ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 14, No. 1, pp. 125-130, January 2024
DOI: https://doi.org/10.17507/tpls.1401.14

The Effect of Subordination on Antecedent Prominence in Anaphora Processing*

Jinmei Chen
Faculty of Foreign Studies, Beijing Language and Culture University, Beijing, China

Wei Cheng[1]
College of Foreign Studies, Jinan University, Guangzhou, China

***Abstract*—One important factor affecting anaphor resolution is the prominence of its antecedent in the discourse. Discourse Prominence Theory (Gordon & Hendrick, 1998) defines antecedent prominence in terms of its height in the phrase structure tree. Two reading time experiments were conducted to test this theory by investigating the effect of subordination on antecedent prominence both within and between sentences. Both experiments had a 2×2 design with Structure (intra- vs. inter-sentential) and Anaphor (pronoun vs. repeated name) as independent variables, and differed in the degree of subordination of the antecedent. The results of the experiments show a main effect of Anaphor such that clauses/sentences containing pronouns were read faster than those containing repeated names, but no interaction between Structure and Anaphor. These results indicate that the repeated name penalty occurs no matter how deeply the antecedent is embedded within subordinate clauses. Taken together, the findings demonstrate that antecedent prominence is not determined by its height in the phrase structure as claimed by Discourse Prominence Theory but rather may be related to the syntactic function (subject) and status (head noun) of antecedents.**

***Index Terms*—subordination, anaphora, prominence, repeated name penalty, Discourse Prominence Theory**

I. Introduction

Cohesion is an important device which links text together and makes it meaningful (Halliday & Hasan, 1976). Various means can be employed to realize cohesion, one of which is anaphora. Speakers or writers often refer backwards to someone or something that is previously mentioned. Such anaphoric reference can be realized linguistically either as full expressions such as definite noun phrases and repeated proper names, or as reduced expressions such as reflexives, pronouns and even null elements. These different types of anaphoric expressions are not used arbitrarily. Rather, they have different functions and roles and follow certain patterns of use under the influence of a variety of factors including syntactic constraints, their referents or antecedents' role in the discourse, etc. (Ariel, 1990). This paper focuses on the factor of antecedent prominence and investigates its effect on anaphora processing.

A. Anaphora Within a Sentence

When occurring within the boundaries of a sentence, any referential expression is subject to syntactic constraints known as Binding Theory in generative syntax (Chomsky, 1981). It consists of three principles which determine the usage of reflexives, pronouns and R-expressions, respectively. Principle A determines that reflexives must have an antecedent within its local governing category. Thus, in (1), *himself* can only refer to the more local referent *Mark* rather than the farther away referent *John*.

(1) John told Mark_i to wash himself_i.

By contrast, Principle B determines that pronouns cannot have a local antecedent. Therefore, in (2), the pronoun *him* cannot refer to *Mark* but rather either the subject *John* or some other entity in the discourse.

(2) *John told Mark_i to wash him_i.

Principe C regulates that R-expressions (e.g., proper names) cannot be coreferential with any entity in the sentence that c-commands it. Hence, in (3) the proper name *John* cannot refer to the same person as the subject pronoun *he* because the pronoun *he* is in a c-commanding relation with *John*.

(3) *He_i told Mark to wash John_i.

B. Anaphora in Discourse

Despite the well-documented operation of Binding Principles in anaphora resolution in the psycholinguistic literature, this theory only applies within sentences and says nothing about reference resolution between sentences. In normal speech, people do not just produce one utterance, but rather a series of utterances. Thus, anaphora may go beyond the

* The research was partially supported by the Key Project Funding of the 2022 International Chinese Education Research Program (Approval Number: 22YH31B).
[1] Corresponding Author. E-mail address: jnucheng@126.com

boundaries of a single sentence and occur at the discourse level. Moreover, even within the boundaries of a sentence, anaphora resolution is subject to other factors than Binding Principles. For example, in (4), both *Mark* and *John* can be coreferential with the pronoun *he* because neither of them is c-commanded by the pronoun. To correctly figure out what *he* refers to, we need to consider discourse information. From this example, therefore, we can see that even anaphora within a sentence is subject to discourse factors.

(4) Before he left, Mark washed John.

One important finding in anaphora resolution in discourse is that, when coreferential with previous entities, repeated names are generally processed more slowly than pronouns, a phenomenon known as the repeated name penalty (Gordon et al., 1993). Numerous studies have confirmed this phenomenon in different discourse conditions (e.g., Gordon & Chan, 1995; Gordon et al., 1999), in other languages such as Chinese (Young et al., 1999, 2003), Spanish (Gelormini-Lezama, 2018; Gelormini-Lezama & Almor, 2011), and Japanese (Shoji et al., 2017), using non-behavioral research paradigms such as ERP (Swaab et al., 2004) and eye tracking (Ledoux et al., 2007).

