

Effects of discourse factors on the interpretation of Korean null pronouns in subject and object position

Jina Song & Elsi Kaiser*

Abstract. The present study investigates whether interpretation of null pronouns in Korean is best captured by a heuristics-based approach, focusing on subjecthood and syntactic parallelism, or a discourse-based coherence-relation approach, which regards pronoun resolution as a side-effect of general inferencing processes during discourse comprehension. We report two experiments where we investigated whether and how the interpretation of null pronouns in subject and object position in Korean is influenced by (i) the nature of the connective between the preceding clause and the pronoun-containing clause (*kuliko* “and” or *waynyahamyen* “because”), (ii) the presence/absence of the additive marker *-to* “also” in the pronoun-containing clause, and (iii) the presence/absence of the topic marker *-(n)un* on the subject of the preceding clause. As a whole, our results support the coherence-relation approach. We find that when a cue indicating a resemblance relation (either the connective *and* or the additive marker *also*) is present, null pronouns in both subject and object position tend to be interpreted as referring to an antecedent in a parallel syntactic position (subject and object respectively). This parallelism bias strengthens when a resemblance relation is signaled by both *and* and *also*. In contrast, when the two clauses are linked with *because* (indicating an explanation relation), null pronouns show no significant preference for either the subject or object antecedent. Topic-marking has no effect, possibly due to lack of context. Our study provides new evidence that both subject- and object-position null pronouns are sensitive to both syntactic and discourse-level factors.

Keywords. pronoun interpretation, reference resolution, null pronouns, Korean, coherence relations, subjecthood, topic marker

1. Introduction. One of the crucial aspects of successful communication is the correct interpretations of underspecified referring expressions such as null and overt pronouns. In contexts where multiple potential antecedents are present, as in the Korean example in (1), how do comprehenders interpret null pronouns which do not, on their own, provide sufficient information to identify the intended referent?

- (1) John-*i* Tom-*ulj* p^wull-ess-ta. Jack-*i* pro_{obj i/j} paysinha-yss-ta.
 John-Nom Tom-Acc call-Past-Dec Jack-Nom betray-Past-Decl
 ‘John_i called Tom_j. Jack betrayed pro_{i/j}.’

Broadly speaking, prior work on reference resolution has tended to pursue two main lines of inquiry: a heuristics-based approach and a coherence relation approach. On the one hand, *heuristics-based* approaches focus on identifying the interpretation biases and/or processing strategies that guide pronoun resolution. Indeed, various heuristic strategies have been suggested including the subject preference, the first-mention preference, and the parallel function preference (e.g.,

* The authors would like to thank the members of the USC Language Processing Lab and the audience at the 94th Annual Meeting of the LSA for helpful feedback and comments on this work. Authors: Jina Song, University of Southern California (jjinas@usc.edu) & Elsi Kaiser, University of Southern California (emkaiser@usc.edu).

Givón 1983, Ariel 1990, Smyth 1994, Stevenson Crawley & Kleinman 1994, Arnold 2001, Crawley, Stevenson & Kleinman 1990). There is often an implicit (sometimes explicit) assumption that these strategies are specific to pronoun resolution and/or that pronoun resolution is a consequence of specific ‘search strategies.’ On the other hand, according to *coherence relation approach*, pronoun resolution is a by-product of more general reasoning to make the most coherent discourse-level interpretation, without postulating strategies specifically for pronoun resolution (e.g., Hobbs 1979, Kehler 2002, Wolf, Gibson & Desmet 2004, Kehler Kertz, Rohde & Elman 2008, Rohde & Kehler 2008, Kaiser 2011, Ueno & Kehler 2016).

While these approaches have informed work on how overt pronouns are interpreted, less attention has been paid to the question of how different factors influence null pronoun resolution. Most of the work on null pronoun resolution has largely focused on (i) null pronouns in subject position – which are crosslinguistically more frequent than null pronouns in object position – and on (ii) only one of the heuristic strategies, namely the subject preference (e.g. Alonso-Ovalle et al. 2002, Filiaci 2010, Kweon 2011, Filiaci, Sorace & Carreiras 2013, Fedele, 2016). Only a few studies have investigated how coherence relations might guide the interpretation of null pronouns (Ueno & Kehler 2016, Kim, Grüter & Schafer 2013). Thus, it is not yet clear to what extent factors other than syntactic position (e.g., the subject preference) guide interpretation of null pronouns, and more generally, whether the underlying mechanism guiding null pronouns resolution is best captured by the heuristics-based approach or the coherence relation approach.

In the present study, we aim to shed light on these issues. To this end we conducted two offline experiments in Korean by employing three semantic/discourse factors that contribute to the establishment of discourse-level coherence relations: (a) the connectives *waynyahamyen* ‘because’ (signals an explanation relation) and *kuliko* ‘and’ (can signal a parallel/resemblance relation), (b) the additive marker *-to* ‘also’ which signals a parallel/resemblance relation, and (c) the topic marker *-(n)un*, which can render nouns more topical. If the interpretation of null pronouns is influenced only by the grammatical role of the antecedent, these factors should have no effect on pronoun interpretation. However, if null pronoun resolution is sensitive to discourse-level information such as coherence relations, we expect to see modulating effects of these three factors. Our studies were conducted in Korean. Korean is well-suited for our aims, as it has null pronouns in both subject and object position, and thus differs from Romance languages which typically only have pro-drop in subject position. Testing null pronouns in both subject and object position is important, as it allows us to dissociate subjecthood effects from parallelism effects.

The structure of this paper is as follows: In Sections 2 and 3, we present the heuristics-based approach and the coherence relation approach to pronoun resolution. In Section 4, we outline our research questions and predictions. Sections 5 and 6 present the experiments that we conducted, and Section 7 is a general discussion that presents our overall results and conclusions.

2. Heuristics-based approach. Prior work on pronoun resolution has identified various preferences that guide pronoun resolution. For instance, one strategy that has been proposed for pronouns in subject position is a preference for antecedents that are also realized in subject position (the ‘subjecthood’ heuristic). Carminati (2002) outlines a specific version of this hypothesis, the Position of Antecedent Hypothesis (PAH) for null pronouns, based on her work on Italian.

- (2) Position of antecedent hypothesis (Carminati 2002)
A null pronoun prefers an antecedent in the Spec IP position, an overt pronoun prefers an antecedent not in the Spec IP position.
- (3) a. Quando Mario_i ha telefonato a Giovanni_j, **pro**_i aveva appena finite di mangiare.
When Mario_i telephoned Giovanni_j **pro**_i had just finished eating.

