Korean Journal of English Language and Linguistics, **Vol 26, April 2026, pp. 564-579**
DOI: 10.15738/kjell.26..202604.564

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Retrieving Antecedents in Noun Phrase Ellipsis and Reflexive Dependencies: New Evidence from a Maze Task Experiment

Nayoun Kim (Sungkyunkwan University)

Received: December 30, 2025
Revised: February 11, 2026
Accepted: March 5, 2026

Kim, Nayoun
Associate Professor
Sungkyunkwan University
25-2 Seonggyungwan-ro, Jongno-gu, Seoul 03063, Republic of Korea
Tel: +82-2-760-0270
Email: nayoun@skku.edu

* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2025S1A5A8008853).

ABSTRACT

Kim, Nayoun. 2026. Retrieving antecedents in noun phrase ellipsis and reflexive dependencies: New evidence from a Maze task experiment. *Korean Journal of English Language and Linguistics 26*, 564-579.

Previous research has shown that the processing of Noun Phrase Ellipsis (NPE) involves the reactivation of structural, semantic, and morphological information associated with the antecedent (Kim et al. 2019), and that reflexive processing adheres to Binding Principle A, requiring a structurally and semantically constrained antecedent (Chomsky 1981, Dillon et al. 2013, Sturt 2003). However, no prior work has examined antecedent retrieval in sentences that contain both ellipsis and reflexive dependencies. The present study tested a strong structural-recovery account, which predicts that sufficiently detailed structural, category-level, and morphological information from the NPE antecedent should guide subsequent reflexive resolution. Using the G-Maze paradigm (Boyce et al. 2020, Forster et al. 2009, Witzel and Forster 2014) and illusions of grammaticality (Wagers et al. 2009) as a diagnostic, we found, contrary to our predictions, that illusions emerged in both grammatical and ungrammatical sentences. This pattern is inconsistent with the predictions of the strong structural-recovery account and suggests that any reactivated structural information does not reliably constrain reflexive resolution in the predicted manner. At the same time, the findings do not isolate a single alternative mechanism and are compatible with multiple explanations, including reliance on linearly local cues, memory-based limitations, or task-related properties of the Maze paradigm. We therefore interpret these results as placing meaningful constraints on strong structural-recovery accounts and as motivating further research to adjudicate among competing explanations.

KEYWORDS

syntactic processing, ellipsis, a maze-task experiment, reflexive, the processing of dependencies, illusion of grammaticality, memory

1. Introduction and Background

1.1 The Processing of Noun Phrase Ellipsis and Reflexive Dependencies

Sentence comprehension unfolds in real time, requiring comprehenders to continuously integrate incoming words into an evolving interpretation. As each word is encountered, the parser incrementally constructs hierarchical structure, revises earlier commitments, and generates expectations about upcoming material based on available linguistic cues. Sentence processing is therefore fundamentally incremental. Reflexives and ellipsis, however, present particular challenges for the parser, since both reflexive dependencies and noun phrase ellipsis (NPE) require the retrieval of an antecedent that is syntactically and semantically licensed.

Comprehending reflexive and elliptical dependencies involves linking two elements, where one depends on the other for its grammatical status and interpretation (Chomsky 1981, Merchant 2001). NPE, in particular, exemplifies this type of dependency: one element relies on a licensed antecedent for its interpretation, and without such an antecedent, the intended meaning cannot be recovered.

(1) Tom's masterpiece was rejected, and Ella's was praised.

In (1), the noun *masterpiece* in the first conjunct is omitted after the possessive noun in the second conjunct. The interpretation and grammatical role of this missing noun can be established only by retrieving its antecedent in the first conjunct. In this construction, the antecedent (*masterpiece*) functions as the controlling element of the NPE dependency, since it determines both the interpretation and the grammatical properties of the ellipsis. Accordingly, we refer to the antecedent as the controlling element and the ellipsis site as the dependent element. We now turn to another type of dependency, namely reflexive dependency.

The reflexive *herself* in (2a) is licensed only if an NP is available to serve as an appropriate antecedent. A mismatch in morphological features between the antecedent and the reflexive renders the sentence ungrammatical, as shown in (2b). Thus, it is well established that the reflexive and its potential antecedent must match in number (and other relevant morphological features such as gender). Crucially, while a reflexive requires an antecedent for licensing and interpretation, the antecedent does not in turn depend on the reflexive. Accordingly, we treat the reflexive as the dependent element and the antecedent as the controlling element in the reflexive dependency.

(2) a. The businesswoman keeps herself busy.
b. *The businesswomen keep herself busy.

In a reflexive dependency, a reflexive depends on its antecedent at both structural and semantic levels. Structurally, the antecedent must occupy a position from which it c-commands the reflexive (Chomsky 1981, Reinhart 1976, Reinhart and Reuland 1993) and must occur within the reflexive's local binding domain such as NP or TP (see Büring 2005 for detailed reviews of cross-linguistic variation in what are considered binding domains). For example, in (2a), the reflexive (*herself*) is licensed by the NP (*the businesswoman*), which c-commands it (Chomsky 1981). Semantically, a reflexive's interpretation is determined entirely by the reference of its antecedent. Accordingly, a reflexive is licensed only under conditions in which both structural and semantic constraints are fulfilled.

As we have seen, both NPE and reflexive dependencies are constrained by their antecedents along structural and semantic dimensions. This observation naturally raises the question of how readers identify these dependencies

and rapidly reactivate the hierarchical structural, semantic, and morphological information associated with the antecedent during online processing (Kim et al. 2019, 2020). To address this question, we now turn to the time course of processing sentences (1) and (2a).

In (1), successful interpretation requires retrieving the nominal *masterpiece* at the ellipsis site in order to reconstruct the intended meaning of the possessive phrase in the second conjunct. Upon encountering the possessive morphology and, subsequently, the verb, the parser is compelled to infer that a syntactically required nominal has been omitted. Because the interpretation of NPE depends on an antecedent that satisfies both structural and semantic licensing conditions, the parser must initiate a memory search for an appropriate antecedent (Lewis and Vasishth 2005). Given that previously processed constituents are maintained in memory, the parser initiates a retrieval process to recover the antecedent (Frazier and Clifton 2001, Martin and McElree 2008, 2009, 2011). Once an appropriate antecedent is retrieved, the NPE dependency can be successfully established. Previous research indicates that antecedent retrieval in NPE resolution relies on the recovery of structural, categorial, and lexical (semantic and morphological) information associated with the antecedent (Ha et al. 2025, Kim et al. 2019, 2020).

