J. Linguistics **58** (2022), 73–109. © The Author(s), 2021.
Published by Cambridge University Press
doi:10.1017/S0022226720000420

Fragments and structural identity on a direct interpretation approach

JOANNA NYKIEL
Kyung Hee University

JONG-BOK KIM
Kyung Hee University

(Received 16 August 2018; revised 23 October 2020)

This paper examines the relationship between merger and sprouting fragments, which are typically taken to involve clausal ellipsis. We argue that structural identity constraints on fragments and their correlates should, where appropriate, make reference to the argument structure of lexical heads in the antecedent clauses. Our proposal is spelled out as part of a direct interpretation approach to clausal ellipsis, but, in addition, it incorporates processing-based preferences as a means to motivate the contrast between merger and sprouting fragments. We propose specifically that phrases which are available to serve as correlates for fragments are maximal categories derived from the argument structure of lexical heads in the antecedents. This proposal successfully predicts form-matching effects that surface under clausal ellipsis, as well as well-known limits on clausal ellipsis regarding the morphosyntactic form of fragments. We take advantage of the fact that fragments are not embedded in unpronounced structures, which allows us to articulate a proposal that avoids the difficulty of having to simultaneously relate a fragment to the structure of the antecedent and to its own unpronounced structure, a difficulty that current PF-deletion accounts face.

KEYWORDS: clausal ellipsis, construction-based HPSG, direct interpretation, form minimization, fragments, merger, sprouting

1. INTRODUCTION

This paper examines clausal ellipsis and, in particular, the contrast between two kinds of fragments, the merger kind and the sprouting kind.[1] We illustrate merger (1) and sprouting (2) fragments for Korean first (see also J.-B. Kim 2015a,b).[2]

[1] We are grateful to three anonymous referees for their constructive feedback and helpful suggestions. We would also like to thank Anne Abeillé for comments on a previous draft of this paper and Jean-Pierre Koenig for discussion of some of the ideas included in this paper. We also thank Jungsoo Kim and Rok Sim for comments and suggestions on earlier drafts. All remaining errors are ours. This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2017S1A2A2041092).

[2] The glosses we use in this paper include ACC (accusative), COP (copula), DAT (dative) DECL (declarative), DGB (dialogue game board), GCASE (grammatical case), GC (general case), INST (instrument), NEG (negation), NOM (nominative), NON.S (non-subject case), PST (past), QUE

(1) A: Mimi-ka nwukwu-lul manna-ss-e?
Mimi-NOM who-ACC meet-PST-QUE
'Who did Mimi meet?'
B: Haha-lul. 'Haha-ACC'
B′: Haha. 'Haha'

(2) A: phyenci-ka wa-ss-e?
letter-NOM come-PST-QUE
'Did the letter come?'
B: Ung, Mimi-lopwuthe. 'Yes, Mimi-SRC'
B′: *Ung, Mimi. 'Yes, Mimi'

The difference between (1) and (2) is that only the fragments in (1B, B′) have an overt correlate (*nwukwu-lul* 'who-ACC'), although all the fragments have correlates in the antecedent clauses.[3] This has the consequence that the fragments in (1B, B′) are both well-formed, one being marked for the same case as its overt correlate (*Haha-lul* 'Haha-ACC') and the other being caseless (*Haha*). Meanwhile, only the fragment in (2B) is well-formed, because it is marked for source (*Mimi-lopwuthe* 'Mimi-SRC'), that is, the case that would be marked on the implicit correlate.[4]

(question), REF (reflexive), SAL-UTT (salient utterance), SCASE (semantic case), SRC (source), and so forth.

[3] The reader will notice another difference between example (1) and example (2) is that (1) features structural case (accusative), while (2) features semantic case (see J.-B. Kim 2016). We will return to this difference, although it does not affect the pattern described in these examples.

[4] An important caveat is order with respect to Korean. While the facts just described hold for unembedded fragments, embedded fragments behave differently (J.-B. Kim 2015a; J. Kim 2017). As example (i) shows, the embedded fragment must be caseless rather than marked for accusative.

(i) Mimi-ka nwukwunka-lul manna-ss-nuntey nwukwu/*nwukwu-lul-i-nci
Mimi-NOM someone-ACC meet-PST-DECL-but who/*who-ACC-COP-QUE
molu-keyss-ta.
not-know-PRES-DECL
'Mimi met someone, but I do not know who.'

This pattern is observed in two structural case markers, nominative and accusative, but not in semantic case markers, which behave alike in embedded fragments and in unembedded fragments. J.-B. Kim (2015a) suggests that this pattern follows from the inability of the copula that is part of the structure of the fragment in (i) to assign structural case. Both structural and semantic case markers are again obligatory for embedded sprouting fragments (ii) and for embedded multiple fragments (iii):

(ii) Mimi-ka senmwul-ul pat-ass-nuntey, nwukwu-*(lopwuthe)-i-nci
Mimi-NOM present-ACC get-PST-but who-SRC-COP-QUE
molu-keyss-ta.
not.know-PRES-DECL
'Mimi got a present, but I do not know from who.'

(iii) John-i ecey mwuenka-lul nwukwunka-hanthey cwu-ess-nuntey, na-nun
John-NOM yesterday something-ACC someone-DAT give-PST-but I-TOP
mwues-ul nwukwu-*(hanthey)-i-nci molu-keyss-ta.
what-ACC who-DAT-COP-QUE not.know-FUT-DECL
'John gave something to someone yesterday, but I do not know what to whom.'

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

Before we proceed, a brief terminological note is in order. Although the term 'correlate' may seem somewhat vague at this point, we propose in Section 4 that a correlate may be realized as any maximal category smaller than S in the terminology of Barton (1991, 2006) (i.e. NP, VP, AdjP, AdvP, or PP) that is licensed by the grammar. We propose further that if there is an overt or implicit correlate for a fragment, it may at a minimum be any maximal category derived from the current argument structure of lexical heads in the antecedent, including optional arguments. We refer to all stranded XPs as fragments, although much of the literature we review uses the term 'remnants' to refer to stranded XPs that are wh-phrases and represent the construction sluicing, as in (3B).

(3) A: Mimi is waiting for someone.
B: For who?

Since Merchant (2004), it has been known that sluicing and other fragments behave similarly. They also exhibit the same sensitivity to the merger/sprouting distinction in English (Chung 2006) and in Korean (J.-B. Kim 2015a,b). Hence, the analysis of fragments that we defend in Section 4 readily carries over to sluicing. In the interest of clarity, however, we refer to stranded wh-phrases as remnants whenever we refer specifically to sluicing.