- b. Quando Mario_i ha telefonato a Giovanni_j, lui_j aveva appena finito di mangiare.
 When Mario_i telephoned Giovanni_j he_j had just finished eating.
 ‘When Mario telephoned Giovanni, (he) had just finished eating.’

The PAH states that in a language with null and overt pronouns, pronoun resolution is guided by the syntactic position of the antecedent. Null pronouns prefer antecedents in subject position (Spec IP); overt pronouns prefer antecedents in a non-subject position that is structurally lower than Spec IP. In the Italian example (3), the subject *Mario* is preferred as the referent of the null pronoun (3a), and the object *Giovanni* is preferred as the referent of overt pronoun (3b). Carminati provides evidence from self-paced reading experiments and questionnaire studies.

However, this account leaves open two key questions that need to be addressed to better understand the mechanism underlying null pronoun resolution. First, most studies on the PAH focused on null pronouns in subject position, and little attention has been paid to null pronouns in other positions. Because researchers have investigated cases in which subject null pronouns always have the same grammatical role as the preferential antecedent (i.e., the preceding subject entity), it is difficult to tell whether the subject antecedent preference is driven by the PAH or the parallel function strategy (PFS). The PFS is another type of heuristics-based approach, defined in (4) (e.g., Smyth 1994, Stevenson et al. 1995, Chambers & Smyth 1998).

- (4) Parallel function strategy (Smyth 1994)

A pronoun with two or more grammatically and pragmatically possible antecedents in a preceding clause will be interpreted as coreferential with the candidate that has the same grammatical role.

- (5) a. Jim_i hit Jason_j, and then he_i kicked Bill.
 b. Jim_i hit Jason_j, and then Bill kicked him_i.

The PFS assumes that the grammatical role of the pronoun relative to its antecedent guides pronoun interpretation: a pronoun tends to be coreferential with a referent with the same grammatical role. Thus, subject- and object-position pronouns are predicted to have different preferences: a subject pronoun prefers to refer to a subject antecedent and an object pronoun prefers to refer to an object antecedent, as in (5). Thus, without investigating null pronouns in non-subject position, we cannot tell whether the previous results of the subject antecedent preference of null pronouns is driven by the PAH or the PFS.

The second key question that is still open concerns the nature of the mechanism that guides pronoun resolution. One aspect of this question has to do with whether the interpretation of a particular referring form is guided by a single constraint/strategy, or whether multiple constraints/strategies can be at play (see also Kaiser & Trueswell 2008 for related discussion). In her discussion of the PAH, Carminati (2002) states that at least in Italian, non-structural factors, such as the semantics of the antecedent, do not significantly influence the antecedent preferences of null pronouns. In her view, the interpretation of null pronouns in Italian is mostly guided by structural/configurational factor (Carminati 2002:58). Similarly, if we assume the strict version of the PFS as expressed in (4), on its own the PFS does not make explicit predictions about specific effects of semantic or discourse factors on null pronoun interpretation.

However, recent evidence shows that null pronoun resolution is affected not only by syntactic but also by semantic and discourse factors (e.g., Kim et al. 2013, Kwon & Sturt 2013). For example, an experiment on Korean null pronouns by Kim et al. (2013), building on work in English by Rohde, Kehler & Elman (2006), found that the aspect of the verb in the preceding

clause (imperfective vs. perfective) had an influence on null pronoun resolution in transfer-of-possession contexts (e.g., *X was handing a book to Y* vs. *X handed a book to Y*). Kim et al. (2013) found that the perfective aspect had a significantly weaker effect on the subject antecedent preference than the imperfective aspect because the Goal (the object) became prominent in the completed event (see also Rohde et al. 2006). Further evidence for effects of semantic factors on null pronoun resolution comes from Japanese (Ueno & Kehler 2016) and Italian (Fedele 2016), who found effects of verb semantics on the interpretation of null pronouns in subject position. In these studies, null subject pronouns tended to refer to the *object* antecedents with object-biased implicit causality verbs (i.e., verbs that impute the cause of the events to the object entity). Moreover, Kwon & Sturt (2013) found that Korean null pronouns in subject position are insensitive to morphological feature mismatch with potential antecedents if the null pronoun is already bound to a discourse topic – in other words, topicality plays a role in guiding null subject interpretation. As a whole, these findings show that null pronoun interpretation is not guided by a single factor or by purely structural considerations.

3. Coherence relations approach. Contrary to the heuristics-based approach, the coherence relations approach explicitly argues that pronoun resolution is a by-product of establishing the most coherent interpretation of utterances in discourse (e.g., Hobbs, 1979, Kehler, 2002, 2008). This approach assumes that a key property of natural language understanding is identifying the most likely relations between sentences. Thus, according to this view, pronouns are interpreted as part of the larger process of making inferences about the most coherent discourse-level interpretation, including what kinds of semantic relations hold between the sentence containing a (null) pronoun and the sentence containing the antecedent.

Among various classes of coherence relations, *explanation* and *resemblance* relations have been widely discussed relative to pronoun resolution (e.g., Wolf et al. 2004, Kertz et al. 2006, Kehler et al. 2008, Rohde & Kehler 2008, Kaiser 2011). The *explanation* relation holds if there is a causal link between sentences and the *effect* sentence precedes the *cause* sentence. It is usually signaled by the connective *because*. This coherence relation is established by inferring the most plausible causal relation between two sentences, and the inference process is “contingent on other properties being true of the world” (Kehler 2002:20). Thus, the *explanation* coherence relation is driven by semantics and world knowledge. In the process of the inference, pronouns are resolved by being bound to their referents. Consider (6) from Winograd (1972):

- (6) a. The city council_i denied the demonstrators_j a permit because they_j advocated violence.
 b. The city council_i denied the demonstrators_j a permit because they_i feared violence.

For example, the subject pronoun *they* in (6a) preferentially refers to *the demonstrators*; whereas *they* in (6b) preferentially refers to *the city council*. This difference between (6a,b) is not predicted by the heuristics we have discussed so far, since there is no change in the grammatical roles of the pronouns or the antecedents from (6a) to (6b). Instead, the difference in interpretation stems from reasoning based on the explanation relation and real-world knowledge.

It is worth noting that, because the establishment of the *explanation* relation is contingent on other properties being true of the world, we expect that without semantic information and available world knowledge, fully-specified *explanation* coherence relations cannot be inferred.