We now turn to the processing of reflexives in (2a). The interpretation of a reflexive is constrained by both structural and featural requirements. On the one hand, grammatical principles require that a reflexive be structurally licensed by an antecedent that c-commands it, indicating that sentence comprehension mechanisms are sensitive to hierarchical structural relations (Cunnings and Sturt 2014, Dillon et al. 2013, Kush 2013). On the other hand, successful dependency formation also depends on compatibility in morphosyntactic and semantic features: only an NP whose gender, number, and animacy specifications align with those of the reflexive can serve as a suitable antecedent (Sturt 2003).

At the point where the reader encounters *herself* in (2a), the parser initiates a retrieval process to identify an appropriate antecedent. Since the antecedent precedes the reflexive, the parser must reactivate previously processed material in memory. Once a suitable antecedent is retrieved, the reflexive dependency can be successfully established (Cunnings and Sturt 2014, Dillon et al. 2013, Sturt 2003).

The question raised above concerns which types of information associated with the antecedent, including structural, semantic, and morphological information, are reactivated during the online processing of NPE constructions and reflexive dependencies. Previous work has shown that the processing of NPE involves the reactivation of structural, semantic, and morphological information associated with the antecedent (Kim et al. 2019). At the same time, the distribution of reflexives is governed by Binding Principle A (Chomsky 1981), suggesting that the parser should retrieve an antecedent that occupies a structurally licit position from which it can c-command the reflexive within its local binding domain (Chomsky 1981). Despite independent research on NPE and reflexive processing, no studies to date have examined NPE and reflexive dependencies occurring in the same sentence. Addressing this gap bears directly on questions concerning how linguistic representations are structured in memory, and how it is retrieved during reflexive processing (Kim et al. 2019, 2020, Martin and McElree 2008, 2009, 2011). It also addresses the extent to which the parser recovers elliptical and reflexive dependencies, both of which are semantically and syntactically constrained by their antecedents. Using a maze-task paradigm (Boyce et al. 2020, Forster et al. 2009, Witzel and Forster 2014), the present study tests a strong structural-recovery account by examining whether antecedent information, including structural, semantic, syntactic, and morphological information, constrains reflexive resolution in sentences containing both elliptical and reflexive dependencies.

1.2 Illusion of Grammaticality in NPE and Reflexive Dependency

To examine which aspects of the antecedent are recovered in NPE and subsequently used to resolve the reflexive

dependency, we use the illusion of grammaticality phenomenon as a probe (Kim et al. 2019, Lago et al. 2015, Tanner et al. 2014, Wagers et al. 2009). Let us first consider the time course of sentences involving NPE. In (3), the noun marked with possessive morphology (*Karen's*) and the subsequent modal signal that an important syntactic element is missing. How information from the left context is encoded in memory gives rise to at least three possible retrieval outcomes at the NPE site (Ha et al. 2025, Kim et al. 2019).

(3) Jenny's brother that worked with the sister will leave soon, and Karen's will certainly give himself some time to relax.

One theoretical perspective on memory encoding proposes that previously processed linguistic material is represented in memory along with its lexical (semantic and morphological) and categorial features. Under this view, stored representations retain content-based information that can later serve as retrieval cues (Lewis and Vasishth 2005). In reactivation, memory items matching the reflexive's retrieval cues are accessed in parallel (McElree 2000). This retrieval mechanism, known as *content-addressable memory* (Wagers et al. 2009), grants access to memory representations mediated by their featural content (Lewis and Vasishth 2005, McElree 2000, Van Dyke and Lewis 2003).

Within this framework, detection of NPE initiates the reactivation of previously encoded linguistic material. In (3), multiple noun phrases occur in the input prior to the ellipsis site. Because retrieval is cue-based, any NP whose lexical and categorial features match the retrieval cues generated at the ellipsis site can be activated as a potential antecedent. Consequently, both the nominal head *brother* governing the meaning of the whole NP (*brother that worked with the sister*) and the nominal head within the modifier PP *sister* may be retrieved at the ellipsis site, insofar as each provides an NP that satisfies the category-level cue for nominal material (Kim et al. 2019).

According to an alternative theory of memory encoding, previously processed linguistic material is stored in memory as structured representations (Kim et al. 2019, 2020), preserving not only lexical (semantic and morphological) and categorial features, but also hierarchical structural relations (Kush 2013, Lewis and Vasishth 2005). Under this account, the entire NP (*brother that worked with the sister*), which contains the embedded NP (*sister*), is retrieved as a structured unit and can serve as the antecedent of the reflexive. Moreover, because *brother* functions as the head of the noun phrase, it projects the category and core interpretation of the entire constituent. As a result, the head noun *brother* occupies a privileged structural position that is syntactically distinct from that of the embedded noun *sister*, which is contained within a modifier and does not project the properties of the larger phrase (Cunnings and Sturt 2018, Kim et al. 2019, Parker and An 2018).

Finally, in accordance with previous literature indicating that the parser can momentarily pursue a locally coherent parse that is later ruled out (Konieczny et al. 2010, Tabor et al. 2004), the linear proximity of *sister* to the ellipsis site may give rise to mis-retrieval. Although *sister* is an NP embedded within the PP, its closeness to the ellipsis site makes it possible for the parser to retrieve it as an antecedent. Thus, the parser may retrieve the linearly closest NP to the ellipsis, namely, *the sister* (see Kim et al. 2019 for a comprehensive account of the mechanisms underlying antecedent recovery in NPE).

