*) are generally associated with less accessible referents, and that more attenuated expressions (e.g. *he*) tend to refer back to highly accessible entities

[1] University of Cologne, *Transilvania* University of Braşov; sofiana.chiriacescu@uni-koeln.de

BDD-A18362 © 2015 Transilvania University Press
Provided by Diacronia.ro for IP 216.73.216.221 (2025-10-20 12:16:02 UTC)

(Hobbs, 1979; Ariel, 1988; Grosz, Joshi and Weinstein, 1995; Arnold, 1998; Kaiser and Trueswell, 2008; Kehler, Kertz, Rohde and Elman, 2008).

Most studies focused on highly accessible referents and investigated the factors that contribute to pronoun use and interpretation. Different syntactic, semantic and information structural factors, such as syntactic function, thematic roles, parallelism and implicit causality have been shown to affect pronoun processing and thus the accessibility of referents (Arnold, 1998; Chiriacescu, 2011a).

The current paper extends previous accounts in several ways. First, we will investigate the role played by and interaction between animacy (i.e. human vs. non-human) and gender (i.e. masculine vs. feminine) on the choice of referring expression and for establishing accessibility. Second, compared to most previous research, we focus on initially less-prominent referents, i.e. on direct object referents and on non-topical referents in English. We will conduct a sentence continuation study to explore these issues. On a more general level, the aspects addressed in this study have potential implications for the better clarification of the notion of accessibility, which is central to studies investigating reference resolution and reference tracking in discourse.

The structure of this paper is as follows. In Section 2, a brief review of the factors that contribute to referential accessibility is offered. Section 3 presents and discusses a sentence-completion study that explores the impact of animacy and gender on the use or pronouns and definite noun phrases. The findings of the study and several directions for further research are discussed in the last section.

2. Accessibility, gender and animacy

Assuming that the process of referring is dynamic, constantly changing, and that referents are more or less accessible as the discourse unfolds, different theories of accessibility or prominence emerged, as an attempt to capture the correlation between a particular type of referring expression and the discourse features the referent associated with it encodes. In verbal communication, participants keep mental representations of the entities introduced in the discourse (Bransford et al., 1972, Clark and Haviland, 1977). These mental representations help discourse participants organize the information at hand by keeping track of the information that was already mentioned and linking the new information to already established discourse representations. In such frameworks, where *accessibility* is regarded as a gradient category rather than a categorical one, a discourse referent can be more or less accessible. The common assumption is that each anaphoric expression is linked to one referent that is selected from a list of discourse referents that are ranked with respect to their accessibility.

BDD-A18362 © 2015 Transilvania University Press
Provided by Diacronia.ro for IP 216.73.216.221 (2025-10-20 12:16:02 UTC)

While full noun phrases significantly rely on their descriptive content for unique referential identification, pronouns like *he* and *it* have little or no descriptive content. Their resolution to a unique referent heavily relies on the accessibility ranking of individuals in the preceding context. Several studies have shown that there is an inverse relation between the explicitness of the anaphoric expression (in terms of descriptive, lexical, and phonological material) and the accessibility of its antecedent expression. In other words, a less explicit type of referring expression often correlates with an accessible referent; the opposite holds for more explicit types of referring expressions (e.g. definite noun phrases). As pronouns were shown to be associated with highly accessible referents, most research has investigated the factors that contribute to the proper use and interpretation of pronouns in English. Recency, givenness, syntactic prominence, semantic prominence, implicit causality are some of the factors that have been shown to boost a referent's accessibility. For example, in (1 a-c), the referent associated to *Paul* is more accessible than the referents associated to *a man, a woman,* and *a vase,* as this referent is mentioned first, in subject position and as the aboutness topic. Thus, a subsequent pronoun should be interpreted more readily as referring back to this referent, than to its less accessible or prominent counterpart.

(1) a. Paul$_1$ saw a man$_2$ when *he*$_{1/?2}$_____
b. Paul$_1$ saw a woman$_2$ when *she*$_2$_____
c. Paul$_1$ saw a vase$_2$ when it$_2$_____

However, besides accessibility, animacy and gender affect pronoun processing as well. Accordingly, the pronouns in (1b) and (1c) should be more easily interpreted compared to the pronoun in (1a), as they unambiguously point to one antecedent expression that encodes different gender or animacy information.