Another common coherence relation is the *resemblance* relation, which holds between sentences if the events in the sentences are similar: The comprehender makes “an inference on the commonalities among corresponding relations and entities in the two sentences” (Kehler, 2002:16). Thus, the *resemblance* coherence relation is made by identifying and pairing sets of

parallel arguments and events. In this coherence relation, the pronouns are interpreted by being bound to the parallel arguments as a way of maximizing similarity.

- (7) Margaret Thatcher_i admires Hillary Clinton_j, and George Bush worships her_j.
 (i) admire (Thatcher, Clinton) (ii) worship (Bush, her) (Kehler 2002:157)

The most likely coherence relation between the sentences in (7) is inferred to be the *resemblance* relation, as signaled by the connective *and* and the verbs with similar meanings. The *resemblance* relation is established by identifying semantically parallel predicates (*admire*, *worship*) and arguments between the propositions. Thus, *her* is preferentially interpreted as *Hillary Clinton*, as they are semantically parallel arguments. This interpretation reveals how strongly the joint presence of ‘and’ plus similar verb semantics signal a resemblance interpretation, because this interpretation actually contradicts world knowledge. (In reality, Bush is more likely to worship Thatcher). In sum, the inference process in the *resemblance* relation focuses on finding semantically parallel entities and events (Kehler 2002:158), seemingly even without real-world knowledge.

Although many studies have investigated how a shift in coherence relations yields a different pattern of *overt* pronoun interpretation (e.g., Kehler 2002, Wolf et al., 2004, Kertz et al., 2006, Kehler et al. 2008, Rohde & Kehler 2008, Kaiser 2011), very few studies have looked at effects of coherence on *null* pronouns (but see Ueno & Kehler 2016). Our work aims to contribute in this area.

4. Research question and predictions. In the present study, the questions in (8) are addressed by investigating how Korean null pronouns are resolved in an ambiguous context, where two potential antecedents (i.e., the subject NP and the object NP) are present. We used a variety of nonce verbs to eliminate effects of verb semantics. (9) shows one of the test items we used. (The Korean verbs are nonce verbs; we have indicated this by using the nonce verbs ‘daxed’ and ‘blicked’ in English, although the nonce verbs are of course all meaningless. Cf. Hartshorne and Snedeker (2013) for use of the nonce word ‘dax’ in pronoun-resolution work on English.)

- (8) a. What is the underlying mechanism for interpreting null pronouns in the object and subject position? Is this mechanism guided by the heuristics-based approach or by the coherence relation approach?
 b. What factors affect the interpretation of null pronouns and how do they interact?

- (9) Tom-i/nun Jack-ul tayksuhayss-ta.
 Tom-Nom/Top Jack-Acc daxed-Dec.
 Waynyahamyen/kuliko Bill-i/-to pullikhayss-ki ttaymwun-i-ta.
 Because/and Bill-Nom/Also blicked-Nml reason-be-Dec
 ‘Tom-Nom/-Top daxed Jack. Because/And Bill-Nom/-Also blicked.’

We manipulated three semantic and discourse factors that have been claimed to contribute to the discourse-level coherence relations: (i) the connective type (*and* or *because*), (ii) the presence/absence of the additive marker, and (iii) the presence/absence of the topic marker. If null pronoun resolution is governed purely by the PAH (i.e., a subjecthood heuristic) and/or purely by a simple version of the PFS that is based on parallel structural positions (i.e., a parallel function heuristic), we expect to see effects of the pronouns’ and antecedents’ structural positions and/or grammatical roles, but no effects of the three semantic and discourse-related factors. In contrast, if null pronoun resolution is governed by discourse-level coherence relations, we expect that the coreferential patterns to be modulated by each of these three factors.

Let us now preview our experimental manipulations. First, we manipulated **the type of the connective** that links the pronoun containing sentence and the preceding sentence that contains

the potential antecedents. This was done to cue different coherence relations, as in (9). In Experiment 1, we used the connective *waynyahamyen* ‘because’ to signal an *explanation* relation. In Experiment 2, we used the connective *kuliko* ‘and’ to signal a *resemblance* relation. Because we used nonce verbs in both clauses, this means that constructing a fully fleshed-out explanation relation is not expected to be possible. (Recall the discussion above regarding the key role of real-world knowledge for reasoning about explanation relation.) In contrast, examples such as (7) show that parallel relations can be constructed even if they go against real-world knowledge.

Thus, given prior work, we predict that – if pronoun resolution *is* sensitive to coherence relations – in sentences like (9) with nonce verbs, pronouns cannot be properly resolved in the presence of a *because* connective but can be resolved with a *resemblance* connective (as long as the basic argument structure of the predicates is provided). We expect no bias toward either of the two potential antecedents in sentences like (9) with a *because* connective¹, but do expect a parallel antecedent bias with the ‘and’ connective that signals a *resemblance* relation.

Second, we manipulated the morphological marker in the sentence with the null pronoun: we used the **additive particle -to ‘also’ or the nominative marker -i/-ka** on the overt noun, as in (9). We assume that the additive particle *-to* ‘also’ signals a *resemblance* relation due to its semantics: It triggers the presupposition that there is an entity distinct from the individual denoted by the *NP-to* that satisfies the predicate, as shown in (10).

(10) $[[x-to]] = \lambda P.\lambda x.P(x) \wedge \exists z: [[P(z)]] = 1$ (Lee, Park & Madigan 2015: 236)

If there is a preceding sentence that denotes the same/similar property as the sentence with the additive marker, the presupposition triggered by the marker will be satisfied. Because the events described by the sentences in (9) (i.e., a presuppositional sentence and a sentence with the additive marker) are similar in their meaning and argument structure, we expect a *resemblance* relation will be inferred. Thus, if the additive marker is used to interpret null pronouns, we expect a parallel antecedent bias with the additive marker, but not with the nominative marker.

Third, we manipulated **the presence/absence of topic marking** in the antecedent-containing sentence by marking the subject NP with the topic marker *-(n)un* or the nominative marker *i/-ka* (ex.9). It has been argued that pronouns prefer topical antecedents (e.g., Ariel, 1990; Gundel et al., 1993; Grosz et al., 1995). This leads us to expect subjects to be interpreted as antecedents more often when they are topic-marked than when they are nominative-marked.

Fourth, we investigate null pronouns in both **subject and object position**.