Against this background, the present study investigates the mechanisms of antecedent retrieval during the processing of NPE constructions that are followed by a reflexive dependency. We draw on grammaticality illusions (Lago et al. 2015, Wagers et al. 2009) as a diagnostic tool to determine what types of information encoded in the antecedent are reactivated (Kim et al. 2019) during the processing of an elliptical dependency followed by a reflexive dependency.

(4) a. Jenny's brother that worked with the woman will leave soon, and Karen's will certainly give himself some time to relax.
b. Jenny's brother that worked with the man will leave soon, and Karen's will certainly give himself some time to relax.
c. Jenny's sister that worked with the woman will leave soon, and Karen's will certainly give himself some time to relax.
d. Jenny's sister that worked with the man will leave soon, and Karen's will certainly give himself some time to relax.

Under a retrieval mechanism based exclusively on lexical and categorial features, both *brother*/*sister that worked with the woman*/*man* and *woman*/*man* may be retrieved. More precisely, the head nouns *brother*/*sister* and *woman*/*man* constitute the likely retrieval targets. Under this theory, any NP bearing the appropriate nominal category (NP) and matching the relevant morphological features (e.g., gender and number) is eligible for retrieval. After the antecedent has been retrieved in the NPE, the parser evaluates candidate antecedents during reflexive processing with respect to the cues provided by the reflexive, including lexical and categorial features such as [+Noun], [+Singular], and [+Feminine]. Under this scenario, both *brother*/*sister* and *woman*/*man*, insofar as they satisfy these cues, may be reactivated in memory. Crucially, because retrieval is cue based and parallel, neither NP is predicted to be prioritized over the other at this stage of processing.

Consequently, memory representations matching only a subset of the retrieval cues may compete for retrieval, thereby giving rise to interference effects when an ultimately incorrect item is retrieved. When multiple candidates share overlapping features, such as gender as in (4c) and (4d), they are jointly activated in memory, and this simultaneous activation induces similarity-based interference effects during retrieval (Jäger et al. 2017). If both *brother/sister* and *man/woman* match the reflexive in gender, the gender cue distributes activation across both antecedents, making it more difficult for the parser to identify the structurally appropriate antecedent. As a result, we expect not only facilitatory interference effect, such that (4d) is processed faster than (4c) under ungrammatical conditions, but also inhibitory interference effect (Jäger et al. 2017, Lewis and Vasishth 2005), whereby (4b) is processed more slowly than (4a) under grammatical conditions. To summarize, once the antecedent has been retrieved for NPE, and the reflexive dependency must subsequently be resolved, we predict greater similarity-based interference when multiple candidates (e.g., *the brother*, *the man*) share the same gender feature as the reflexive (*himself*) in grammatical conditions (Jäger et al. 2017, Lewis and Vasishth 2005, Van Dyke 2007, Van Dyke and McElree 2006), along with stronger facilitatory interference in ungrammatical conditions.

Second, when retrieval is based on lexical, categorial, and structural information represented in memory (Kim et al. 2019), the NP *brother that worked with the woman*, including the embedded *woman*, may be retrieved, with emphasis placed on the head noun of the entire constituent, namely *brother*. When reflexive binding fails due to a gender mismatch between the head noun and the reflexive, the parser may instead retrieve the embedded noun in the modifier, as cue-based retrieval mechanisms attempt to resolve the mismatch by accessing alternative elements that partially satisfy the reflexive's retrieval cues (Kim et al. 2019, Lago et al. 2015, Tanner et al. 2014, Wagers et al. 2009). Thus, we expect a grammatical illusion in (4d) relative to (4c). By contrast, no difference is predicted for grammatical sentences, since successful gender matching in grammatical conditions precludes the need for the parser to initiate repair (Lago et al. 2015, Wagers et al. 2009).

Lastly, if the parser is attracted to a locally coherent parse that is ultimately incompatible with the overall structure (Gibson et al. 1996, Tabor et al. 2004), only the NP segment immediately adjacent to the reflexive (e.g., *man*, *woman*) will be retrieved. Consequently, when the retrieved NP is associated with the reflexive, reflexive

binding succeeds provided that the NP adjacent to the reflexive matches it in gender. This effect is expected to arise in both grammatical and ungrammatical sentences. That is, when the reflexive dependency is processed subsequent to the resolution of the NPE dependency, the locally matching noun is predicted to benefit from a processing advantage in both grammatical and ungrammatical conditions, with (4b) expected to be faster than (4a) and (4d) faster than (4c).

As illustrated above, we examine real-time processing in cases where two distinct dependencies are both structurally and semantically linked to a single antecedent. In such configurations, the ellipsis must be resolved prior to the establishment of the reflexive dependency. We investigate what types of information can be recovered when both dependencies must be successfully resolved, using a novel maze task experiment.

2. The Current Experiment

In the following experiment, we investigate the antecedent retrieval processes for NPE and subsequent reflexive dependencies, both of which crucially require an antecedent for interpretation.

2.1 Participants

A total of 58 native speakers of English (aged 20–50) were recruited via Prolific. All participants gave their consent to participate and received $4 as compensation. Once participants entered the experiment, they were required to complete it. The experiment lasted approximately 15–20 minutes.

2.2 Materials

The materials followed a 2 × 2 factorial design in which the *Gender* of the local noun and *Grammaticality* were manipulated as independent variables. The *Gender* factor manipulated whether the noun in the first conjunct linearly adjacent to the reflexive matched or mismatched the reflexive in gender. The *Grammaticality* factor manipulated whether sentences were grammatical or ungrammatical, as determined by gender compatibility between the nominal head and the reflexive. The experimental materials consisted of 24 items. Following a Latin square design, each participant read one condition from each item, along with 48 filler sentences that contained neither ellipsis nor reflexives. Participants also completed a preliminary three-item practice session to become familiar with the task. Table 1 presents a sample set of stimuli used in the current maze task experiment.

Table 1. **A Sample Set of Stimuli Used in the Experiment**

Condition	Example sentence
A. Grammatical, Match	Christopher's sister that studied with the girl will leave soon, and John's will probably dress herself for the party.
B. Grammatical, Mismatch	Christopher's sister that studied with the boy will leave soon, and John's will probably dress herself for the party.
C. Ungrammatical, Match	Christopher's brother that studied with the girl will leave soon, and John's will probably dress herself for the party.
D. Ungrammatical, Mismatch	Christopher's brother that studied with the boy will leave soon, and John's will probably dress herself for the party.