Initially, Gordon et al. (1993) provided an account for the repeated name penalty from the perspective of Centering Theory (Grosz et al., 1983, 1995). The basic assumptions of Centering Theory are that every non-initial utterance has a backward-looking center and a set of forward-looking centers, with the backward-looking center establishing a link between the current utterance and previous discourse and the forward-looking centers making a potential connection between the present utterance and subsequent discourse. Under this account, pronouns incur less processing effort than repeated names because they are the most natural vehicle to take up the role of backward-looking center.

Gordon and Hendrick (1998) further developed Centering Theory and proposed Discourse Prominence Theory (DPT) as a general framework for anaphora processing both within and between sentences. This theory maintains that anaphors are resolved by construction rules mapping syntactic representation to discourse representation. Pronouns and repeated names have different effects on anaphora processing because the construction rules for them are not the same. When the referring expression is realized as a pronoun, the construction rule directly relates it to an entity in previous discourse which matches in number, gender and meaning. However, when a repeated name is used to make reference, the relevant construction rule will first establish a new entity, which is further matched up with another entity predicted on the same name in previous utterances through a rule of equivalence. Thus, repeated names involve more cognitive processes than pronouns, resulting in slower processing time.

C. Antecedent Prominence in Anaphora Processing

The advantage of pronouns over repeated names in anaphora processing does not always happen. A major factor moderating the repeated name penalty is the degree of prominence of the antecedent. For example, in Gordon et al. (1993), pronouns are read faster than repeated names only when their antecedent is in a prominent position in local discourse. This prominence effect is, according to DPT (Gordon & Hendrick, 1998), due to the difference between construction rules for pronouns and repeated names as sketched above. On the one hand, it is easier to establish a coreferential relation between a pronoun and a prominent antecedent because upon occurrence of pronominal reference the construction rule for pronouns will directly search for appropriate discourse entity. The more prominent the entity is, the more easily it would match up with the pronoun. On the other hand, repeated names trigger the projection of a new entity, which is further equated with previous entity. This equation process works more effectively with non-prominent than prominent entities.

According to DPT, antecedent prominence is defined as follows: "The syntactic prominence of an NP is related to its height in a syntactic tree and therefore inversely related to its depth of embeddedness" (Gordon & Hendrick, 1998, p. 406). This definition of referent prominence finds support in some studies. In a reading time study, Gordon et al. (1993) found that if the antecedent is the subject of previous utterance, there will be a reading time advantage of reference with pronouns over repeated names. Besides the antecedent's grammatical role, they also found an effect of surface order. When the antecedent is the first-mentioned entity, the repeated name penalty will occur even if it is not the subject. In a study on both intra-sentential and inter-sentential anaphora, Gordon et al. (1999) discovered that, when referring to an independent noun phrase in the subject position, clauses containing pronouns are read faster than those containing repeated names; by contrast, no significant difference in reading time was found when the antecedent is a component of a conjoined noun phrase or a possessive. Swaab et al. (2004) did a similar study with ERP and found the same effect.

Informative as these studies are, there are some limitations. For example, the test items in Gordon et al. (1999) may violate Principle C of Binding Theory (Chomsky, 1981). As mentioned above, this syntactic principle determines that any R-expression cannot refer to an antecedent that c-command it. In (5), an example in their study (Gordon et al., 1999, p. 358), the subject of the main clause is in a c-commanding relation with the subject of the subordinate clause. Therefore, the repeated name *John* cannot refer backwards to the first *John*. Although Gordon and Hendrick (1997) found that native English speakers' intuition about Principle C is not as accurate as Principles A and B, they did find that coreference is more acceptable when the antecedent name and the referential name are not in a c-commanding relation than when they are. Thus, the longer reading time of the clause containing the repeated name may not necessarily reflect the repeated name penalty as observed in other contexts but rather the interference of syntactic constraints.

(5) John went to the store so that he/John could buy some candy.

Another possible confounding factor in previous studies is that the relationship between embeddedness and prominence is entangled with constituent headedness in the test stimuli. For instance, in (6) (Gordon et al., 1999, p. 358), *John* is embedded in a conjunct and is not a head noun in a strict sense. Thus, it is not clear whether the results support their claims of antecedent prominence in DPT, or whether they simply reflect the unprivileged status of non-headedness of the referent.

(6) John and Mary went to the store so that he/John could buy some candy.