- (11) Tom-i/nun Jack-ul tayksuhayss-ta.
 Tom-Nom/Top Jack-Acc daxed-Dec.
 Waynyahamyen/kuliko Bill-i *pro_{obj}* pullikhayss-ki ttaymwun-i-ta.
 Because/and Bill-Nom blicked-Nml reason-be-Dec
 ‘Tom-Nom/-Top daxed Jack. Because/And Bill-Nom blicked *pro_{obj}*.’ [object position]
- (12) Tom-i/nun Jack-ul tayksuhayss-ta.
 Tom-Nom/Top Jack-Acc daxed-Dec.
 Waynyahamyen/kuliko (*pro_{subj}*) Bill-to (*pro_{obj}*) pullikhayss-ki ttaymwun-i-ta.
 Because/and Bill-Also blicked-Nml reason-be-Dec
 ‘Tom-Nom/-Top daxed Jack. Because/And *pro_{subj}* blicked Bill-Also.’ [subject position]
 ‘Tom-Nom/-Top daxed Jack. Because/And Bill-Also blicked *pro_{obj}*.’

¹With an *explanation* relation, a subsequent pronoun can refer to a preceding subject *or* object, depending on verb meaning (e.g. implicit causality, Caramazza et al. 1977, McKoon et al., 1992, Rohde 2008, i.a.). Thus, a meaningless nonce verb in this context could act like a subject- or object-biased implicit causality verb. Thus, in an ‘X because Y’ frame, presence of a nonce verb in X means a pronoun in Y could refer to the subject or object in X.

Specifically, by using the additive particle *-to* ‘also’ and the nominative marker *-i/-ka* in the sentence containing a null pronoun, we are able to manipulate (i) whether the null pronoun is clearly in object position or (ii) whether the grammatical role of the null pronoun is *ambiguous* between subject and object. When the overt noun has the nominative marker *-i*, the null pronoun is clearly in object position, as (11). However, when the overt noun has the additive particle *-to*, its grammatical role could be subject or object, as in (12). Consequently, the grammatical role of the null pronoun could also be object or subject. In other words, the grammatical position of the null pronoun is ambiguous in the presence of the additive marker.² We used images (see Fig.1) to make clear that the nonce verb in the second clause is *transitive*, not intransitive. This made it clear to participants that the second clause has a null argument in subject or object position

By examining both subject and object null pronoun resolution, we can see whether pronoun resolution is being guided by the PAH or the PFS, or by discourse-level coherence information. If the PAH is the underlying mechanism, we expect a subject antecedent preference with both subject and object null pronouns. However, under the PFS, we predict a subject antecedent preference with subject null pronouns, but an object antecedent preference with object null pronouns.

5. Experiment 1: ‘Because’ connective.

5.1. PARTICIPANTS. Forty-two self-identified native speakers of Korean, educated in Korea, participated over the internet via Qualtrics (www.qualtrics.com). None had lived outside Korea for more than 10 years. Two people who did not complete the questions or had lived outside Korea for over 10 years were excluded.

5.2. MATERIALS AND DESIGN. The two sentences in each target item are connected with the connector *waynyahamyen* ‘because.’³ We used a 2 × 2 design and manipulated the topicality of the subject in the first sentence and the additive/ nominative status of the noun in the second sentence (Table 1). First, the subject NP of the first sentence had either a nominative or a topic marker. We refer to this nominative marker as **Nom1** and the topic marker as **Top**. Second, the NP of the second sentence had a nominative marker or the additive particle *-to*. We refer to this nominative marker as **Nom2**, and the additive particle as the **Also** marker. We used nonce verbs in all target items to exclude effects of verb semantics.⁴ The nonce verbs were designed to be pronounceable in Korean, and were combined with the active form of the verb endings to ensure an active interpretation (i.e. could not be interpreted as passives). The images (Fig.1) indicated the verbs are transitive. The gender of the two names in each sentence was the same. Twenty-four target and 36 filler sentences were presented using a Latin-Square design.

² In (12), when *NP-also* is interpreted as in subject position, it has a presupposition that *somebody other than Bill blicked pro*. When *NP-also* is interpreted as being in object position, it means *pro blicked somebody other than Bill*.

³ As shown in Table 1, our Korean target items consist of two different sentences, and not two different clauses: In Korean, both “And” and “Because” can start a new sentence. *Waynyahamyen* ‘because’ is best translated into English as “The reason for this is...”. We used two distinct sentences in Korean because the causal connective *waynyahamyen* appears most naturally at the start of a new sentence.

⁴ Effects of verb semantics are a key question for future work. The current studies focus on semantic and coherence-related cues from the connectives and the topic-marker, but verb semantics should be addressed in future work.

Sentence1	Sentence2	Examples
Nom1	Nom2	Tom- i Jack-ul tayksuhayss-ta. waynyahamyen Bill- i pullikhayss-ki ttaymwun-i-ta. Tom- Nom Jack-Acc daxed-Dec. Because Bill- Nom daxed-Nml reason-be-Dec 'Tom daxed Jack. Because Bill blicked <i>pro</i> .'
	Also	Tom- i Jack-ul tayksuhayss-ta. waynyahamyen Bill- to pullikhayss-ki ttaymwun-i-ta. Tom- Nom Jack-Acc daxed-Dec. Because Bill- Also daxed-Nml reason-be-Dec 'Tom daxed Jack. Because Bill also blicked <i>pro</i> . // <i>pro</i> also blicked Bill'
Top	Nom2	Tom- un Jack-ul tayksuhayss-ta. waynyahamyen Bill- i pullikhayss-ki ttaymwun-i-ta. Tom- Top Jack-Acc daxed-Dec. Because Bill- Nom daxed-Nml reason-be-Dec 'As for Tom, Tom daxed Jack. Because Bill blicked <i>pro</i> .'
	Also	Tom- un Jack-ul tayksuhayss-ta. waynyahamyen Bill- to pullikhayss-ki ttaymwun-i-ta. Tom- Top Jack-Acc daxed-Dec. Because Bill- Also daxed-Nml reason-be-Dec 'As for Tom, Tom daxed Jack. Because Bill also blicked <i>pro</i> // <i>pro</i> also blicked Bill.'

Table 1. A Set of Example Stimuli in Experiment 1

5.3. PROCEDURE. We used a picture-writing task that asks people to write down the names of the participants of the event in the picture (Figure 1). Every item consisted of simultaneous presentation of (i) the two sentences and (ii) a schematized picture that depicted the event of the underlined part of the sentence (Fig. 1). In targets, the second sentence was underlined. Target pictures consisted of two person-shaped stick figures that represented the event participants, and an arrow denoting the action of the verb in the second sentence. Participants were instructed that the stick figure at the start of the arrow is the initiator of the action (Agent), and the icon at the end of the arrow is the entity that underwent the action (e.g. Theme).⁵ As target images always included two figures, it was clear that the nonce verbs express an interpersonal *transitive* action.