The examples in (1)–(2) pose a challenge for theories of ellipsis in terms of formulating identity conditions on fragments. It is well known that fragments exhibit case-matching effects in languages with overt case-marking systems, such as German (Ross 1969; Merchant 2001, 2004). Whatever case is marked on a fragment's correlate, the same case must be marked on the fragment. This holds true of Korean fragments in the sense that, while they may be caseless, they may not be marked for a different case than the correlate's case. So the fragments in (1) and (2) have to be marked for accusative and source, respectively. However, example (1) suggests that case identity is not always enforced, or to be more precise, that a fragment can undergo reduction of sorts that renders it caseless just in case its correlate is overt. This pattern is not unique to Korean; it in fact parallels the pattern observed in English and several other languages (see Merchant 2001, 2004; Chung 2006, 2013). We illustrate it for an English merger fragment in (4) and a sprouting fragment in (5), of which only the former may be realized in two ways, as a PP or NP.

(4) A: Who did Mimi wait for?
B: For Harvey./Harvey.

(5) A: Did the letter come?
B: Yes, from Mimi./*Yes, Mimi.

The PP fragment *For Harvey* in (4B) corresponds to the PP *who ... for* in the antecedent, and the NP fragment *Harvey* seems to be a reduced form of the PP *For Harvey* arising via P(reposition) drop. An identity requirement akin to the case-matching one we saw in Korean holds here as well, since it is not possible to change

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

the preposition *for* in the fragment to any other that mismatches the preposition present in the antecedent.

In addition to the above patterns, the syntactic category of a fragment can differ from that of the correlate. Example (6) shows an NP correlate paired with a PP fragment in Dutch when both categories are subcategorized by the verbal head (see Levelt & Kelter 1982).

(6) A: Wie laat Paul zijn viool zien?
who lets Paul his violin see
'Who does Paul show his violin to?'
B: Aan Toos.
to Toos
'Toos.'

We discuss these and similar examples separately in Section 3.2.

These data raise two questions relating to structural constraints on fragments: (1) what morphosyntactic forms fragments may take, and (2) why it is possible for fragments to alternate between what appears to be a more complex and a less complex structure only in merger cases. We address them here by first reviewing the existing proposals regarding syntactic constraints on fragments. This type of constraint has been the focus of much research within the direct interpretation (DI) approach to fragments, which generates fragments 'as is' and within which our analysis is also couched (Ginzburg & Sag 2000; Culicover & Jackendoff 2005; Sag & Nykiel 2011; Ginzburg 2012; J.-B. Kim 2015a,b; Jacobson 2016; Ginzburg & Miller 2018). A competing line of research that assumes underlying sentential structures for fragments places syntactic constraints on these underlying structures, but typically not on fragments (Merchant 2001, 2004, 2005, 2013; Weir 2014), though Chung (2013) and Abels (2017) are attempts to do so. We argue for a syntactic constraint on fragments that ensures that fragments have access, via their correlates, to the argument structure of lexical heads that are part of the antecedents. But this constraint is not enforced on all fragments (see Section 4). Those fragments whose morphosyntactic features are not licensed by the argument structure of any heads in their antecedents, because they either correspond to (overt or implicit) adjuncts or do not correspond to any phrases within the antecedents, are not subject to our syntactic constraint. We also assume that merger fragments may take as their correlates any maximal categories evoked from the antecedent, subject to constraints that the grammar already places on fragments and to processing-based constraints operating on anaphors. This assumption allows us to provide a principled motivation for constraints that previous research (Chung 2006, 2013; J.-B. Kim 2015a,b) has proposed to impose on sprouting fragments. Our proposal is based in a large part on English and Korean data, but it extends to other languages, as we illustrate by offering further crosslinguistic examples. We develop this proposal in Section 4. Before this, we first walk the reader through the existing theoretical proposals regarding syntactic identity under clausal ellipsis in Section 2

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

and then review, in Section 3, available evidence in favor of the identity constraint that we propose. Section 5 concludes.

2. Fragments and structural identity

There are two strands of research that have attempted to deal with the question of how syntactically parallel fragments and their antecedents must be. Our particular concern is the extent to which the merger/sprouting distinction has been central to formulating syntactic constraints on fragments. We begin with the strand of research that assumes underlying syntactic structure for fragments and then turn to the strand that does not do so.

2.1. *Movement and PF-deletion accounts*

Merchant (2004) proposes that fragments derive from sentential sources, such as (7B, B′), with fronting of the fragment (here either as PP or NP) and PF-deletion of the rest of the clause, but otherwise, there is only a semantic entailment condition on ellipsis.[5] It follows that there are no syntactic constraints on fragments themselves.

(7) A: Who did Mimi wait for?
B: For Harvey ~~Mimi waited~~.
B′: Harvey ~~Mimi waited for~~.

In later work, Merchant (2005, 2013) recognizes that certain argument structure mismatches (e.g. voice mismatch, as in (8)) are unavailable for clausal ellipsis, but his analysis of such mismatches relates to the underlying structures that embed fragments, and not to fragments themselves.

(8) A: Someone cleaned up Ben's room.
B: *By who ~~was Ben's room cleaned up~~?

Case-matching effects are also the consequence of fragments being embedded in underlying sentential structures, with case being assigned by appropriate lexical heads, which must then be the same as the lexical heads assigning case to fragments' correlates. Whether fragments surface as PPs or NPs is similarly linked to the underlying structures that embed them. Two kinds of underlying structures are found in the literature: declarative clauses (or interrogative clauses in the case of sluicing) that match the structure of antecedents (Merchant 2004) and copular clauses that do not (Rodrigues, Nevins & Vicente 2009; Van Craenenbroeck 2010; Barros 2014; Griffiths & Lipták 2014). One of these structures is expected to apply to derive NPs or PPs as fragments, as needed. For instance, in languages that do not tolerate fronting of prepositional objects for lack of a preposition-stranding rule, and hence cannot derive

[5] For more on this condition, see Merchant (2001).

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

NPs as fragments from declarative clauses like (7B′), copular clauses (9) could in principle be used as underlying sources.[6]

(9) A: Who did Mimi wait for?
B: Harvey ~~it was~~.

Implicit in these accounts is an assumption that fragments may freely take either NPs or PPs as their correlates (but see Kobele 2015 for an explicit statement to this effect). This assumption, however, leaves it unclear why such freedom exists for merger fragments, but not for sprouting fragments. Indeed, accounting for the merger/sprouting difference has proven difficult, since it is well known that sprouting requires that a fragment share the semantic and syntactic features of the parallel position in the structure of its antecedent or, in other words, that it be no less complex a phrase than that position allows (see Chung 2006, 2013; Culicover & Jackendoff 2012). To capture this requirement, Chung (2013) proposes a syntactic condition on DP fragments such that they be case-licensed by the same lexical heads that case-license their correlates (see further Section 4.3). Chung (2013) imposes further argument structure identity on antecedents and underlying structures that embed fragments, but her proposal conflicts with the possibility of using copular clauses as underlying structures, which clearly violate any kind of identity requirement based on argument structure (see e.g. Thoms 2015 for a critique of Chung 2013).