2.3 Procedure

We employed a grammatical maze (G-Maze) task (Boyce et al. 2020, Forster et al. 2009, Witzel and Forster 2014) instead of a standard self-paced reading paradigm. As in self-paced reading, the G-Maze task provides word-by-word reading-time measures; however, it additionally requires participants to make explicit grammaticality decisions at each step of the sentence (Boyce et al. 2020). Each trial presented participants with two alternative continuations: one grammatically licit continuation and one ungrammatical distractor (referred to as a foil). Participants chose the option that fit into the unfolding sentence. Trials in which an ungrammatical option was selected were immediately terminated and excluded from further analyses.

The maze paradigm affords several methodological advantages for studying real-time sentence processing. Previous research has shown that maze tasks yield highly localized measures of processing difficulty, in contrast to self-paced reading, with effects often manifesting in spillover regions rather than at the critical word itself (Boyce et al. 2020, Forster et al. 2009, Witzel and Forster 2014). In addition, the task structure enforces sustained attention throughout each trial: since incorrect responses trigger immediate trial termination, participants cannot afford to disengage at any point during the trial (Forster et al. 2009). Maze tasks consistently produce low trial-failure rates while preserving robust sensitivity to experimental manipulations, rendering them particularly well suited for web-based psycholinguistic research (Boyce et al. 2020, Forster et al. 2009, Witzel and Forster 2014), although some variation in distractor difficulty in the G-maze task may warrant further consideration.

The experiment was administered using PC-Ibex (Zehr and Schwarz 2018). At each step of the task, participants were presented with two candidate sentence continuations, one grammatical continuation and one ungrammatical foil. Note that the ungrammatical foils are not pseudo-words; rather, they are composed of real lexical items that are neither semantically nor grammatically compatible with the unfolding sentence. This property of the foils requires participants to go beyond simply recognizing words and instead to fully integrate each word with the immediate context and the already established parse.

As illustrated in (5), in the G-Maze paradigm, foils must be structurally plausible at each step and composed entirely of grammatical forms. For example, participants first choose between *Christopher's* and a distractor option (e.g., x-x-x). If they correctly select *Christopher's*, they then proceed to choose between *sister* and *find*. Upon correctly selecting *sister*, the next decision involves choosing between *that* and *nor*. If they correctly select *that*, they then proceed to choose between *studied* and *ago*. This process continues until the end of the sentence, at which point a new trial begins. Participants were instructed to choose the continuation that could be integrated into the sentence in progress, responding as accurately as possible by pressing either the "e" or "i" key. Figure 1 presents screenshots from the G-Maze experiment used in the present study.

(5) Target sentence example employed in the current experiment
a. Christopher's sister that studied with the girl will leave soon, and John's will probably dress herself for the party.

Foils
b. x-x-x find nor ago miss else nor which miles cent, anti than etc congress woods receive fund died whose.

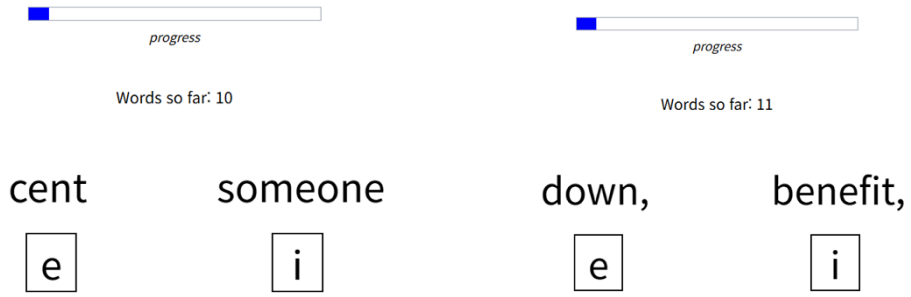

Figure 1. Screenshots from the Present G-Maze Experiment

2.4 Predictions

Our predictions are based on the sample set of stimuli illustrated in Table 1. As previewed above, under accounts in which retrieval is guided primarily by morphological, semantic, and category-level features, the parser is expected to retrieve multiple potential antecedents that satisfy the requirements of the ellipsis. In the present configuration, both the head of the NP (e.g., *sister* in *sister that studied with the girl/boy*) and the embedded NP contained in the modifier PP (*girl*) share the same category (NP), with no principled priority assigned to either candidate. Thus, when the reflexive dependency remains to be resolved, the reader is expected to select the noun that matches the reflexive in gender rather than the gender-mismatching noun, particularly when the head noun (*brother*) does not match the reflexive. As a result, we predict faster processing in Condition C compared to Condition D.

At the same time, when multiple candidates share the same gender features, they compete for activation during retrieval. This competition gives rise to similarity-based interference effect (Jäger et al. 2017, Lewis and Vasishth 2005, Van Dyke 2007, Van Dyke and McElree 2006), resulting in slower processing in conditions with gender overlap. Accordingly, we predict that Condition A will be processed more slowly than Condition B at the reflexive.

Under a strong structural-recovery account, in which antecedent retrieval is guided by a fully specified structural representation along with associated morphological and semantic features, condition C is expected to be processed faster than condition D at the reflexive, with no difference predicted between conditions A and B. This is because the head noun of the NP *brother that studied with the girl*, namely *brother*, is initially considered, and the noun *girl* is considered as a potential antecedent only when reflexive binding with the nominal head of the entire NP (*brother*) and the reflexive (*herself*) fails (Kim et al. 2019, Lago et al. 2015, Wagers et al. 2009). At this stage, the parser searches for a morphologically matching noun as a fallback strategy to resolve the dependency.

Finally, if only the noun phrase adjacent to the ellipsis site or reflexive is retrieved (Konieczny et al. 2010, Tabor et al. 2004), the parser should then reactivate the noun in the first conjunct that is closest to the reflexive. In this case, a local noun whose gender matches that of the reflexive would be treated as a potential antecedent, leading to facilitation when the local noun and reflexive match in both grammatical and ungrammatical conditions. Accordingly, Condition A should be read faster than Condition B, and Condition C should be read faster than Condition D at the reflexive region.