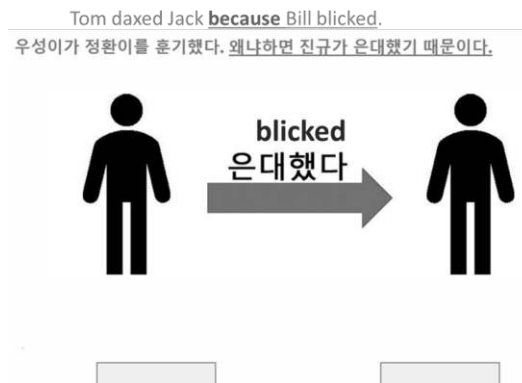


Figure 1. Sample target stimuli used in Experiment 1. (The English translations are included here for ease of exposition, and were not shown to participants).

Participants were asked to type the name of each person in the associated text box so that the picture describes the meaning of the underlined the sentence. In Figure 1, if a participant thought that *the person who blicked* is *Bill* and *the person who was blicked* is *Jack*, they would write *Bill* in the left box and *Jack* in the right box. Thus, based on people's responses, we can determine how they interpreted the null pronoun in the critical sentence.

5.4. PREDICTIONS. The main research question of the present study is whether null pronouns in subject and object position are interpreted based on a grammatical-role based heuristics-based

⁵ The direction of the arrows (i.e., leftward or rightward) is counterbalanced in both experiments.

strategy (specifically, PAH or PFS) or discourse-level coherence relations between sentences.

We predict that if the null pronoun searches for its antecedent simply based on grammatical role (a structural-heuristics based approach), semantic and discourse information should not matter. PAH predicts a uniform subject antecedent preference. The PFS predicts a uniform parallel grammatical role preference: Object null pronouns should prefer the preceding object, and subject null pronouns should prefer the preceding subject, equally in all conditions.

However, if null pronoun resolution is a side effect that results from general discourse-level processing (the coherence relation approach), we expect that semantic and pragmatic information can influence pronoun resolution. This point will also become important once we compare Experiments 1 and 2. For now, recall that all targets in Experiment 1 use *waynyahamyen* ‘because’ (signalling an *explanation* relation). As discussed in Section 4, since we purposefully used nonce verbs, the connective *because* is not expected to elicit a clear bias for the subject or object.

In addition to the connectives, if the additive marker *-to* ‘also’ triggers a *resemblance* relation between the sentences, we expect null pronouns in the Also condition to be coreferential with the semantically parallel argument. However, in the Nom2 condition, we predict no preference for either of the two antecedents, since neither the connective *because* nor the nominative marker in the second sentence mark the subject or object as a semantically-parallel argument.

What about the topicality of the first-sentence subject, marked with the topic marker or the nominative marker? If discourse-level factors, and not just grammatical role information, guide pronoun resolution, one might expect null pronouns to show a stronger preference for subjects marked as topics than subjects not marked as topics (see discussion in Ueno & Kehler 2016).

5.5. RESULTS AND DISCUSSION. Overall, the results support the coherence relation approach. There was no uniform pattern of null pronoun interpretation across conditions. Figure 2 shows the proportion of subject and object referents for all conditions averaged over participants. This figure provides information both about (i) the grammatical role of the antecedent of the null pronoun and (ii) the grammatical role of the null pronoun itself:

First, as regards the **grammatical role of the antecedent**, the light grey parts show the proportion of trials on which participants interpreted the null pronoun as referring to the preceding object. The darker parts of the bars show the proportion of trials on which participants interpreted the null pronouns as referring to the preceding subject position. Thus, we separate the data into *sbj_antecedent* (participant interpreted null pronouns as referring to preceding subject) and *obj_antecedent* (participant interpreted null pronouns as referring to preceding object).

Second, as regards the **grammatical role of the null pronoun itself**, the label immediately to the left of each bar indicates the grammatical role assigned to the null pronoun. In the conditions where the NP in the second sentence is marked with nominative (*Nom2*), the null pronoun is unambiguously in object position (*Nom2 + Pro_Obj*). When the NP in the second clause has the additive *-to* ‘also’ marker, the null pronoun could be interpreted as subject (*Also + Pro_Sbj*) or as object (*Also + Pro_Obj*), so we show these responses separately.

The Nom1 and the Top conditions are collapsed in Figure 2: Topic marking had no effect on the rate of subject choices in either the Nom2 ($p=.26$, glmer) or the Also ($p=.63$) conditions.

As Figure 2 shows, when the noun in the second sentence is nominative, the object position null pronoun has no clear preference for the preceding subject (44%) or object (56%) (*Nom2* bar at the top of Fig. 2). These patterns do not differ from chance (50%; intercept-only glmer, $p=0.1$): There is no overarching subject preference, contrary what one might expect based on the PAH, and no overarching object preference, contrary to the PFS.

In contrast, when the noun in the second sentence has the additive marker *-to* ‘also’ – i.e.,

when the grammatical role of the null pronoun could be subject or object – the antecedent choice patterns are quite different depending on the grammatical role that participants assign to the null pronoun (two lower bars, Fig. 2). Here, we observe *parallelism effects*: Subject-position null pronouns are biased towards subject antecedents (67% subject choices); object-position null pronouns are biased towards object antecedents (83% object choices). Indeed, the rate of subject choices for subject-position null pronouns is significantly higher than chance, and the rate of object choices for object-position pronouns is significantly higher than chance (p 's<.001).⁶ The additive marker triggers parallelism effects, as we predicted.