An attempt to reconcile merger and sprouting fragments with a more relaxed identity condition than Chung's (2013) is found in Abels (2017). His fit condition, as stated in (10), is formulated for sluicing.

(10) *Fit condition*
Modulo agreement in the antecedent and wh-movement, replacing the correlate by the remnant in the antecedent must lead to a syntactically well-formed structure with the right meaning or – for sprouting – adding the correlate into the antecedent and making no further changes must lead to a syntactically well-formed structure with the intended thematic interpretation.

(10) requires by implication that all remnants fit into the structure of the antecedents as well as into the structure of the underlying clauses that host them, because Abels (2017) admits the possibility that underlying structures are either interrogative or copular structures. The data that constitute the basis for formulating (10) come from

[6] The expectation that NPs cannot appear as fragments unless a language permits preposition stranding is part of the preposition-stranding generalization (Merchant 2001) and constitutes one of the reasons why copular clauses have been suggested as alternatives to isomorphic clause sources. This reasoning remains controversial, however, given a body of crosslinguistic evidence demonstrating that copular clauses are not available sources for NP fragments in several languages without preposition stranding, such as Greek (Molimpakis 2018), Emirati Arabic (Leung 2014), Russian (Philippova 2014), Polish (Nykiel 2013), and Serbo-Croatian (Stjepanović 2008, 2012). A related difficulty arises for Korean, where adequate underlying sources cannot be found for fragments, including for fragments that undergo case drop (J.-B. Kim 2015a,b; Kim 2017). We return to this issue in Korean in Section 3.

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

Bulgarian case markers available for pronominal objects. To illustrate, (11) shows two case-marking options for the verbal object in the antecedent clause: general case (GC) or non-subject case (NON.S). Abels reports that the same options are inherited by the sluicing remnant, with a preference for case match such that both the remnant and the verbal object are marked for either GC or NON.S.

(11) Ivan sreshtna njakoi/njakogo no ne znam
Ivan met someone-GC/someone-NON.S but not I-know
koi/kogo.
who-GC/who-NON.S
'Ivan met someone but I do not know who.'

The Bulgarian data are interesting for two reasons. First, they show that the morphosyntactic form of a remnant and the morphosyntactic form of its correlate are both constrained by the argument structure of the lexical head that licenses the correlate. Abels (2017) makes this point only indirectly by pointing to the necessity of a remnant fitting into the structure of its antecedent. Second, the remnant tends to follow the morphosyntactic form of its overt correlate once a choice has been made, that is, once one of the available cases has been selected by the correlate.

As far as P-drop, Abels reports that a remnant may be realized as an NP only if it is marked for GC, as in (12), and otherwise, it will surface as a PP.

(12) Ivan tancuva s njakoi/njakogo no ne znam
Ivan danced with someone-GC/someone-NON.S but not I-know
koi.
who-GC
'Ivan danced with someone but I do not know who.'

Abels's analysis of this pattern is that remnants marked for GC, unlike remnants marked for NON.S, could derive from copular clauses as underlying sources (see example (9) above), since copular clauses require GC-marked subjects. Abels argues that an NP remnant marked for GC fits into the structure of its antecedent as well. Taking the syntactic category of the remnant as our guide, we locate the phrase that matches that category in the antecedent (*njakoi*), replace it with the remnant, and confirm that the resultant structure is well-formed (we leave it to the reader to check that (10) is also satisfied if the remnant in (12) is realized as a PP).

The fit condition also predicts correctly that for sprouting cases, nothing smaller than a PP fits the structure of an antecedent if there is an implicit PP slot in it, and hence a remnant must be a PP, as in (5) repeated here as (13).

(13) A: Did the letter come?
B: Yes, from Mimi./*Yes, Mimi.

However, we will see examples of sprouting fragments in Section 4 that do not fit into the structure of their antecedents and are therefore clearly problematic for the fit condition.

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

To summarize, the movement and PF-deletion approach relies on two kinds of underlying structures to be available for clausal ellipsis, while at the same time attempting to reconcile the behavior of merger fragments with the behavior of sprouting fragments. The result is a tension between allowing a variety of underlying structures (but recall the problem that legitimate underlying structures are not available in every language, see fn. 6.) and constraining how fragments may be realized, which an account like Abels (2017) is an attempt to solve.

2.2. *Direct interpretation accounts*

The central idea behind DI accounts of fragments is that they have no more structure than is visible and hence are not constituents of full clauses (Levin 1982; Ginzburg & Sag 2000; Culicover & Jackendoff 2005, 2012; Sag & Nykiel 2011; Ginzburg 2012; J.-B. Kim 2015a,b; Kim & Abeillé 2019). We view this characteristic of DI accounts as an advantage over movement-and-deletion accounts. We can implement the idea that the argument structure of lexical heads present in the antecedent determines in all relevant cases the range of XPs that are available to serve as fragments regardless of whether we find appropriate underlying sources for the fragments. Movement-and-deletion accounts cannot do the same without losing the ability to derive fragments that require nonisomorphic sources, as discussed in Section 2.1.

The semantic and morphosyntactic features of a fragment are licensed non-locally: they are inherited from the fragment's correlate in the antecedent clause (which Ginzburg & Sag 2000 term salient utterance and which Culicover & Jackendoff 2005 term target) via a mechanism that integrates the fragment into a proposition constructed on the basis of the antecedent. Fragments are aligned with overt (merger) or implicit (sprouting) phrases functioning as their correlates, and the transfer of relevant semantic and syntactic features (e.g. case and preposition choice) is an integral part of this process (see, in particular, Culicover & Jackendoff 2005). An early proposal that makes use of the idea that fragments inherit the grammatical functions of their correlates along with the relevant morphosyntactic features non-locally is found in Levin's (1982) Lexical Functional Grammar analysis of sluicing. In the proposals of Levin (1982) and Culicover & Jackendoff (2005) there is no requirement that the syntactic category and/or case of a fragment match those of its correlate; rather, featural match follows from the fact that a constituent with a certain grammatical function must carry the features that are licensed for that function. However, Culicover & Jackendoff (2005) distinguish between direct NP arguments and oblique NP arguments (NPs functioning as prepositional objects), which they use to ensure that fragments whose correlates are implicit PPs are realized as oblique NP arguments together with the correct prepositions. This makes syntactic category a syntactic feature that is relevant for fragments.

An important aspect of Culicover & Jackendoff's analysis is their treatment of optionality in the argument structure of lexical heads in the antecedents. There are fragments that cannot be matched with any overt correlates in the current argument

structure of a head. For instance, we must interpret the fragment in (14B) as an oblique NP argument of the verb *flirt* despite the fact that no such oblique argument is licensed by the intransitive use of *flirt* in (14A).