2.5 Analysis

We analyzed the data using linear mixed-effects regression implemented in the *lme4* package in R (Bates et al. 2015). The models were specified with a maximal random-effects structure, involving crossed random intercepts and slopes for participants and items (Baayen et al. 2008, Bates et al. 2015). When a model did not reach convergence, the random structure was progressively simplified until convergence was reached. Reading times were log-transformed to reduce skew.

2.6 Results

In this section, we report log reading times at the critical region and the critical spillover region. These two regions are particularly relevant to our study, as the G-Maze task is specifically designed to capture fine-grained localization effects (Boyce et al. 2020, Forster et al. 2009, Witzel and Forster 2014). The log reading times at the reflexive are presented in Figure 2, and the log reading times at the critical spillover region (the region following the reflexive) are presented in Figure 3. Results of the linear mixed-effects models for log reading times at the critical region and the spillover region are presented in Table 2.

Table 2. Results of Linear Mixed-effects Models for Log Reading Times at the Critical Region and Spillover Region

	Estimate	SE	*t*-value	*p*-value
Critical Region (*himself/herself*)				
(Intercept)	6.950	0.038	184.637	
Grammaticality	0.056	0.023	2.409	0.016 *
Gender	–0.063	0.023	–2.728	0.006 **
Grammaticality × Gender	0.075	0.046	1.626	0.105
Critical Region Spillover				
(Intercept)	6.811	0.043	157.746	
Grammaticality	0.026	0.020	1.305	0.192
Gender	–0.017	0.020	–0.876	0.381
Grammaticality × Gender	–0.023	0.040	–0.571	0.568

Note. ** $p < .01$, * $p < .05$.

At the critical region (*himself*/*herself*), we observed a main effect of *Grammaticality* ($\beta = 0.056$, $SE = 0.023$, $t = 2.409$, $p < 0.05$), such that grammatical sentences were read faster than ungrammatical sentences. We also found a main effect of *Gender* ($\beta = –0.063$, $SE = 0.023$, $t = –2.728$, $p < 0.001$): sentences in which the local noun matched the reflexive were read faster than those in which it did not. However, there was no interaction between *Grammaticality* and *Gender* ($\beta = 0.075$, $SE = 0.046$, $t = 1.626$, $p = 0.105$), indicating that the gender-matching effect did not selectively enhance only the ungrammatical conditions. Instead, the effect was observed uniformly across conditions.

At the critical spillover region, we observed no main effect of *Grammaticality* ($\beta = 0.026$, $SE = 0.020$, $t = 1.305$, $p > 0.05$) or *Gender* ($\beta = –0.017$, $SE = 0.020$, $t = –0.876$, $p > 0.05$), nor any interaction between the two ($\beta = –0.023$, $SE = 0.040$, $t = –0.571$, $p > 0.05$). This pattern can be attributed to a defining property of the maze task, namely that it is designed for fine-grained localization of processing difficulty, with effects emerging directly at the critical region (Boyce et al. 2020, Forster et al. 2009, Witzel and Forster 2014).

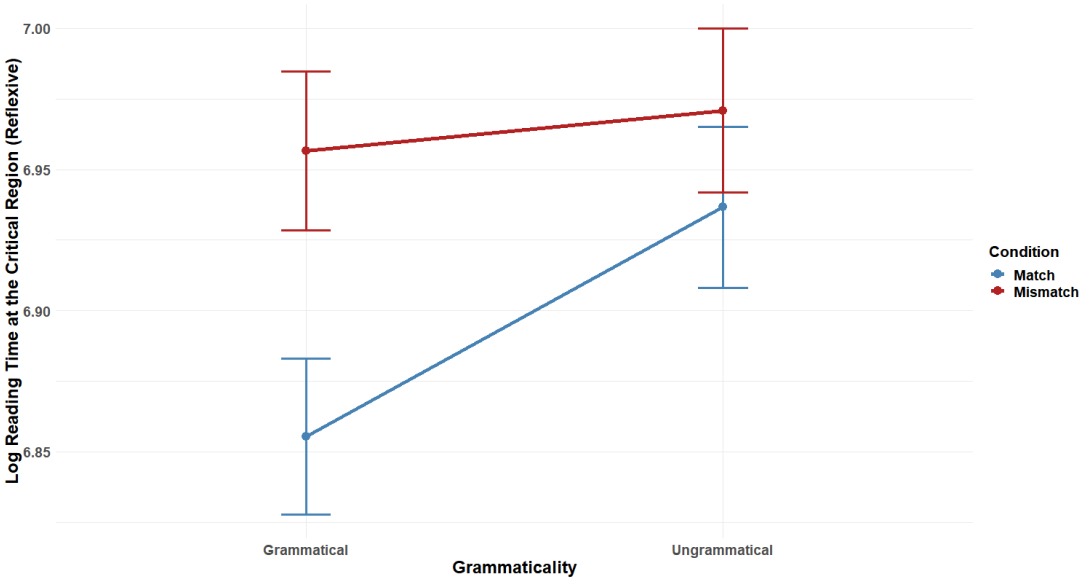

Figure 2. Log Mean Reading Times at the Critical Region (e.g., *himself/herself*)