A question that arises at this point is how different types of information (i.e. accessibility, animacy and semantic gender) compete in pronoun resolution. In other words, what are the roles of accessibility, animacy and gender in choice of referring expression? Most studies that have focused on pronoun interpretation offered divergent responses to this question. For example, proponents of the Minimalist Hypothesis (McKoon and Ratcliff, 1992) suggested that accessibility information is used first, while gender information comes at play at a later stage in pronoun processing. Another body of research (Crawley et al., 1990; Ehrlich, 1980), on the contrary, has argued that gender is used rapidly and efficiently in pronoun disambiguation and that accessibility is used at a later stage, but only when it is needed.

Summarizing, previous research did not provide consistent results on the rapid use of these semantic factors in the disambiguation of pronouns. Moreover, most studies focused on the interpretation of pronouns in on-line processing by employing probe word recognition tasks or eye-tracking reading experiments. This raises

BDD-A18362 © 2015 Transilvania University Press
Provided by Diacronia.ro for IP 216.73.216.221 (2025-10-20 12:16:02 UTC)

several questions. For example, it is unclear, whether animacy and gender show the same effects in off-line production. Additionally, we do not know whether the various factors associated with accessibility (e.g. syntactic position) and other semantic cues (e.g. animacy and gender) affect the next-mention biases and the choice of subsequent referring expressions in a similar way. In the following sections, we report on an experiment, which showed that animacy and gender should be accounted for while investigating referential preferences in production.

3. The story completion study

To explore the effects of animacy and gender on the choice of referring expression, a sentence-completion task was conducted.

Participants. One hundred and one self-reported native English speakers participated in the study. They were recruited using Amazon's Mechanical Turk service. It took about twenty minutes for each participant to complete the experiment.

Materials and Design. Twenty-four critical items were constructed, which consisted of scene-narrative pairs, as illustrated in Table 1 below. The first sentence of each experimental item set up a particular context (e.g., walking down the street in example (a)) and introduced a character (e.g., *James* in (a)), which was always human and masculine. This character was introduced by a proper name in subject position and was the clearly established aboutness topic[2]. The second sentence rementioned the first character twice by means of pronouns, in syntactic subject position and as the aboutness topic. The second sentence additionally introduced the critical referent in direct object position. The realization of the second referent was manipulated with respect to animacy (i.e. human vs. non-human) and gender (i.e. masculine vs. feminine), which resulted in three conditions: (i) [hum/masc]; (ii) [hum/fem]; (iii) [non-hum], as illustrated in Table 1.

(a) James$_1$ decided to hang out in the park.		
(b) On his$_1$ way there, he$_1$ saw___	(i) a man$_2$.	[hum/masc]
	(ii) a woman$_2$.	[hum/fem]
	(iii) a bench$_2$.	[non-hum]

Table 1. Sample experimental items

[2] We will use the term 'topic' as an information structural device, which stands for the entity the current sentence makes a predication (Reinhart 1982). I assume that most sentences (with the exception of thetic sentences, such as *the dog is barking)* can have at most one aboutness topic.

BDD-A18362 © 2015 Transilvania University Press
Provided by Diacronia.ro for IP 216.73.216.221 (2025-10-20 12:16:02 UTC)

We asked participants to read each two-sentence story and continue it by providing one natural-sounding and plausible sentence continuation. Participants were told that there was no correct or incorrect way to continue the stories and that there was no connection between the different stories.

As participants' continuations were not restricted in any way (e.g. by pronouns, sentence connectives, etc.), their responses were made on the basis of the mental representations they developed during reading the stories.

Procedure and data analysis. The continuation sentences (including subordinate ones, if there were any) provided by the participants, were coded by two independent judges with respect to the type of anaphoric referring expressions used to pick up the referent associated with the critical item. We additionally coded the type of referring expression used to remention the first referent, which was introduced in topic and subject position. Please note that we excluded from the present analysis non-referentially used noun phrases, bare noun phrases, generic noun phrases, bridged or linked noun phrases, plurals, appositives and predicate nominatives[3]. The two judges agreed upon 92% of the cases; in the case of disagreement, differences were resolved through discussion.