(14) A: Harriet has been flirting again.
B: Yeah, with Ozzy. (Culicover & Jackendoff 2005: 259)

Culicover & Jackendoff propose that the lexical entry of *flirt* includes the optional oblique NP argument and is activated in memory along with the intransitive variant whenever this verb is encountered in discourse. This lexical entry of *flirt* can therefore license the fragment in (14B) as an oblique NP argument. This idea will be relevant for our discussion of sprouting fragments in Section 4.3.

Ginzburg & Sag (2000), working within the framework of Head-driven Phrase Structure Grammar (HPSG), constrain fragments to be the category nominal (i.e. NPs or PPs), and correlates, whether implicit or overt, must match this category. This is a requirement that consistently reappears in much later work (Sag & Nykiel 2011; J.-B. Kim 2015a,b; Jacobson 2016; Abeillé & Hassamal 2019). However, in none of these works is it clear how one would go about separating out merger fragments from sprouting fragments with respect to the (im)possibility of dropping prepositions (or case markers in Korean). The usual assumption that NP fragments, unlike PP fragments, take NPs as correlates (see e.g. Sag & Nykiel 2011) leaves it unclear why NP correlates are possible only for merger fragments. J.-B. Kim (2015a) attempts to block case drop under Korean sprouting by formulating an additional constraint, which has a similar function to the lexico-semantic constraint of Chung (2006) blocking elements not present overtly in the antecedent from being dropped under clausal ellipsis. We will formulate a constraint that is sensitive to the merger/sprouting distinction and provide a processing-based motivation for it.

The general consensus in the DI literature is that syntactic constraints are not enforced on fragments across the board. Discussions of this point center on fragments that have adjuncts as their correlates and whose syntactic features therefore do not depend on any lexical heads in the antecedent (see Culicover & Jackendoff 2005; Ginzburg 2012). To illustrate, consider (15) from COCA.

(15) A_1: You write about that?
B_1: I did.
A_2: In what?
B_2: American Scholar.

The fragment in ($15A_2$) represents sprouting of a VP adjunct, excluding the possibility that any syntactic features of the fragment are licensed by the verbal head. Similarly, Ginzburg (2012:236) notes that locative and temporal expressions impose no syntactic category restrictions on merger fragments. A variety of syntactic categories are permitted as fragments corresponding to the adverb *when*, as in (16).

(16) A: When did Bo arrive?
B: At 2./Recently./The day after Mo left.

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

We will likewise not enforce any syntactic constraints on all fragments, but rather ensure that relevant fragments have access to all the morphosyntactic information available from the argument structure of lexical heads in the antecedents.

In the remainder of this paper, we defend a syntactic condition on fragments involving both syntactic category and case features, and we let processing principles interact with it to capture the merger and sprouting facts. Our analysis, articulated in Section 4, follows the DI approach, which has been convincingly argued both for English (Ginzburg & Sag 2000; Culicover & Jackendoff 2005, 2012; Jacobson 2016) and for Korean (J.-B. Kim 2015a,b; J. Kim 2017). It remains unclear how structural identity of this kind could be implemented on movement-and-deletion approaches without sacrificing the possibility of deriving fragments from a variety of underlying sources. Like Ginzburg & Sag (2000), we use the framework of HPSG. This choice is dictated by the explicitness of representing various aspects of the argument structure of lexical heads, on the one hand, and the admissibility of (extra-grammatical) construction-specific constraints on fragments, on the other (see Section 4).

3. Evidence for structural identity

This section reviews crosslinguistic evidence in favor of imposing on some fragments structural identity constraints involving both syntactic category and case features based on the argument structure of lexical heads in the antecedents. We first turn to sprouting fragments and then address merger fragments standing in two kinds of relationship with their antecedents.

3.1. *Sprouting fragments*

It is well known that sprouting fragments are subject to strong syntactic identity constraints if they pick out arguments as their correlates (Ross 1969; Chung, Ladusaw & McCloskey 1995, 2011; Chung 2006, 2013; Culicover & Jackendoff 2012). We mentioned in the introduction that this holds true of both English and Korean. Consider again (2) and (5) repeated below:

(17) A: Did the letter come?
B: Yes, from Mimi.
B′: *Yes, Mimi

(18) A: Phyenci-ka wa-ss-e?
letter-NOM come-PST-QUE
'Did the letter come?'
B: Ung, Mimi-lopwuthe. 'Yes, Mimi-SRC.'
B′: *Ung, Mimi. 'Yes, Mimi.'

The English fragment in (17B) corresponds exactly to a PP argument and the Korean fragment in (18B) to an NP marked for source, and if these correspondences do not hold, ungrammaticality arises, as in (17B′) and (18B′). In the oft-cited

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

examples of sprouting fragments the correlate is usually an argument within VP, but sprouting can reach into constituents that are more deeply embedded in the antecedent structure. Consider (19)–(20) from COCA.

(19) A: You're a case study, Sy.
B: A case of what?

(20) A_1: With trembling fingers the fellow gathered up his scraps from the desk.
B: Trembling with what?
A_2: With timidity.

In both examples, modifiers within NP (*case* and *trembling*) license the material being sprouted, which again must preserve the syntactic category restrictions imposed by the licensers.

These patterns demonstrate clearly that sprouting fragments must be structurally no smaller than the implicit positions in the current argument structure of given lexical heads in the antecedents, with some flexibility in terms of which heads fragments actually access. To the best of our knowledge, there are no challenges to this generalization that is termed Chung's Generalization, after Chung (2006), as far as languages with prepositions are concerned (though see Section 4.3).

As for Korean, Nykiel et al. (2018) reported, as expected, severely degraded acceptability ratings for caseless sprouting fragments like (18B′), and this also applies to omission of structural case markers, as seen in the degraded caseless fragment corresponding to an implicit accusative-marked NP in (21B′).

(21) A: Chelswu-ka pat-ass-ney.
Chelswu-NOM receive-PST-DECL
'Chelswu received (it).'
B: Ung, sangkum-ul.
yes, prize-ACC
'Yes, the prize.'
B′: *Ung, sangkum.
yes, prize
'Yes, the prize.'

Korean sprouting fragments thus constitute a further robust argument for requiring morphosyntactic identity between them and their correlates in the event that the correlates are members of the argument structure of some lexical head in the antecedent.

However, sprouting reveals yet another aspect of the correspondence between fragments and their antecedents. Two case markers *-lopwuthe* and *-eykey* can express the source case in Korean. Either of them may be marked on the sprouting fragments in (22B, B′). This example indicates that the status of these fragments is radically different than the status of caseless sprouting fragments like (18) and (21), which are judged as unacceptable. We take the behavior of the sprouting fragments

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

in (22) as evidence that sprouting permits variation in case marking only if this variation is licensed by the argument structure of a given lexical head.