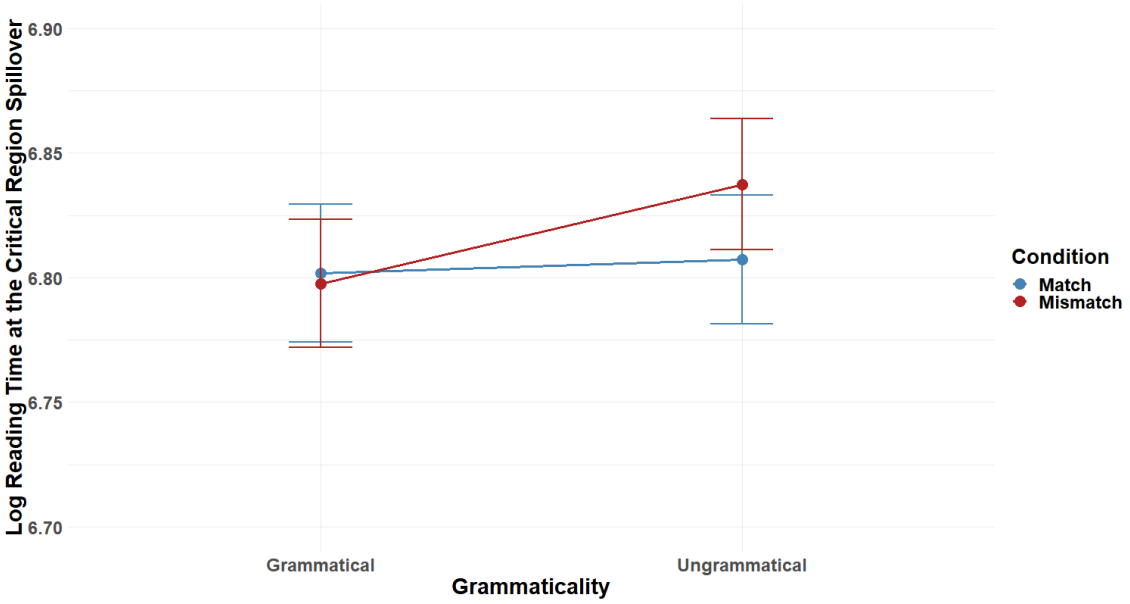

Figure 3. Log Mean Reading Times at the Critical Region Spillover

3. General Discussion

It is well established that the processing of NPE involves the reactivation of structural information associated with the antecedent (Ha et al. 2025, Kim et al. 2019, Merchant 2001, 2013). Likewise, the processing of a reflexive requires adherence to Binding Principle A, which stipulates that the antecedent must c-command the reflexive within its local binding domain (Chomsky 1981). Thus, the parser must identify an antecedent in a structurally

appropriate position from which it can c-command the reflexive (Cunnings and Sturt 2014, Dillon et al. 2013, Sturt 2003). Because both the reflexive and the ellipsis dependency are semantically and syntactically controlled by the antecedent, understanding these dependencies requires recovering the semantic, syntactic, morphological, and structural features linked to the antecedent. In this paper, we investigate what types of linguistic information are retrieved at the reflexive when two dependencies are present, both of which are semantically and structurally constrained by the same antecedent. To address this question, we embed NPE and reflexive dependencies within the same sentence.

We adopted the G-Maze paradigm, which serves as a complementary method to self-paced reading for investigating dependency processing. Previous studies indicate that G-Maze involves word identification, contextual integration, and monitoring the completion of integration (Boyce et al. 2020, Orth et al. 2025). At the same time, G-Maze strongly promotes incremental processing by requiring participants to choose between two alternatives, and it has been shown to yield larger effect sizes and more precise temporal localization compared with other methodologies (Boyce et al. 2020, Forster et al. 2009, Witzel and Forster 2014). This design promotes structural prediction and reveals commitment points in real time, making G-Maze useful for probing online comprehension.

To reiterate, our findings revealed main effects of both *Grammaticality* and *Gender* at the reflexive region, with no interaction between them. Specifically, ungrammatical sentences were read more slowly than their grammatical counterparts, and gender-mismatch sentences were read more slowly than gender-match sentences. There was no evidence of an illusion of grammaticality in the ungrammatical constructions (Wagers et al. 2009), as the gender-match effect emerged in both grammatical and ungrammatical contexts.

These findings are not compatible with an account in which only semantic, morphological, and category-level information is reactivated independently of the structural information of the antecedent. If such a retrieval mechanism were operative, we would expect a similarity-based interference effect (Jäger et al. 2017, Lewis and Vasishth 2005, Van Dyke 2007, Van Dyke and McElree 2006) in grammatical conditions, since additional elements occurring between the recovered antecedent and the reflexive should introduce inhibitory interference rather than the facilitatory interference effect driven by the locally adjacent noun to the reflexive. These results are likewise inconsistent with strong accounts that posit that fully specified structural, semantic, morphological, and category-level information is reactivated and reliably constrains reflexive resolution. Under that account, we would expect facilitatory interference only in ungrammatical conditions, not in grammatical conditions (Lago et al. 2015, Wagers et al. 2009).

One might be tempted to conclude that the absence of an illusion of grammaticality arises because the local noun adjacent to both the ellipsis and the reflexive is retrieved irrespective of grammaticality, in contrast to previous findings (see Kim et al. 2019, 2020). On this view, readers are predicted to experience robust interference from local elements across conditions, because they tend to select the closest or most linearly adjacent element in the first conjunct to the ellipsis or reflexive (Konieczny et al. 2010, Tabor et al. 2004). Related proposals further argue that activation of a nearby antecedent can remain sufficiently strong to induce temporary confusion and lead readers to select a linearly closer but grammatically inappropriate antecedent (Ferreira and Patson 2007).

The present findings may also be interpreted within the unrestricted race framework of sentence processing (Van Gompel et al. 2000, Logačev and Vasishth 2016), which predicts an ambiguity advantage whereby multiple parses can be pursued in parallel, often yielding greater processing ease than when a single analysis is forced (Swets et al. 2008). Because unambiguous attachments may require reanalysis whereas ambiguous structures avoid such revision costs (Van Gompel et al. 2000, Logačev and Vasishth 2016), the G-Maze task's emphasis on strictly

local, word-by-word interpretation may promote reliance on locally coherent yet ambiguous nouns, consistent with this model's predictions.

Although interference from local elements is therefore one plausible interpretation of the present findings, the observed pattern is also compatible with several alternative explanations. These include accounts invoking the demands of resolving multiple dependencies and those attributing the effects to task-driven sensitivity to local lexical compatibility induced by the Maze paradigm.