D. The Present Study

In this study, two self-paced reading experiments were conducted to investigate the effect of antecedent prominence on anaphora processing. In particular, we aimed to test the predictions of DPT (Gordon & Hendrick, 1998) regarding antecedent prominence from the perspective of the repeated name penalty. Different from previous studies, we disentangled embeddedness and constituent headedness by considering constructions where the antecedent is embedded within subordinate clauses yet still occupies the subject head noun position. Examples (7) and (8) demonstrate the stimuli used in the two experiments. In (7), the antecedent is embedded within a temporal subordinate clause, while in (8) it is first embedded within a complement clause, which is further embedded in a subordinate clause introduced by *although*. Since the subordinate clause precedes the main clause, the possible antecedent does not c-command the referential expression, thus avoiding violation of Principle C. In both experiments, the antecedent in the test items appears in a non-prominent position in accordance with the definition of antecedent prominence in DPT, thus providing a new testing ground for the prediction of the theory. Moreover, since DPT is claimed to apply both within and between sentences, this study also examined anaphora processing both inter-sententially and intra-sententially.

(7) When John went to the store, he/John saw a squirrel.

(8) Although it was said that Alex would come, he/Alex did not show up.

II. METHOD

A. Participants

Fifty native English speakers took part in the study, 24 in Experiment One and 26 in Experiment Two. All were undergraduate students from a public university in the U.S. and received course extra credit for participation.

B. Materials

Thirty-two items were created for Experiment One. Each item consisted of four conditions as shown in Table 1. In the intra-sentential condition, the main character was introduced as a subject head noun of the subordinate clause while the referring expression appeared in the subject position of the following main clause. It was either a pronoun (matched with the antecedent in number and gender) or a repeated name coreferential with the subject of the first clause. The inter-sentential version was constructed by breaking down the intra-sentential version into two independent sentences. The inter-sentential version was similar to the intra-sentential version except for the absence of the subordination word such as *when* in the sample and the change of punctuation and capitalization. Certain minor changes were also made to the first sentence to ensure discourse coherence between the two sentences. The second sentence remained the same as the second clause in the corresponding intra-sentential version.

TABLE 1
SAMPLE TEST ITEMS OF EXPERIMENT ONE

	Intra-sentential	Inter-sentential
Pronoun	When John went to the store, he saw a squirrel.	John went to the store. He saw a squirrel.
Repeated name	When John went to the store, John saw a squirrel.	John went to the store. John saw a squirrel.

Every item occurred in just one condition across the participants so that every participant read only one version of an item. There were a total of 32 test stimuli and 27 fillers for every participant. Filler items were similar to test items but differed in syntactic structure. Moreover, to ensure that participants read the item, every stimulus was followed by a yes/no question.

Experiment Two also contained 32 test items, some of which were adapted from Experiment One. The crucial difference between the stimuli of the two experiments lied in the degree of embeddedness. While the antecedent was embedded within one subordinate clause in the intra-sentential condition in Experiment One, it was embedded within two subordinate clauses in Experiment Two. Other designs were the same between the two experiments. Sample items are presented in Table 2.

TABLE 2
SAMPLE TEST ITEMS OF EXPERIMENT TWO

	Intra-sentential	Inter-sentential
Pronoun	Although it was said that Alex would come, he did not show up eventually.	It was said that Alex would come. He did not show up eventually.
Repeated name	Although it was said that Alex would come, Alex did not show up eventually.	It was said that Alex would come. Alex did not show up eventually.

C. Procedure

The two experiments had the same procedure. The participants were tested on a Window-based PC running E-Prime 2.0. They read the test item clause by clause or sentence by sentence, depending on the types of stimuli. The participants controlled their own reading speed by pressing the space bar to read the next clause/sentence. When they pressed the space bar, the previous clause/sentence disappeared from the screen. After they read each item, a yes/no question appeared on the screen, which was answered by pressing the 'B' or 'N' key. At the end of each trial, they received feedback as to whether their answer was right or wrong. A practice block consisting of five fillers different from those in the experiment was implemented before the experiment started.

III. RESULTS AND DISCUSSION

The dependent variable was the reading time of the second clause/sentence containing the referring expressions, that is, the time latency from the presentation of the second clause/sentence to the participants' pressing the space bar. If a participant's accuracy rate was below 80%, the data of that participant would be dropped, and the trials in which the reading time was three *SD*s away from the mean of each participant would be considered as outliers and dropped. In the end, the data of two participants in Experiment One and one participant in Experiment Two were eliminated. Less than 1% of the data were considered as outliers in each experiment.