Predictions. Previous work has shown that pronominalization biases are generally modulated by accessibility and that human referents and sentences containing referents of different gender correspond to higher rates of immediate subsequent mention and pronominalization compared to their non-human and same-gender counterparts (Givón, 1983; Ariel, 1899; Arnold 1998). In light of these observations we make several predictions:

- As the first referent is highly accessible by virtue of being realized in subject position, as a topic, and as the semantic Agent of the target sentence, the prediction is that this referent will be mentioned next by means of less elaborated expressions (e.g. pronouns). This tendency is not expected to vary with the manipulation of the second referent.
- As the critical referent is associated with a lower degree of accessibility compared to the first, the prediction is that we find overall lower rates of pronominalization for the first. Gender is not expected to impact this preference.
- As previous research has shown that referents associated with human concepts are more accessible than those associated with non-human referents, we expect participants to produce more elaborated expression for the latter.

[3] It is still under debate whether the morphologically unrealized subject position in the second conjunct of a sentence like *Paul came in and Ø sat down* should be considered an anaphoric element, or not. In the present study, we coded the deleted subject referent in similar examples as an anaphoric instance of that referent, without adopting a particular syntactic analysis for that position (but see Van Valin (1986) or Brandner and Fanselow (1992) for a discussion of several approaches to subject lacking in finite structures in English).

BDD-A18362 © 2015 Transilvania University Press
Provided by Diacronia.ro for IP 216.73.216.221 (2025-10-20 12:16:02 UTC)

- With respect to gender, we expect more pronouns for the different-gender condition (i.e. [hum/fem)] than for the same-gender condition (i.e. [hum/masc]). This prediction is based on the reported tendency to avoid ambiguity (Ariel, 1988; Givón, 1983).

Results and discussion. Ten continuations were excluded from the analysis, because they were not coherent (*n=8*) or because participants did not write full sentences (*n=2*). This left us with 91 responses that could be coded. The two independent judges coded only those continuation sentences that mentioned at least one of the referents introduced in the experimental items.

The findings show that both animacy and gender affect language production in terms of likelihood of pronominalization. First, the results support previous findings on the likelihood of highly accessible referents to be pronominalized, as the first referent introduced in subject position was picked up as a pronoun in the continuations in 64% of the cases across conditions. That is, calculating the type of referring expression used more frequently to pick up the first referent, we observe that participants predominantly opted for pronouns and proper names (e.g. 64% and 29% respectively, as reported in Table 2) to fulfil this function.

Referents	Pronoun	Name	Def. unmod. NP	Def. mod.NP
Subject (1st referent)	64%	29%	6%	n.a.
[hum/masc]	14%	2%	33%	51%
[hum/fem]	22%	n.a.	47%	31%
[non-hum]	30%	n.a.	45%	25%

Table 2. Type of referring expressions used to pick up the referents (mean values)

Second, as expected, we find overall lower rates of subsequent pronominalization for the referents in the critical conditions, which were realized as direct objects in the target sentences. The manipulation of the critical referent with respect to animacy and gender did not reverse this pattern. Participants chose definite (un)modified referring expression to pick up the critical referents in the subsequent discourse. The results confirm previous studies, which have shown that less accessible referents (in terms of syntactic position, thematic role, topichood) do not tend to be pronominalized (Arnold, 1988; Chiriacescu, 2011b).

Third, contrary to the initial predictions, findings reveal lower rates of pronominalization for referents introduced in the human condition compared to the non-human condition, as illustrated in the second column of Table 2. In other words, participants used less pronouns for referents in the [hum/masc] and [hum/fem] condition compared to the [non-hum] one.

BDD-A18362 © 2015 Transilvania University Press
Provided by Diacronia.ro for IP 216.73.216.221 (2025-10-20 12:16:02 UTC)

Fourth, the manipulation of gender did impact the type of referring expression used in the subsequent discourse to pick up the critical referent, as predicted. As reported in Table 2, participants used *slightly* more pronouns for referents introduced in the different-gender condition compared to referents in the same-gender condition. Moreover, participants did not only use fewer pronouns for referents in the [hum/masc] condition, compared to referents in the [hum/fem] condition, but they more often used definite modified noun phrases to pick up the referent (i.e. in 51% of the cases for [hum/masc] referents, vs. 31% for referents in the [hum/fem] condition, as illustrated in the last column of Table 2).

A closer inspection of the continuations reveals another interesting pattern with respect to the type of sentence (i.e. transitive vs. intransitive) participants opted for to pick up the referents in the three conditions, as illustrated in Table 3.