(22) A: Phyenci-ka wa-ss-e.
letter-NOM come-PST-DECL
'A letter came.'
B: Ung, Mimi-eykey. 'Yes, Mimi-SRC.'
B′: Ung, Mimi-lopwuthe. 'Yes, Mimi-SRC.'

Similar variation, this time in terms of syntactic category, is found in Polish sprouting fragments. Consider (23), where either a dative-marked NP (23B) or a PP (23B′) is available as a fragment.

(23) A: Harvey stawia drinki.
Harvey-NOM buys drinks-ACC
'Harvey is buying drinks.'
B: Tak, wszystkim.
yes everybody-DAT
'Yes, for everybody.'
B′: Tak, dla wszystkich.
yes for everybody-GEN
'Yes, for everybody.'

These possibilities follow from the fact that either a dative NP or a PP may realize the indirect object left implicit in (23A). The indirect object, whether appearing as an NP or PP, can either precede or follow the direct object *drinki* ('drinks') in (23A), and this may be contributing to the acceptability of both of the Polish fragments in (23B, B′) compared to their English counterparts. The English counterparts of these fragments depicted in (24B, B′) show a different acceptability pattern: the fragment in (24B′) seems unacceptable although an indirect object NP alternates with a PP (e.g. *Harvey is buying everybody drinks* vs. *Harvey is buying drinks for everybody*) in full clauses (see also Chung et al. 1995 and Merchant 2001 for related examples from English).

(24) A: Harvey is buying drinks.
B: Yes, for everybody.
B′: *Yes, everybody.

We assume that this difference between English and Polish is due to the fact that, while in Polish a single argument slot permits two syntactic categories without any accompanying syntactic rearrangement of the other arguments (see Section 3.2 for similar data from Dutch), the choice between an NP and PP in English is linked to different word orders and two different kinds of argument structure. Hence, the structure in (24A) commits the indirect object to being realized as a PP.

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

In sum, the sprouting facts discussed above indicate that a fragment has access to all the slots in the current argument structure of lexical heads in its antecedent, including to slots filled by elements optionally realized overtly and by elements that alternate with other elements in terms of case or syntactic category. Recall from Section 2.2 that Culicover & Jackendoff's (2005) analysis of fragments already permits optionality in the realization of particular slots in the argument structure of verbal heads to influence the range of fragments that are licensed.

3.2. *Merger fragments and variation in argument structure*

This section is concerned with fragments of the merger type similar to examples like (22) and (23) above. Judgments about fragments corresponding to alternating phrases change when one of the options is selected explicitly in the antecedent. We illustrate this point first with Korean merger fragments in (25B, B′) and then with Polish merger fragments in (26B, B′). The source argument is expressed by means of the -lopwuthe case marker (not -eykey) in the antecedent in (25A), and the mismatching fragment (25B) is degraded (though not ruled out completely) compared to the matching one (25B′).

(25) A: Nwukwu-lopwuthe phyenci-ka wa-ss-e?
who-SRC letter-NOM come-PST-QUE
'From whom did the letter come?'
B: ?Ung, Mimi-eykey. 'Yes, Mimi-SRC.'
B′: Ung, Mimi-lopwuthe 'Yes, Mimi-SRC.'

The Polish fragment in (26B), which mismatches the recipient argument's syntactic category (the PP *dla nich* 'for them' in (26A)) is also degraded.

(26) A: Harvey stawia dla nich drinki.
Harvey-NOM buys for them-GEN drinksACC
'Harvey is buying drinks for them.'
B: ?Komu?
who-DAT
'For who?'
B′: Dla kogo?
for who-GEN
'For who?'

The Bulgarian sluicing data we discussed in Section 2.1 fall into the same category. Recall Abels's (2017) example in (11), repeated here as (27).

(27) Ivan sreshtna njakoi no ne znam kogo/koi.
Ivan met someone-GC but not I-know who-NON.S/who-GC
'Ivan met someone but I do not know who.'

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

The remnant may surface as marked for either of the cases licensed for the verbal object, which functions as the correlate. Abels (2017) notes that there is a preference for matching the case features of the remnant and correlate, but mismatch is also possible.[7]

There is experimental support for the matching effects noted above. Levelt & Kelter (1982) reported on Dutch data where they manipulated the syntactic category of arguments selected for by the Dutch verb *laten* 'let' in questions like (28) and (29). This verb selects for an NP (28A) or a PP (29A) and permits either NP fragments or PP fragments to be used as answers, as in (28B) and (29B). Levelt & Kelter's main finding is what they term a 'correspondence effect' such that matching syntactic categories (the NP fragment in (28B) and the PP fragment in (29B)) were consistently favored over mismatching ones.

(28) A: Wie laat Paul zijn viool zien?
who lets Paul his violin see
'Who does Paul show his violin to?'
B: Aan Toos / Toos.
To Toos / Toos
'Toos.'

(29) A: Aan wie laat Paul zijn viool zien?
to who lets Paul his violin see
'Who does Paul show his violin to?'
B: Aan Toos / Toos.
To Toos / Toos
'Toos.'

These data show that a fragment has access to the same argument structure options as those available to its correlate at the same time as it preferably replicates the morpho-syntactic features chosen for its correlate. The data thus strengthen our argument that some fragments are subject to structural constraints imposed on their correlates on the basis of the argument structure of the head that licenses the correlates.

We suggest that the degradation associated with mismatching fragments and correlates (whether it is case mismatch, as in Bulgarian and Korean, or syntactic category mismatch, as in Dutch or Polish) follows from a violation of the preference for matching features, even if variation is permitted. It is important to note that the preference for matching features is a preference that can be violated only in cases where the correlate may be realized in more than one way in the antecedent; otherwise, the features of the fragment and the correlate will always match (we are not aware of any acceptable examples of fragments

[7] Jacobson (2016) reports on Hungarian examples where a single verb selects two different case markings, similar to the Bulgarian facts, but where native speakers almost always opt for matching the cases marked on a fragment and its correlate. It appears then that case matching in Hungarian is closer to a requirement than a preference.

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

mismatching the features of their correlates when the argument structure of a given lexical head does not license any variation). Matching overt features can be interpreted as facilitating the process of locating a fragment's correlate. This process is known to involve a direct-access mechanism that relies on the features of a fragment and finds the best match for it among content-addressable linguistic representations previously stored in memory (Martin & McElree 2011). More generally, this type of direct search for the target phrase among stored memory representations is part of the cue-based models of sentence processing (see McElree 2000; McElree, Foraker & Dyer 2003; Lewis & Vasishth 2005; Lewis, Vasishth & van Dyke 2006; Van Dyke 2011; Van Dyke & Johns 2012; Caplan & Waters 2013) and is affected by the 'cue specificity' of the anaphor that initiates the search (here, the fragment) such that the more features that are shared between the target and the anaphor, the easier the search. To illustrate, Harris (2015) offered evidence for processing benefits associated with matching correlate/remnant features under sluicing. Sluicing remnants with fully specified nominal heads like *which tourists* in (30) were better cues to their correlates (i.e. were read faster) than remnants with partially specified nominal heads like *which ones* in (30).