Table 3 shows the mean reading times of the second clause/sentence in different conditions in Experiment One. A 2×2 repeated measures ANOVA was conducted with Structure (sentence vs. discourse) and Anaphor (pronoun vs. repeated name) as within-participants factors using SPSS. There was no main effect of Structure [$F1$, $F2 < 1$] but a main effect of Anaphor such that reading times of the second clause/sentence were longer in the repeated name condition than the pronoun condition [$F1(1, 21) = 7.38$, $p < .05$; $F2(1, 31) = 6.30$, $p < .05$]. Moreover, there was no significant interaction between Structure and Anaphor [$F1 < 1$; $F2(1, 31) = 3.00$, $p = .093$].

TABLE 3
MEAN READING TIMES OF THE SECOND CLAUSE/SENTENCE IN MILLISECONDS IN EXPERIMENT ONE

Structure	Referring expression	
	Repeated name	Pronoun
Intra-sentential	1626 (429)	1460 (393)
Inter-sentential	1565 (341)	1468 (428)

Note. Reading time standard error of the mean in parentheses.

The mean reading times of the second clause/sentence in different conditions in Experiment Two are presented in Table 4. The statistic analysis was the same as in Experiment One. The results show that there was a main effect of Structure [$F1$ (1, 24) = 4.85, $p < .05$; $F2$ (1, 31) = 6.51, $p < .05$], with the second line being read faster in the intra-sentential condition. There was also a main effect of Anaphor by item [$F2(1, 31) = 5.23$, $p < .05$], such that clauses/sentences containing pronouns were read faster than those containing repeated names; the effect was marginally significant by subject [$F1(1, 24) = 4.06$, $p = .055$]. However, there was no significant interaction between the two factors [$F1$, $F2 <1$].

TABLE 4
MEAN READING TIMES OF THE SECOND CLAUSE/SENTENCE IN MILLISECONDS IN EXPERIMENT 2

Structure	Referring expression	
	Repeated name	Pronoun
Intra-sentential	1574 (391)	1454 (328)
Inter-sentential	1679 (370)	1606 (417)

Note. Reading time standard error of the mean in parentheses.

As the results show, both experiments observed main effects of Anaphor. When the referring expression was a repeated name, the reading time of the second clause/sentence was significantly slower than when it was a pronoun. This suggests that the repeated name penalty occurs when the antecedent is embedded within subordinate clause(s). As for the factor of Structure, the two experiments had different results. A main effect was found in Experiment 2 but not in Experiment 1. Moreover, no interaction between the two factors of Structure and Anaphor was observed in either experiment, showing that anaphora processing occurs similarly intra-sententially and inter-sententially.

Part of the results of the study is consistent with previous findings. Like Gordon et al. (1999), the repeated name penalty was found within the boundary of sentences. As described in the Introduction section, their results were possibly confounded by the ungrammaticality of test sentences which violate Principle C of Binding Theory. In this

study, this potential problem was avoided by introducing the antecedent in a subordinate clause at the beginning of a sentence. Under this circumstance, the repeated name penalty was still borne out, confirming that the repeated name penalty may be observed both within and between sentences.

However, the main results of the study contradict previous predictions and findings concerning factors affecting antecedent prominence in anaphora processing. In particular, DPT assumes that a referent's prominence is dependent on its height in the phrase structure, and is inversely related to its depth of embeddedness (Gordon & Hendrick, 1998). Under this definition, the antecedent in the stimuli in the present study is not syntactically prominent, for they are embedded within one or two subordinate clauses. DPT also maintains that the repeated name penalty can only occur with reference to a prominent entity in the local discourse, predicting that pronouns have no processing advantage over repeated names when the antecedent is embedded within subordinate clauses. Yet, as the results show, the repeated name penalty was observed in both experiments, contradicting the prediction of DPT.

Therefore, current findings challenge DPT's definition of antecedent prominence and suggest that embeddedness may not be the factor at play. Other factors should be considered, particularly whether the referent is a syntactic subject and whether it is a head noun. This further casts doubt on whether phrase structure should be adopted as the vehicle of determining prominence. In Gordon and Hendrick's (1998) original accounts, syntactic prominence is formally defined in terms of x-command relations, a kind of syntactic relation dependent on a sentence's tree structure. If a referent's syntactic function (subjecthood) and status (headedness) are sufficient to determine its prominence, as suggested by this study, their resort to x-command relations may be unnecessary.

Finally, this study also examined anaphora processing both intra-sententially and inter-sententially as Gordon and Hendrick (1998) claimed that DPT applies both within and between sentences. This argument is based on the assumption that anaphora processing is a mapping from syntactic representation to discourse representation. One of the main predictions following from this argument is that the repeated name penalty occurs not only intra-sententially but also inter-sententially. This finds support in this study as repeated names are processed differently than pronoun whether in the intra-sentential or inter-sentential condition.