Referents	Transitive sentences	Intransitive sentences
[hum/masc]	43%	57%
[hum/fem]	47%	53%
[non-hum]	67%	33%

Table 3. Transitive vs. intransitive sentences used by participants in all conditions

The results show that participants chose to continue talking about *both* referents (i.e. the subject and the critical referent) more often than not, when the critical referent was associated to a non-human individual. A typical continuation sentence for the [non-hum] condition is the following: *Indeed, he saw a bench not too far from where he was, and [...]*. We furthermore notice that referents in this condition were continued in a parallel syntactic position, i.e. in direct object position. The opposite seems to hold for the [human] condition, as less than 50% of the participants picked up the subject and the critical referent in the same continuation sentence. More intransitive sentences were chosen to remention the critical referents and participants' continuations focused more on the initial subject referent. Furthermore, compared to the [non-human] condition, participants mentioned the critical referents more often in a syntactic non-parallel position. More topic shifts occurred in participants' continuation for these conditions (e.g. *The woman was singing "Au Clair de la Lune"*). The difference in gender for human referents did not reverse the tendencies mentioned above.

As a whole, participants' responses revealed that both animacy and gender affect the subsequent choice of referring expression. While the latter factor showed only minimal effects, animacy showed clear effects in terms of likelihood of pronominalization and subsequent sentence type.

BDD-A18362 © 2015 Transilvania University Press
Provided by Diacronia.ro for IP 216.73.216.221 (2025-10-20 12:16:02 UTC)

4. General discussion

The study presented in this paper used a sentence completion task to explore how people employ linguistic information in the process of establishing referential chains. We used animacy and gender as a testing ground into this issue. Although these two factors play a pivotal role in reference processing, previous studies have come to conflicting results about the way in which speakers integrate this information in the process of reference production. The experiment presented in this paper was intended to shed some light on these aspects, even though many questions still remain open for further research.

We find overall lower rates of subsequent pronominalization for the referents in the critical conditions, which were realized as direct objects in the target sentence. As subject referents were made more prominent than the direct objects, being realized in topic position, as the semantic agent of the sentence that introduced them, it was no surprise to find overall stronger subject pronominalization scores. This finding is in line with previous investigations, which showed that, under controlled conditions, subjects and aboutness topics have a strong impact upon subsequent pronoun production (e.g. Arnold, 1998; Chiriacescu 2014, among many others).

Even though the participants in the experiment reported here could have opted for more reduced types of referring expressions for the conditions [hum/fem] and [non-hum], as there was no competition between the referents in terms of gender or animacy, they nevertheless chose more elaborated types of referring expressions (i.e. definite (modified) noun phrases). However, the findings show that it was not only the competition in terms of gender that hindered participants to use more pronouns for the critical referents, but that the presence of multiple referents in a sentence describing a transitive event reduces the rate of pronominalization. The results furthermore show that the low pronoun rates for the critical referents are due to the competition of the two referents in terms of animacy. This observation converges with a body of neurolinguistic research, which showed that transitive sentences are easier to process when the arguments are different in terms of their distinctive features such as animacy, for example (McElree, 2006).

The results further show that the matching in animacy of two referents introduced in a transitive event impacts not only the subsequent choice of referring expression, but also the subsequent choice of event type. Transitive sentences were chosen as continuations for sentences that introduced referents, which differed in animacy, and intransitive sentence were preferred for referents matching in animacy.

The general findings show that the choice of referring expression is sensitive to multiple constraints. On the one hand, accessibility was shown to influence choice of subsequent expression. Referents associated with high levels of accessibility or prominence were more likely to be picked up by pronouns, whereas less accessible referents were predominantly mentioned again by definite (modified) noun phrases. On the other hand, manipulating the animacy and gender of referents realized in

BDD-A18362 © 2015 Transilvania University Press
Provided by Diacronia.ro for IP 216.73.216.221 (2025-10-20 12:16:02 UTC)

direct object position, had an impact upon the discourse structure of the subsequent discourse in terms of event type. Recent investigations (Chiriacescu, 2014) showed that other factors, such as the type of referring expression affect the discourse structure of the subsequent discourse. One path for future research would be to examine other factors, such as information status, that might impact the structure of the subsequent discourse, independently of accessibility.

Acknowledgements

I would like to thank to two anonymous reviewers and to the editors of this volume for their comments and suggestions.