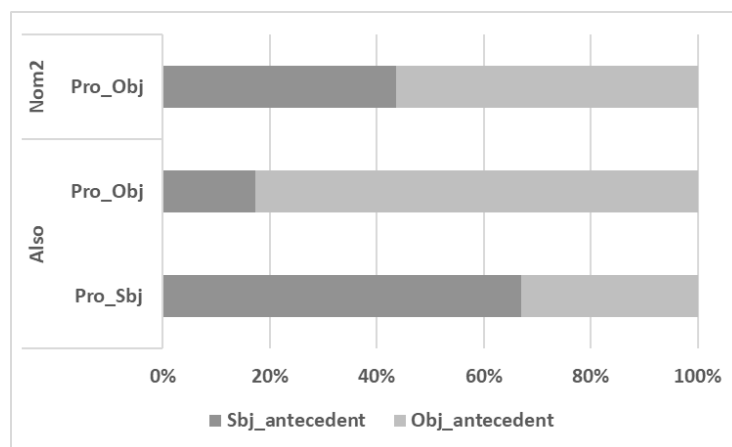


Figure 2. Experiment 1: Percentage of subject and object antecedent choices. (The Nom1_Nom2 and Top_Nom2 conditions, and the Nom1_Also and Top_Also conditions are collapsed)

In sum, then, we find that (i) the additive marker, signaling a *resemblance* relation, triggers parallelism effects, but (ii) on its own, without verb semantics, the connective *because* does not elicit a clear referential preference – which is what we expected, given the absence of verbal semantic information and world knowledge.

However, we found no effects of the topic marker, which is unexpected if pronoun resolution is sensitive to discourse-level information. The absence of a topicality effect might stem from the subject being both an Agent and also a Causee (due to the *explanation* relation). Because a Causee is less topical than an Agent (Hawkinson and Hyman 1974, Givón 1976), this might have obscured topicality effects. The lack of a topicality effect may stem from the absence of discourse context (see also Ueno & Kehler 2016 on Japanese). We return to this below.

6. Experiment 2: ‘And’ connective. Experiment 2 uses the connective *and* (indicating a *resemblance* relation) instead of the connective *because*. This allows us to further test how null pronoun resolution is guided by discourse-level coherence, and how null pronoun resolution varies by coherence relation type.

6.1. PARTICIPANTS. Forty-two self-identified native speakers of Korean, educated in Korea, participated over the internet. None had lived outside Korea for more than 10 years. Six participants

⁶ To test the statistical significance of this pattern, we used one-sample *t*-tests to assess whether the rate of subject vs. object antecedent choices differs from chance. Because there are four different interpretational options, we define chance as 0.25: Participants could interpret null as being (i) in subject position and referring to preceding subject, (ii) as being in subject position and referring to preceding object, (iii) as being in object position and referring to preceding subject, or (iv) as being in object position and referring to preceding object.

were excluded because they did not complete the questions, and four participants were excluded to balance the number of participants per list.

6.2. MATERIALS AND PROCEDURE. We used the same materials and procedure as in Experiment 1 (Table 1), except for the connective type: Now, the two sentences of each target item were connected with the sentential connector *kuliko* ‘and’, instead of *waynyahamyen* ‘because’.

6.3. PREDICTIONS. The predictions regarding effects of grammatical role-based resolution heuristics are the same as in Experiment 1: Under the PAH, a uniform subject antecedent preference is expected; under the PFS, a uniform parallel grammatical role preference is expected.

However, the coherence relation approach predicts different patterns of null pronoun resolution in Experiments 1 and 2 due to the connective. The connective *kuliko* ‘and’ will trigger a *resemblance* relation between the events described in the sentences. Given that in the *resemblance* relation, null pronouns are interpreted to maximize similarity between sentences, we predict they will tend to be coreferential with the parallel entity across conditions. We predict that not only in the Also condition (where additive *-to* ‘also’ signals the *resemblance* relation) but also in the Nom2 condition, parallel antecedents will be preferred over non-parallel ones.

At first glance, this uniform parallel preference may appear to be the same as the prediction of the PFS. However, what is crucially different from the PFS is that under a coherence relation approach, the strength of the parallelism can vary depending on connective type. The PFS (at least in its traditional form) does not explicitly predict differences in the strength of the parallelism effect between Experiment 1 (*because*) and Experiment 2 (*and*). More specifically, prior work investigating the PFS has focused on the ‘and’ connective and does not make any claims about different kinds of connectives. Thus, the PFS (at least in its simplest form) makes no predictions about differences between connectives. In contrast, the coherence relation approach predicts that stronger parallelism effects will be found in Experiment 2 than Experiment 1: In Experiment 2, the *resemblance* relation is signaled even in the Nom2 condition by the connective *and*. Furthermore, in the Also condition of Experiment 2, the *resemblance* relation is signaled by the use of multiple cues, namely the additive marker *-to* ‘also’ and the connective *and*. Thus, we expect an overarching parallel-antecedent preference in both the Nom2 condition as well as the Also condition, perhaps further strengthened in the Also condition.

Similar to Experiment 1, we had also expected to see a preference for topical subjects to be interpreted as antecedents more often than non-topical subjects. In Experiment 2, the subject antecedent does not perform a Causee role but only an Agent role, which may make effects of topicality easier to detect (see Section 5.5). However, in light of the results of Experiment 1, it may be that we will again fail to detect effects of topicality, perhaps due to a lack of context.

6.4. RESULTS AND DISCUSSION. The results of Experiment 2 support the coherence relation approach. Figure 3 shows the percentages of subject and object antecedents for each condition, collapsing the Top and Nom1 conditions. We find that there is no overarching subject antecedent bias; a bias for subject-position antecedents emerged only when null pronouns were also in subject position. Furthermore, we observe a preference for parallel antecedents in all of the conditions in Experiment 2, as predicted by the PFS and the coherence relation approach.

Let’s take a closer look at Figure 3. When the null pronoun was unambiguously in object position, it was interpreted as having an object-position antecedent 90% of the time (*Nom2* condition, topmost bar), which is significantly higher than chance (0.5, intercept-only glmer, $p > .005$). What about the *Also* conditions, where participants can interpret the null pronoun as being in subject *or* object position? When participants interpreted the null as being in object

position (middle bar), it was interpreted as referring to the preceding object 87% of the time. Conversely, when participants interpreted the null as being in subject position (lowest bar), it was interpreted as referring to the preceding subject 98% of the time. The parallelism effect is statistically confirmed: the rate of subject interpretation for subject-position null pronouns is significantly higher than chance, and the rate of object interpretation for object-position pronouns is significantly higher than chance ($p's < .01$) (see footnote 8 for statistics details).