If the same processes underlying anaphora processing operate both within and between sentences, it should be expected to find no reading time difference between the intra-sentential and inter-sentential conditions. This prediction is supported by Experiment One, in which there was no main effect of Structure, but not by Experiment Two, in which the second line was read faster in the intra-sentential condition than in the inter-sentential condition. One possible reason for the main effect of Structure in Experiment Two could be the semantic awkwardness of the stimuli in the inter-sentential condition as a result of the lack of connectives, making participants spend more time integrating the meaning. There was no connective in the inter-sentential condition in Experiment One either, but since items in Experiment Two had more embedded clauses that items in Experiment One, the possible negative effect of lack of connectives may be larger in Experiment Two. Further research is needed to investigate this issue.

IV. Conclusion

This study intended to examine the effect of antecedent prominence on anaphora processing. By employing subordinate constructions, we tested DPT's (Gordon & Hendrick, 1998) definition of antecedent prominence, which lays emphasis on structural embeddedness. Under the account of DPT, the antecedent embedded in subordinate clauses is in a non-prominent place, thus less likely to incur the repeated name penalty in anaphora processing. However, the results of self-paced reading experiments show the opposite. The repeated name penalty is observed no matter how deeply the antecedent is embedded in subordinate clauses, calling into question DPT's definition of antecedent prominence. Given that the embedded antecedent is a head noun in the subject position, a referent's subjecthood and headedness may play a more important role in determining its prominence than its height in the phrase structure. Future studies can further compare the effects of these two factors in anaphora processing.

References

[1] Ariel, M. (1990). *Accessing noun-phrase antecedents*. Routledge.
[2] Chomsky, N. (1981). *Lectures on government and binding*. Foris Publications.
[3] Gelormini-Lezama, C. (2018). Exploring the repeated name penalty and the overt pronoun penalty in Spanish. *Journal of Psycholinguistic Research*, *47*, 377–389.
[4] Gelormini-Lezama, C., & Almor, A. (2011). Repeated Names, overt pronouns, and null pronouns in Spanish. *Language and Cognitive Processes*, *26*, 437–454.
[5] Gordon, P. C., & Chan, D. (1995). Pronouns, passives and discourse coherence. *Journal of Memory and Language*, *34*, 216–231.
[6] Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*, 311–347.
[7] Gordon, P.C., & Hendrick, R. (1997). *Intuitive knowledge of linguistic coreference*. *Cognition*, *62*, 325–370.
[8] Gordon, P. C., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science*, *22*, 389–424.
[9] Gordon, P. C., Hendrick, R., Ledoux, K., & Yang, C. L. (1999). Processing of reference and the structure of language: An analysis of complex noun phrases. *Language and Cognitive Processes*, *14*, 353–379.

[10] Grosz, B. J., Joshi, A. K., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. In M. Marcus (Ed.), *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics* (pp. 44–50). Association for Computational Linguistics.
[11] Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, *21*, 203–225.
[12] Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.
[13] Ledoux, K., Gordon, P.C., Camblin, C.C., & Swaab, T. Y. (2007). Coreference and lexical repetition: Mechanisms of discourse integration. *Memory & Cognition*, *35*, 801–815.
[14] Shoji, S., Dubinsky, S. & Almor, A. (2017). The repeated name penalty, the overt pronoun penalty, and topic in Japanese. *Journal of Psycholinguistic Research*, *46*, 89–106.
[15] Swaab, T. Y., Camblin, C. C., & Gordon, P. C. (2004). Electrophysiological evidence of reversed lexical repetition effects in language processing. *Journal of Cognitive Neuroscience*, *16*, 715–726.
[16] Yang, C. L., Gordon, P. C., Hendrick, R., & & Hue, C. H. (2003). Constraining the comprehension of pronominal expressions in Chinese. *Cognition*, *86*, 283–315.
[17] Yang, C. L., Gordon, P. C., Hendrick, R., & Wu, J. T. (1999). Comprehension of referring expressions in Chinese. *Language and Cognitive Processes*, *14*, 715–743.

Jinmei Chen is currently working at Beijing Language and Culture University, China. She holds a PhD degree from the University of South Carolina, USA, where she taught English and Spanish courses. Her research interests include discourse analysis, translation and cultural studies.

Wei Cheng is an associate professor in the College of Foreign Studies at Jinan University, Guangzhou, China, where he teaches undergraduate and graduate courses in business English and linguistics. His research interests are psycholinguistics, second language acquisition, and business language.