References

Ariel, Mira. 1988. "Referring and Accessibility." *Journal of Linguistics* 24: 65–87.

Arnold, Jennifer. 1998. *Reference form and discourse patterns*. Stanford, CA: Stanford University dissertation.

Brandner, Ellen, and Gilbert Fanselow. 1992. *Coordinate structures and pro*. Ms. University of Passau.

Bransford, D. John, Richard J. Barclay, and Jeffery J. Franks. 1972. "Sentence memory: A constructive versus interpretive approach." *Cognitive Psychology* 3: 193–209.

Chafe, Wallace. 1976. "Givenness, contrastiveness, definiteness, subject, topics, and point of view." In *Subject and Topic*, ed. Li, C., 25-55. New York: Academic Press.

Chiriacescu, I. Sofiana, 2011a. "Factors contributing to the salience of referents." *Studia linguistica et philologica*. Bucharest: Bucharest University Press.

Chiriacescu, I. Sofiana. 2011b. "Effects of Reference Form on Frequency of Mention and Rate of Pronominalization." In *Anaphora Processing and Applications. Lecture Notes in Computer Science,* ed. S.L. Devi, A. Branco and R. Mitkov, 132-142. Heidelberg: Springer.

Chiriacescu, I. Sofiana. 2014. *The discourse structuring potential of indefinite noun phrases. Special markers in English, German and Romanian*. Stuttgart: University of Stuttgart.

Chiriacescu, I. Sofiana, and Klaus von Heusinger. 2010. "Discourse prominence and *pe*-marking in Romanian." *The International Review of Pragmatics* 2(2): 298–332.

Clark, H. Herbert, and Elma S. Haviland. 1977. "Comprehension and the given-new contract". In *Discourse Production and Comprehension, advances in research and theory*, ed. Freedle, R. O., 1–40. Hillsdale, NJ: Lawrence Erlbaum Associates.

BDD-A18362 © 2015 Transilvania University Press
Provided by Diacronia.ro for IP 216.73.216.221 (2025-10-20 12:16:02 UTC)

Crawley, A. Rosalind, Rosemary J. Stevenson, and David Kleinman. 1990. "The use of heuristic strategies in the interpretation of pronouns." *Journal of Psycholinguistic Research* 194: 245–265.

Ehrlich, Kate. 1980. "Comprehension of pronouns." *Journal of Experimental Psychology* 32: 247–255.

Gernsbacher, M. Allan, and Sandra Shroyer. 1989. "The cataphoric use of the indefinite 'this' in spoken narratives." *Memory and Cognition* 17: 536–540.

Givón, Talmy. 1983. "Topic continuity in discourse: An introduction." In *Topic continuity in discourse: A quantitative cross-language study*, ed. Talmy Givón, 1-42. Amsterdam: John Benjamins.

Givón, Talmy. 1992. "The grammar of referential coherence as mental processing instructions." *Linguistics* 30: 5–55.

Grosz, Barbara, Josh Aravind, and Steven Weinstein. 1995. "Centering: framework for modelling the local coherence of discourse." *Computational Linguistics* 21: 203-226.

Gundel, Janette, Nancy Hedberg, and Ron Zacharski. 1993. "Cognitive status and the form of referring expressions in discourse." *Language* 69: 274-307.

Hobbs, Jerry. 1979. "Coherence and coreference." *Cognitive Science* 2(1): 67–90.

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

Kehler, Andrew, Laura Kertz, Hanna Rohde, and Jeff Elman. 2008. "Coherence and coreference revisited." *Journal of Semantics* 25: 1–44.

McElree, Brian. 2006. "Accessing recent events." In *The psychology of learning and motivation,* ed. Ross, B, 155–200. San Diego, CA: Academic Press.

McKoon, Gail, and Roger Ratcliff. 1992. "Inference during reading." *Psychological Review* 993: 440–466.

Reinhart, Tanya. 1982. "Pragmatics and linguistics: An analysis of sentence topics." *Philosophica* 27: 53–94.

Van Valin, Ron Jr. 1986. "Pragmatics, island phenomena and linguistic competence." *Papers from the Parassession on Pragmatics and Grammatical Theory*, 223–233.

BDD-A18362 © 2015 Transilvania University Press
Provided by Diacronia.ro for IP 216.73.216.221 (2025-10-20 12:16:02 UTC)