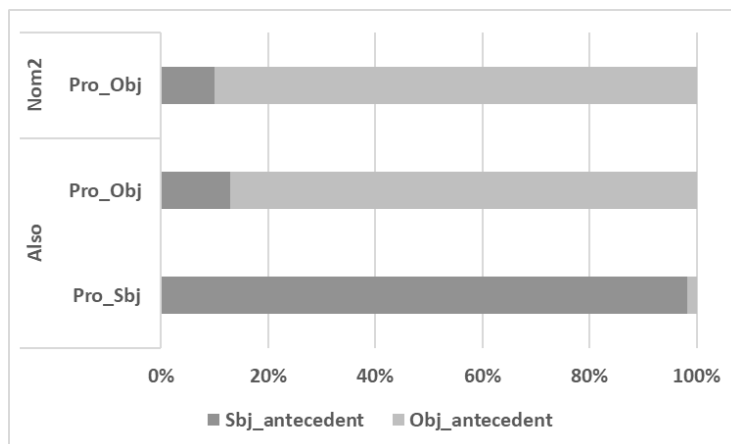


Figure 3. Experiment 2: Percentage of subject and object antecedent choices. (The Nom1_Nom2 and Top_Nom2 conditions, and the Nom1_Also and Top_Also conditions are collapsed)

The results of the parallel function preference we observed in Experiment 2 appear to fit with both the PFS and the coherence relation approach. However, when the strength of the parallelism effect is compared with Experiment 1 (see Figure 2), we find a significantly stronger parallelism preference in Experiment 2 than Experiment 1 (glmer, $p < .001$ in the Nom2 and the Also condition). Thus, null pronouns were more likely to be coreferential with the antecedent with a parallel function in Experiment 2. This supports the predictions of the coherence approach, but not the PFS: only the former predicts that the semantic and discourse information of connectives takes part in guiding null pronoun resolution. Thus, we conclude that the information from connectives and the additive marker guides inferencing about discourse relations between sentences, and null pronouns are interpreted as part of this inferencing process.

As in Experiment 1, we found no topicality effect in Experiment 2 (glmer, $p = .24$ in the Nom2 condition, $p = .16$ in the Also condition). Because the topicality effect was not found in either the *resemblance* relation or the *explanation* relation, we assume that the topicality effect is absent not because the topicality was obscured by less topical semantic role of the subject antecedent (i.e., Causee), but because the marker itself was not strong enough (in the absence of further context) to determine the topic entity compared to the discourse topic. This seems to echo the results of Ueno & Kehler (2016)'s work on subject-position null pronouns in Japanese.

7. General discussion. We conducted two studies on Korean to investigate (i) the general mechanisms that guide the interpretation of null pronouns in subject and object position, and (ii) the specific factors that are part of this mechanism. As a whole, our results fit with a coherence-based approach – i.e., pronoun resolution is a result of establishing discourse-level coherence.

One of the crucial findings in favor of a coherence relation approach is that a shift in coherence relations yields different resolution patterns: Recall that in discourses with an *explanation* relation (Experiment 1, Nom2 condition), null pronouns that are unambiguously in object position showed no preference for the preceding subject or object. However, in discourses with a

resemblance relation (Experiment 2, Nom2 condition), null pronouns that are unambiguously in object position showed a very strong preference for the parallel antecedent. Moreover, Experiment 2 also shows that the preference for the parallel antecedent is strengthened when the *resemblance* coherence relation was signaled by multiple cues (the additive marker and the connective *and*). These patterns cannot be explained by a heuristics-based approach that relies only on a Position of Antecedent Hypothesis (PAH) or a simple version of the Parallel Function Strategy (PFS). Even if the PFS captures the parallel antecedent bias, it does not straightforwardly predict the varying strengths of parallelism depending on which coherence relation holds between sentences (Experiments 1 vs. 2). Our findings imply that null pronoun resolution is a by-product of more general reasoning to make the most coherent discourse-level interpretation.

Let us now consider more closely the contributions of specific factors. Our results are not compatible with a view that *a single factor* (e.g. syntactic position or grammatical role parallelism) guides pronoun resolution. Instead, our results suggest that *multiple factors* contribute to establishing the coherence relations that guide null pronoun interpretation (see also Kaiser & Trueswell 2008). First, we find effects of connective type: As in previous work on overt pronouns, the semantic information in connectives guides interpretation of null pronouns. *Kuliko* ‘and’ indicates a resemblance relation and elicited a parallel antecedent preference, while *waynyahamyen* ‘because’ did not elicit a clear preference for either antecedent. This is as we predicted: due to our use of nonce verbs, a meaningful *explanation* relation could not be established. Second, the semantics of the additive particle *-to* ‘also’ played a role in determining coreference: presence of this particle elicits a parallel antecedent preference. This supports our claim that a *resemblance* relation is inferred to hold between the sentence with the additive particle and the preceding sentence, thereby satisfying the presupposition associated with the additive particle.

Moreover, when the connective *waynyahamyen* ‘because’ (explanation) and the additive particle *-to* ‘also’ (resemblance) co-occur (Also condition, Experiment 1), null pronouns are interpreted in accordance with the *resemblance* relation. This fits with the pattern we already saw with (7) where the parallel antecedent preference conflicts with our world knowledge but the pronoun tends to be coreferential with the parallel antecedent. This confirms Kehler’s (2002) claim that the process of establishing a resemblance relation is so strong that it overrides can override plausibility or real-world information (or lack thereof). Taken together, our results suggest that the semantic information conveyed by connectives and the additive marker are actively used to establish discourse-level coherence relations used to interpret null pronouns.

However, contrary to our expectation, topicality did not affect null pronoun resolution in either experiment. This fits with Ueno & Kehler (2016) whose work on Japanese subject position null pronouns (using transfer-of-possession verbs) found at best marginal effects of topicality.⁷ Thus, without discourse context, the topic marker seems insufficient to elicit topicality effects.

In light of our claims, one might wonder how our findings fit with the PAH. It could be that the subjecthood effects found for subject-position nulls by Carminati (2002) and others stem from an interaction of various coherence relations and verbal semantics used in the prior work. Indeed, research by Fedele (2006) and others has found that null pronouns can shift away from subject antecedents due to verb semantics. Alternatively, these findings may be related the typologically distinct properties of null pronouns in Romance languages vs. languages like Korean and Japanese. In Korean, unlike Romance, overt pronouns are rarely used. Null pronouns are the ‘unmarked’ option that is predominantly used, especially in conversation. Moreover, a previous

⁷ Ueno & Kehler (2016) used real verbs, not nonce verbs. Like us, they used sentences in isolation, without context. The fact that they also found no clear effects of topicality suggest our findings are not due to use of nonce verbs.

study on subject null pronouns in Korean (Kwon and Sturt 2013) argued that null pronoun resolution depends more on discourse cues than morphological cues. If the subject preference reflected in the PAH stems from language-specific traits, we would not find a clear division of labor effect between null and overt pronouns (i.e., subject-bias with null pronouns; object-bias with overt pronouns) even if we were to investigate various coherence relations with overt and null pronouns in Korean. Future work can provide a clearer answer to these questions.