(30) Some tourists sampled the wines but I've forgotten which tourists/which ones.

It follows from this research that the search for the correlate for a fragment becomes harder in the event that the morphosyntactic features of the fragment are underspecified or mismatched with the features of the correlate.[8]

We see the bias toward matching features under merger even in code switching. Gonzalez-Vilbazo & Ramos (2014) found a preference for case matching in sluicing remnants where an antecedent was in Spanish and a clause that embeds a remnant, and the remnant itself, in German. In (31), the preference is for the accusative-marked remnant (*wen*), which matches the accusative-marked Spanish correlate. The dative-marked remnant (*wem*), which would be selected for by the German equivalent of *amenazar* ('threaten'), is degraded in comparison (as is the nominative-marked remnant (*wer*), which neither the Spanish nor the German verb selects for).

(31) Xavi amenazó a alguien. Rate
Xavi threatened to someone-ACC guess
*wer/ wen/ *wem.
*who-NOM/WHO-ACC/*WHO-DAT
'Xavi threatened someone. Guess who.'

[8] It is relevant to note here that syntactic parallelism has been observed as a processing preference both elsewhere in ellipsis and outside of it. For instance, Parker (2018) demonstrates that verb phrase ellipsis is subject to retrieval-based preferences, such that different degrees of match between an ellipsis and its antecedent in terms of morphosyntactic features lead to differential processing difficulty and gradient acceptability patterns reported in the ellipsis literature. Carlson (2002) demonstrates structural parallelism effects in the processing of the elliptical construction gapping. Outside of ellipsis, research starting with Frazier et al. (1984) shows that structural parallelism aids in the processing of coordinated sentences (see e.g. Dubey, Sturt & Keller 2005 and Callahan et al. 2010).

We conclude that while variation in case/syntactic category features is generally permitted for merger fragments whenever it is also permitted by the argument structure of an appropriate lexical head, features mismatching the features of the correlates are degraded for independent reasons to do with the ease of processing.

3.3. *Merger fragments and form minimization*

In this section, we return to the patterns of case drop and P-drop that we saw in the Korean and English examples in (1)–(5) in Section 1. We argue that both of these phenomena have a purely processing-based motivation and are restricted to easy-to-process environments.

Given the data we have reviewed thus far, we propose that in the event that a fragment picks out as its correlate a member of the argument structure of a lexical head in its antecedent, it may pick out any member of the current argument structure, including optional arguments and alternations within a single argument slot involving case or syntactic category. The fragment must then inherit all the semantic and morphosyntactic features that are appropriate for the correlate.[9] However, we also need to take into account that fragments can vary in size, moving along the structure of the antecedent. For instance, all three XPs in (32B, B′, B″) and (33B, B′, B″) are possible fragments, but only the B″-fragments avoid repeating any parts of the antecedents.[10] The B-fragments are authentic (from *City of Lies*, R. J. Ellory 2010 and COCA respectively), and the rest are constructed.

---

[9] An anonymous referee points to example (i), where the B fragment picks out an argument that does not belong to the current argument structure of the predicate and indeed refers to a different meaning of *tired* than that featured in the antecedent.

(i) A: Are you tired?
B: Yes – of you!

Below is another example.

(ii) A_1: She's out.
B: Out of what?
A_2: Her freaking mind. (TV Corpus)

It seems to us that such examples are of somewhat humorous nature and require the interlocutors to accommodate the new argument structure and meanings of the antecedent predicates.

[10] Larger-sized fragments like (33B) have been argued to be illicit if both embedded and following the matrix clause, as in (i) (Ross 1969; Bechhofer 1976; Abels 2019).

(i) *Harvey's become mixed up in something, but I do not know mixed up in what.
Reversing the order of the fragment and matrix clause produces a grammatical result, as in (ii).
(ii) Harvey's become mixed up in something, but mixed up in what, I don't know.

Consider, however, the sprouting fragments below following the predicates *wonder* and *want to know* instead of *know*. Of these, (vB) appears to come closer to embedding than (iii) and (iv), which, as an an anonymous referee notes, are perhaps better characterized as parentheticals.

(iii) I remember watching TV when I was little, and watching commercials about things that would help adults do things faster, because they knew how 'busy' people were. I couldn't help

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

(32) A: What the hell are you talking about?
B: Talking about one of your people.
B′: About one of your people.
B″: One of your people.

(33) A: I've become mixed up in something.
B: Mixed up in what?
B′: In what?
B″: What?

The fragments in (32B″) and (33B″) correspond to the arguments of the prepositional heads *about* and *in*. They can be seen as minimal XPs that elaborate on the propositions raised by the antecedents by picking out specifically the focused elements as their correlates (the wh-phrase *what* and the NP *something*). In contrast, the B- and B′-fragments pick out larger phrases containing the focused elements. We follow Barton (1991, 2006) in assuming that the sizes generally permitted for fragments are maximal categories that the grammar can generate, i.e. NP, PP, VP, AdjP, or AdvP, which also explains why certain reductions in fragment size (object drop in (34B) and Nominal ellipsis in (35B)) are not allowed in English.

(34) A: What did they do to my car?
B: *Commandeered.

(35) A: What kind of dog would you get?
B: *Small/*A small.

We could also invoke here Lambrecht's (1994) assumption that a fragment must be phrasal, not lexical, in form, corresponding to the constituent of the antecedent that

wondering, busy with what? (http://www.crossworldsbooks.com/blog/wednesday-word-thoughts-on-upside-down-in-a-laura-ingalls-town-by-leslie-tall-manning)

(iv) As the days passed, I was so afraid. Well, you might wonder, afraid of what? (https://www.saintfrancisborgia.org/parish-activities-and-outreach/uganda-twinning/uganda-twinning-experiences-2013)

(v) A: 'I was afraid.' Her response was barely audible above the sounds of the water lapping on the shore. Liam put one finger under her chin, forcing her to look at him. 'Afraid of what?' Tears shimmered in his eyes as she saw bewilderment win out over outrage for a second.
B: I also want to know afraid of what because he still seems like a very sweet guy. (https://nanreinhardt.com/weekend-writing-warriors-8-sentence-sunday-8)

Predicates like *wonder* and *want to know* are grouped into a distinct class of rogative predicates while predicates like *know* belong to the class of responsive predicates (e.g. Lahiri 2002; Uegaki 2015; Theiler, Roelofsen & Aloni 2019). This classification follows from distinct semantic restrictions that these predicates place on their complements. But it has also been proposed that the syntax of the complements that these predicates take is nonuniform. One suggestion is that rogative predicates take CPs as their complements while responsive predicates take free relative clauses in the form of DPs (see Turnbull-Sailor 2007 for this classification, along with a set of (syntactic) diagnostics). We leave it as an open question whether or how the possibility of embedding fragments under various predicates is mediated by the semantics of these predicates, but note that material following rogative predicates appears to behave like matrix phenomena (see Turnbull-Sailor 2007 for further discussion).