From the perspective of dependency resolution, the absence of an illusion of grammaticality may reflect interactions between the two dependencies examined in this study. Both ellipsis and reflexive dependencies require the retrieval of an antecedent from memory, with the ellipsis dependency being resolved prior to the reflexive dependency. In our materials, as illustrated in (6), the ellipsis site is signaled by a noun bearing possessive marking, followed by a modal, prompting the reader to retrieve an antecedent such as *sister that studied with the girl* (specifically the head noun *sister*) or, alternatively, *girl*.

(6) Christopher's sister that studied with the girl will leave soon, and John's will probably dress herself for the party.

Once the ellipsis dependency has been resolved, the reflexive dependency must subsequently be processed when the reflexive is encountered. At that point, the antecedent previously retrieved for the ellipsis must be reactivated from memory in order to satisfy the interpretive requirements of the reflexive. Because antecedent recovery is triggered only when the reflexive becomes overtly signaled, the parser does not continuously maintain antecedent information over an extended span of processing (Gibson 1998, 2000). This timing difference may account for the divergence between our results and previous findings on NPE resolution in sentences lacking a reflexive dependency.

The crucial difference between that work and the present study concerns the predictability of the dependent element. When only NPE is present in a sentence, as in Kim et al. (2019), the recovered NP allows readers to anticipate the upcoming verb together with its associated morphological properties. By contrast, in the current experiment, although structural information is recovered at the NPE site, the presence of a reflexive is not predictable in advance. Consequently, when the verb is encountered, the parser may no longer maintain or prioritize the retrieved structural information associated with the antecedent, rendering it more susceptible to decay or to processing demands associated with resolving an additional dependency once the reflexive appears (see Kim et al. 2020 and Wagers and Phillips 2014 for discussion of how distance effects and syntactic configuration can modulate illusions of grammaticality).

On a similar note, in Kim et al. (2019), the sentences involved NPE but did not include a reflexive. In that design, once the antecedent had been recovered, the sentence strongly supported prediction of the upcoming verb in the second conjunct (e.g., *Molly's key to the cabinets must be in the kitchen and Haley's probably are on the carpet*). Under a top-down predictive processing account, the head noun *key* predicts the verb and its morphological realization (e.g., singular agreement), and the comprehender subsequently engages a repair strategy to resolve number mismatch between the antecedent and the verb. By contrast, in our current experiment, the verb does not carry number or gender morphology, and therefore the representation of the recovered antecedent may not be sufficiently strengthened before the reflexive appears. As a result, the relevant features of the subject may need to be retrieved only when the reflexive is encountered (see Keshev and Meltzer-Asscher 2024 for recent evidence that subject representations may be updated based on verbal agreement, thereby attenuating attraction effects at later processing sites).

Another explanation for the unexpected results is that the lack of an illusion of grammaticality may partly stem from the experimental methodology. In the G-Maze task, readers make a grammaticality decision at every step, enforcing strictly incremental processing and enabling precise localization of processing difficulty, with effects detected exactly where they arise rather than spilling over into later regions (Boyce et al. 2020, Orth et al. 2025). Consistent with this, we observed a robust main effect of grammaticality at the critical region in the current experiment.

However, a potential limitation of this methodology, which is also methodologically very interesting to pursue, is its particular sensitivity to local influences, which may cause the observed pattern to reflect task-driven sensitivity to local lexical compatibility. Participants may have adopted strategies that favor highly predictable continuations and locally coherent interpretations. From this perspective, the absence of an illusion of grammaticality could arise from such task-induced strategies. At the same time, to our knowledge, the present study is unique in employing a fine-grained Maze-based design to examine illusions of grammaticality in the context of multiple simultaneously active dependencies, which have previously been shown to yield robust effects for each dependency independently (Cunnings and Sturt 2014, Dillon et al. 2013, Kim et al. 2019, Sturt 2003) but have not been tested simultaneously.

We remain cautious in interpreting the present findings as favoring any particular account of antecedent recovery for the multiple dependencies tested in this study. The extent to which decision-making differences across methodologies influence real-time processing remains unresolved and will require systematic cross-methodological comparisons and revised theoretical interpretations as new data emerge (Boyce et al. 2020). Future research employing other online methodologies, including an eye-tracking while reading experiment, will help clarify whether interference from local coherence can override the recovery of structural, categorial, and lexical information during the resolution of multiple dependencies (Kim et al. 2020, Wagers and Phillips 2014), or whether such effects instead arise from properties of the experimental task. We leave this question open for future investigation.

4. Conclusion

Previous studies have shown that the processing of NPE involves the reactivation of structural information associated with the antecedent (Kim et al. 2019, Merchant 2001, 2013). Reflexive interpretation, in turn, must conform to Binding Principle A, requiring the antecedent to c-command the reflexive within the same local binding domain (Chomsky 1981). A strong structural-recovery account therefore predicts that sufficiently detailed structural information from the NPE antecedent should guide subsequent reflexive resolution (Cunnings and Sturt 2014, Dillon et al. 2013, Sturt 2003). However, no prior work has examined antecedent recovery in sentences containing both elliptical and reflexive dependencies. The present study addressed this gap by investigating reflexive resolution following NPE using an adapted version of the G-Maze paradigm (Forster et al. 2009, Witzel and Forster 2014). Using illusions of grammaticality as a diagnostic, we tested the prediction of a strong structural-recovery account, according to which sufficiently detailed structural, semantic, category-level, and morphological information from the antecedent should yield selective illusions of grammaticality at the reflexive (Kim et al. 2019, Lago et al. 2015, Wagers et al. 2009). Contrary to our predictions, illusions emerged in both grammatical and ungrammatical sentences. This pattern is not fully consistent with the predictions of the strong structural-recovery hypothesis as formulated here. Although structural information may be reactivated during NPE processing, the present findings suggest that such information does not reliably guide reflexive resolution in the predicted manner.

The results do not point to a single mechanism for resolving multiple dependencies. Instead, the observed pattern is compatible with several explanations, including reliance on linearly local cues at the reflexive, partial maintenance or decay of the antecedent representation, increased integration demands associated with multiple dependencies, and task-related properties of the Maze paradigm. Taken together, these findings constrain strong structural-recovery accounts while remaining compatible with multiple retrieval mechanisms. Future research with more direct manipulations of structural accessibility will be necessary to adjudicate among these competing explanations.