References

- Alonso-Ovalle, Luis, Susana Fernández-Solera, Lyn Frazier & Charles Clifton. 2002. Null vs. overt pronouns and the topic-focus articulation in Spanish. *Italian Journal of Linguistics* 14. 151–170.
- Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Arnold, Jennifer E. 2001. The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes* 31. 137–162.
https://doi.org/10.1207/s15326950dp3102_02.
- Carminati, Maria N. 2002. *The Processing of Italian Subject Pronouns*. University of Massachusetts dissertation.
- Caramazza, Alfonso, Ellen Grober, Catherine Garvey & Jack Yates. 1977. Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behaviour* 16. 601–609.
[https://doi.org/10.1016/s0022-5371\(77\)80022-4](https://doi.org/10.1016/s0022-5371(77)80022-4).
- Chambers, Craig G. & Ron Smyth. 1998. Structural parallelism and discourse coherence: a test of centering theory. *Journal of Memory and Language* 39. 593–608.
<https://doi.org/10.1006/jmla.1998.2575>.
- Crawley, Rosalind A., Rosemary J. Stevenson & David Kleinman. 1990. The use of Heuristic Strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research* 19. 245–264. <https://doi.org/10.1007/bf01077259>.
- Fedele, Emily. 2016. *Discourse level processing and pronoun interpretation*. University of Southern California dissertation.
- Filiaci, Francesca. 2010. Null and overt subject biases in Spanish and Italian: A cross-linguistic comparison. In C. Borgonovo et al. (eds.), *Selected Proceedings for the 12th Hispanic Linguistics Symposium*. 171–182. Somerville, MA: Cascadilla Proceedings Project.
- Filiaci, Francesca, Antonella Sorace & Manuel Carreiras. 2013. Anaphoric biases of null and overt subjects in Italian and Spanish: A cross-linguistic comparison. *Language and Cognitive Processes* 29(7). 825–843.
- Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. In Charles. N. Li (ed.), *Subject and Topic*. 149–188. New York: Academic Press.
- Grosz, Barbara J., Scott Weinstein & Aravind K. Joshi. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21. 203–225.
<https://doi.org/10.21236/ada324949>.
- Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69. 274–307. <https://doi.org/10.2307/416535>.
- Hawkinson, Annie K., & Larry M. Hyman. 1974. Hierarchies of natural topic in Shona. *Studies in African linguistics* 5(2). 147.
- Hobbs, Jerry R. 1979. Coherence and coreference. *Cognitive Science* 3. 67–90.
https://doi.org/10.1207/s15516709cog0301_4.

- Kaiser, Elsi & John C. Trueswell. 2008. Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes* 23. 709–748. <https://doi.org/10.1080/01690960701771220>.
- Kaiser, Elsi. 2011. On the relation between coherence relations and anaphoric demonstratives in German. In I. Reich et al. (eds.), *Proceedings of Sinn und Bedeutung*. 337–351.
- Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford: CSLI Publications.
- Kehler, Andrew, Laura Kertz, Hannah Rohde & Jeffrey L. Elman. 2008. Coherence and coreference revisited. *Journal of Semantics (Special Issue on Processing Meaning)* 25. 1–44. <https://doi.org/10.1093/jos/ffm018>.
- Kertz, Laura, Andrew Kehler & Jeffrey Elman. 2006. Grammatical and coherence-based factors in pronoun interpretation. *Proceedings of the Annual Conference of the Cognitive Science Society* 28. 1605–1610.
- Kim, Kitaek, Theres Grüter & Amy J. Schafer. 2013. Effects of event-structure and topic/focus-marking on pronoun reference in Korean. Poster presented at the 26th Annual CUNY Conference on Human Sentence Processing. Columbia, SC.
- Kwon, Nayoung & Patrick Sturt. 2013. Null pronominal (pro) resolution in Korean, a discourse-oriented language. *Language and Cognitive Processes* 28.3. 377–387. <https://doi.org/10.1080/01690965.2011.645314>.
- Lee, EunHee, Sean Madigan & Mee-Jeong Park. 2015. *An introduction to Korean linguistics*. New York: Routledge. <https://doi.org/10.4324/9781315678016>.
- McKoon, Gail, Steven B. Greene & Roger Ratcliff. 1992. Discourse models, pronoun resolution, and the implicit causality of verbs. *Journal of Experimental Psychology* 19. 1040–1052. <https://doi.org/10.21236/ada259740>.
- Rohde, Hannah, Andrew Kehler & Jeffrey L. Elman. 2006. Event structure and discourse coherence biases in pronoun interpretation. *Proceedings of the Annual Meeting of the Cognitive Science Society* 28. 697–702.
- Rohde, Hannah & Andrew Kehler. 2008. The bidirectional influence between coherence establishment and pronoun interpretation. Poster presented at the 21st Annual CUNY Conference on Human Sentence Processing.
- Smyth, Ron. 1994. Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research* 23. 197–229. <https://doi.org/10.1007/bf02139085>.
- Stevenson, Rosemary J., Rosalind A. Crawley & David Kleinman. 1994. Thematic roles, focusing and the representation of events. *Language and Cognitive Processes* 9. 519–548. <https://doi.org/10.1080/01690969408402130>.
- Stevenson, Rosemary J., Alexander W. Nelson & Keith Stenning. 1995. The role of parallelism in strategies of pronoun comprehension. *Language and Speech* 38. 393–418. <https://doi.org/10.1177/002383099503800404>.
- Ueno, Mieko. and Andrew Kehler. 2016. Grammatical and pragmatic factors in the interpretation of Japanese null and overt pronouns. *Linguistics* 54(6). 1165–1221. <https://doi.org/10.1515/ling-2016-0027>.
- Winograd, Terry. 1972. Understanding natural language. *Cognitive psychology* 3(1). 1–191. [https://doi.org/10.1016/0010-0285\(72\)90002-3](https://doi.org/10.1016/0010-0285(72)90002-3).
- Wolf, Florian, Edward Gibson & Timothy Desmet. 2004. Discourse coherence and pronoun resolution. *Language and Cognitive Processes* 19(6). 665–675. <https://doi.org/10.1080/01690960444000034>.