References

Baayen, R. H., D. J. Davidson and D. M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390-412.

Bates, D., M. Mächler, B. Bolker and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1-48.

Boyce, V., R. Futrell and R. P. Levy. 2020. Maze made easy: Better and easier measurement of incremental processing difficulty. *Journal of Memory and Language* 111, 104082.

Büring, D. 2005. *Binding Theory*. Cambridge, UK: Cambridge University Press.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht, Netherlands: Foris Publications.

Cunnings, I. and P. Sturt. 2014. Coargumenthood and the processing of reflexives. *Journal of Memory and Language* 75, 117-139.

Cunnings, I. and P. Sturt. 2018. Retrieval interference and semantic interpretation. *Journal of Memory and Language* 102, 16-27.

Dillon, B., A. Mishler, S. Sloggett and C. Phillips. 2013. Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language* 69(2), 85-103.

Ferreira, F. and N. D. Patson. 2007. The 'good enough' approach to language comprehension. *Language and Linguistics Compass* 1(1-2), 71-83.

Forster, K., C. Guerrera and L. Elliot. 2009. The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods* 41(1), 163-171.

Frazier, L. and C. Clifton, Jr. 2001. Parsing coordinates and ellipsis: Copy α. *Syntax* 4(1), 1-22.

Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68, 1-76.

Gibson, E. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita and W. O'Neil, eds., *Image, Language, Brain*, 95-126. Cambridge, MA: MIT Press.

Gibson, E., N. Pearlmutter, E. Canseco-Gonzalez and G. Hickok. 1996. Recency preference in the human sentence processing mechanism. *Cognition* 59(1), 23-59.

Ha, Y., M. Frazier and N. Kim. 2025. Grammaticality illusions in the real-time processing of backward noun phrase ellipsis. *Korean Journal of English Language and Linguistics* 25, 277-288.

Jäger, L. A., F. Engelmann and S. Vasishth. 2017. Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language* 94, 316-339.

Keshev, M. and A. Meltzer-Asscher. 2024. The representation of agreement features in memory is updated during sentence processing: Evidence from verb-reflexive interactions. *Journal of Memory and Language* 135, 104495.

Kim, N., L. Brehm and M. Yoshida. 2019. The online processing of noun phrase ellipsis and mechanisms of antecedent retrieval. *Language, Cognition and Neuroscience* 34(2), 190-213.

Kim, N., L. Brehm, P. Sturt and M. Yoshida. 2020. How long can you hold the filler: Maintenance and retrieval. *Language, Cognition and Neuroscience* 35(1), 17-42.

Konieczny, L., H. Weldle, S. Wolfer, D. Müller and P. Baumann. 2010. Anaphora and local coherences. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, 1204-1209.

Kush, D. W. 2013. *Respecting Relations: Memory Access and Antecedent Retrieval in Incremental Sentence Processing*. Doctoral dissertation, University of Maryland, College Park, MD, USA.

Lago, S., D. E. Shalom, M. Sigman, E. F. Lau and C. Phillips. 2015. Agreement attraction in Spanish comprehension. *Journal of Memory and Language* 82, 133-149.

Lewis, R. L. and S. Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29(3), 375-419.

Logačev, P. and S. Vasishth. 2016. A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension. *Cognitive Science* 40(2), 266-298.

Martin, A. E. and B. McElree. 2008. A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language* 58(3), 879-906.

Martin, A. E. and B. McElree. 2009. Memory operations that support language comprehension: Evidence from verb-phrase ellipsis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35(5), 1231-1239.

Martin, A. E. and B. McElree. 2011. Direct-access retrieval during sentence comprehension: Evidence from Sluicing. *Journal of Memory and Language* 64(4), 327-343.

McElree, B. 2000. Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research* 29, 111-123.

Merchant, J. 2001. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford, UK: Oxford University Press.

Merchant, J. 2013. Diagnosing ellipsis. *Diagnosing Syntax* 1, 537-542.

Orth, W., D. M. Nemeskey and E. Ronai. (2025). Extending the maze task to Hungarian: New insights in relative clause processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication.

Parker, D. and A. An. 2018. Not all phrases are equally attractive: Experimental evidence for selective agreement attraction effects. *Frontiers in Psychology* 9, 1566.

Reinhart, T. M. 1976. *The Syntactic Domain of Anaphora*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA.

Reinhart, T. and E. Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24(4), 657-720.

Sturt, P. 2003. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language* 48(3), 542-562.

Swets, B., T. Desmet, C. Clifton and F. Ferreira. 2008. Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition* 36(1), 201-216.

Tabor, W., B. Galantucci and D. Richardson. 2004. Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language* 50(4), 355-370.

Tanner, D., J. Nicol and L. Brehm. 2014. The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language* 76, 195-215.

Van Dyke, J. A. 2007. Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(2), 407-430.

Van Dyke, J. A. and R. L. Lewis. 2003. Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language* 49(3), 285-316.

Van Dyke, J. A. and B. McElree. 2006. Retrieval interference in sentence comprehension. *Journal of Memory and Language* 55, 157-166.

Van Gompel, R. P., M. J. Pickering and M. J. Traxler. 2000. Unrestricted race: A new model of syntactic ambiguity resolution. In A. Kennedy, R. Radach, D. Heller and J. Pynte, eds., *Reading as a Perceptual Process*, 621-648. Amsterdam, Netherlands: North-Holland.

Wagers, M. W., E. F. Lau and C. Phillips. 2009. Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language* 61, 206-237.

Wagers, M. W. and C. Phillips. 2014. Going the distance: Memory and control processes in active dependency construction. *Quarterly Journal of Experimental Psychology* 67(7), 1274-1304.

Witzel, J. and K. Forster. 2014. Lexical co-occurrence and ambiguity resolution. *Language, Cognition and Neuroscience* 29(2), 158-185.

Zehr, J. and F. Schwarz. 2018. PennController for Internet Based Experiments (IBEX).

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary