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

he terms the 'focus domain', a role that may be filled by a predicate, an argument, or an adjunct.

To return to examples (32) and (33), the minimal NPs in (32B″) and (33B″) are grammatically well-formed fragments, but the question is why they are permitted only as merger-type fragments. We are particularly concerned with P-drop, that is, the difference between the B′- and B″-fragments, as this difference is criterial for the merger/sprouting distinction. Before we proceed, note that (36) illustrates a similarly wide range of sizes for Korean fragments, from the largest (36B) to the smallest, a caseless NP (36B‴).

(36) A: salamtul-i mwuenka-lul wihay kitohako-iss-e.
people-NOM something-ACC for pray-PROG-DECL
'A: People are praying for something.'
B: Ung, phyenghwa-lul wihay kitohay.
yes, peace-ACC for pray
'Yes, praying for peace.'
B′: Ung, phyenghwa-lul wihay. 'Yes, peace-ACC for'
B″: Ung, phyenghwa-lul. 'Yes, peace-ACC '
B‴: Ung, phyenghwa. 'Yes, peace'

Now, recall that the direct-access mechanism that retrieves a correlate for a fragment, which we discussed in Section 3.2, relies on the specificity of the cues provided by the fragment. This means in turn that the fragment that contains more overt material also carries more retrieval cues that match the antecedent, reducing the difficulty of first finding the antecedent and then integrating the fragment into the proposition expressed by the antecedent. The smallest fragments, which carry minimally few retrieval cues, are expected to be the hardest to process from this perspective. This expectation finds clear crosslinguistic support in the sense that languages employ the smallest fragments (NPs) less than they do larger-sized PPs (e.g. Stjepanović 2008 for Serbo-Croatian, Szczegielniak 2008 for Polish, and Rodrigues et al. 2009 for Spanish and French), and acceptability judgment experiments show a degradation for NP fragments compared to PPs (Nykiel 2013 for Polish, Merchant et al. 2013 and Lemke 2016 for German, and Molimpakis 2018 for Greek). [11] [12] If we consider that the fragment is also a whole processed independently of its antecedent, we could argue that there are additional advantages to using PPs rather than NPs. Certain syntactic and semantic dependencies that are resolved within PP (i.e. case assignment, subcategorization relations and various semantic ambiguities) might prefer its constituents to be adjacent within the fragment (for more detail, see Nykiel & Hawkins 2020). But then, why is P-drop permitted at all

[11] These patterns do not hold in present-day English, where NPs are more frequent as fragments than PPs (Nykiel 2017). See Nykiel & Hawkins (2020) for an explanation of why P-drop is exceptionally common in English that relies on numerous semantic dependencies that prepositions contract with verbal heads.

[12] There is some variation here regarding the kind of clausal ellipsis tested. Nykiel (2013) and Molimpakis (2018) focus on sluicing, Lemke (2016) focuses on fragment answers to wh-questions, and Merchant et al. (2013) employ fragment answers to yes/no questions.

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

(and more widespread than what movement-and-deletion accounts would predict based on the pattern of preposition stranding, recall fn. 6)?

To explain the possibility of P-drop we appeal to Hawkins's (2004) principle of Minimize Forms, one of his principles of efficient language processing, given in (37).

(37) Minimize Forms (MiF) (Hawkins 2004: 31)
The human processor prefers to minimize the formal complexity of each linguistic form F (its phoneme, morpheme, word, or phrasal units) and the number of forms with unique conventionalized property assignments, thereby assigning more properties to fewer forms. These minimizations apply in proportion to the ease with which a given property P can be assigned in processing to a given F.

Nykiel & Hawkins (2020) argue that MiF applies to fragments by reducing their complexity from a PP (or a larger-sized XP) to an NP if the required syntactic and semantic properties can be easily assigned to them. Crosslinguistic data relating to both fragments and anaphora more generally show that the ease with which these properties are assigned is mediated by the accessibility of the antecedent. One way to think of the accessibility of the antecedent is in terms of the semantic and syntactic richness of the phrase serving as the correlate for a fragment. For instance, the minimal NP fragment has the indefinite pronoun *someone* as its correlate in (38) but the lexical NP *a lawyer* in (39). This difference results in the lexical NP correlate in (39) being a semantically and syntactically richer phrase, which makes it a more accessible correlate for the minimal fragment. There is a greater preference for an NP fragment in (39) than in (38) due to the greater ease of matching it with its accessible correlate.

(38) A: I went to talk to someone.
B: Who?

(39) A: I went to talk to a lawyer.
B: Who?

This preference has been reported in several studies of fragments (Stjepanović 2008 for Serbo-Croatian; Szczegielniak 2008, Sag & Nykiel 2011, and Nykiel 2013 for Polish; Rodrigues et al. 2009 for Brazilian Portuguese, Spanish, and French; Caha 2011 for Czech; and Nykiel 2015, 2017 for English) and aligns well with independent research results on memory retrieval and anaphora.

The accessibility difference between pronouns and lexical NPs can be thought of in terms of the strength of the mental representations that they receive once they have been processed. During processing semantically and syntactically richer expressions like lexical NPs receive stronger mental representations than expressions that are less contentful (e.g. pronouns), which boosts their accessibility for future retrieval from memory (Craik & Tulving 1975; Fisher & Craik 1980; Marks 1987; Gallo et al. 2008; Hofmeister et al. 2007, 2013; Hofmeister 2011). We would

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

expect therefore that phrases with stronger mental representations are more accessible antecedents for anaphors, with the result that fewer cues are required for their successful retrieval and hence less contentful anaphors are preferred. Consistent with this expectation, Karimi et al. (2014) found that the length (in words) of a nominal antecedent affects the choice of anaphor (lexical NP vs. pronoun) that refers back to it: the longer the antecedent the less contentful the anaphor.

When we think of accessibility this way, the burden of providing sufficient retrieval cues shifts from the fragment to the correlate in cases where the fragment is minimally contentful (an NP). The correlate can support the retrieval process here if it is contentful, and we expect to see the kind of gradience that is reported in the crosslinguistic literature on P-drop, with lexical NPs being preferred as correlates for NP fragments. But if the correlate has no content at all, as is the case with sprouting, then its degree of accessibility is at its lowest and the retrieval process becomes difficult, if not impossible, on account of neither the correlate nor the NP fragment providing unambiguous retrieval cues. More generally, we would predict an overall processing advantage for overt material over non-overt material that precedes ellipsis regardless of P-drop. Such an advantage has indeed been reported in reading comprehension studies of ellipsis, such that sprouting fragments were read slower than the corresponding merger fragments (Frazier & Clifton 1998; Dickey & Bunger 2011). The crosslinguistic ban on reducing sprouting fragments to NPs could thus be seen as a construction-specific constraint on fragments that has arisen as a conventionalization of what MiF heavily disfavors (i.e. pairing minimal fragments with low-accessibility correlates) as a grammatical principle,[13] along the lines of (40).[14]

(40) *No form minimization* (NFM)
The application of MiF is blocked if the resulting minimal fragment corresponds to a subconstituent of an implicit argument in the antecedent.

We will keep NFM as a constraint in our proposal, but, alternatively, we could leave it to MiF to rule out minimal fragments under sprouting.

By applying MiF to fragments, we can also account for English merger examples like those in (41)–(42).

(41) A: What does it look like?
B: Like keys. (Seinfeld)

[13] See Hawkins (2004, 2014) for the Performance-Grammar Correspondence Hypothesis, which states that grammars are impacted by the strength of performance preferences, such as processing pressures, and respond by conventionalizing them to various degrees. For instance, grammars may resolve processing pressures by making the disfavored variant ungrammatical rather than keeping it as a low-frequency option alongside other, more frequent, options.

[14] While there exist other formulations of a constraint on sprouting (e.g. Chung 2006, 2013; J.-B. Kim 2015a; see Section 4.3 for more detail), they are rather descriptive, while ours is motivated as a limiting case of MiF.

(42) A: On what does the well-being of the EU depend?
B: A stable currency. (Ginzburg & Sag 2000:301)

In (41B) the PP fragment is a larger phrase than what its correlate appears to be (the NP *what*), and conversely, the NP fragment in (42B) is a smaller phrase than its apparent correlate (the pied-piped PP *on what*) (note that the status of (42B) is not the same as that of the NP fragment in the Dutch example in (29B), because an NP cannot replace a PP in the argument structure of the verbal head *depend* in (42A)). This is possible because form minimization is optional here: the fragment may be realized either as a larger maximal category than the minimal phrase (41) or it may simply be the minimal phrase (42). The question that remains, and which we address in Section 4, is how to determine what the fragment's correlate is.

Before concluding this section, we offer additional evidence that P-drop is an instance of form minimization and that this kind of form minimization is specific to fragments. The best indicators of form minimization are found in languages with overt case marking, such as Greek. Molimpakis (2018) demonstrated experimentally that in the Greek example in (43), the NP fragment must still be marked for the caselicensed by the dropped preposition (and hence, nominative is illicit here).

(43) Sto proavlio I neari mathitria krivotan apo kapjous
In-the yard the young student was-hiding from someone-ACC
alla kanis den katalave pjous/*pji.
but noone-NOM NEG realized who-ACC/*who-NOM
'In the yard the young student was hiding from someone, but no one realized who.'

This pattern is expected on the assumption that the fragment undergoes purely processing-based form minimization. But it is problematic for Abels's fit condition. If there is a well-formed underlying structure for the fragment in (43), it must include a PP within which the case on the fragment is licensed and out of which the fragment can move, but such a structure is unavailable in Greek.

While it is clear that the kind of P-drop discussed in this section is specific to anaphora, there is evidence that although Korean case drop is not, it behaves differently in anaphoric contexts than in non-anaphoric ones. Nykiel et al. (2018) demonstrated experimentally that semantic case markers may be dropped from Korean merger-type fragments, as in (44).

(44) A: nwukwu-lopwuthe ton-ul pat-ass-e?
who-SRC money-ACC receive-PST-QUE
'From whom did you receive money?'
B: Mimi-lopwuthe. 'Mimi-from'
B′: Mimi. 'Mimi'

Ratings for caseless fragments trended in the direction of lower acceptability than ratings for case-marked fragments, although this difference failed to reach

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

statistical significance. Nykiel et al. (2018) obtained similar results for Korean structural case markers. The fragment in (45) may be realized as either an accusative-marked NP or a caseless NP, with no reliable acceptability difference between the two options.

(45) A: Mimi-ka mwues(-ul) masy-ess-ni?
Mimi-NOM what(-ACC) drink-PST-QUE
'What did Mimi drink?'
B: Cwusu-lul 'juice-ACC'
B′: Cwusu 'juice'.

There is one configuration in which minimal (caseless) fragments are better than larger ones. In examples like (45A), the fragments' correlate (*mwues-ul* 'what-ACC') can undergo case ellipsis such that it is realized as the caseless NP *mwues*. The most acceptable way to combine the fragment and the correlate is by matching their structural case features by making the fragment caseless as well (Nykiel et al. 2018) (this result resembles the matching effects we discussed in Section 3.2).

The fact that Korean structural case, as opposed to semantic case, tolerates case ellipsis is relevant for our argument. Structural case is limited to accusative, genitive and nominative and is arguably specified at the constructional level in the sense of Goldberg (1995, 2006) rather than at the lexical level (Kim 2016). In other words, an NP that fills a given slot in a syntactic structure receives the case required by that slot. It is therefore possible, given a certain syntactic structure, to predict what structural cases will be assigned to particular slots in that structure. The predictability of structural case is a likely reason why NPs can drop their structural case markers in full clauses (see Lee 2016). However, Korean distinguishes between the structural cases accusative and genitive, on the one hand, and nominative, on the other. While the former may be dropped from full clauses, the latter may not. Regardless of this difference, merger fragments can drop nominative case markers in addition to accusative and genitive, as shown for the fragment in (46B′) (see Morgan 1989 for this observation regarding the nominative case).

(46) A: Nwukwu-ka/*Nukwu ku chaek-ul sa-ass-ni?
who-NOM/*who the book-ACC buy-PST-QUE
'Who bought the book?'
B: Yongsu-ka 'Yongsu-NOM'
B′: Yongsu 'Yongsu'.

This pattern suggests that merger fragments can undergo form reduction that is not always available to NPs in full clauses. This is an important property of fragments because we could assume that the possibility of dropping structural case that exists for fragments simply follows from the fact that argument NPs have the same possibility in full clauses. That is, case drop in fragments could be reduced to the way argument NPs can be realized in full clauses (see Merchant 2004 for this

suggestion). However, (46) demonstrates that there is no such correspondence between fragments and argument NPs, and hence no possibility of deriving caseless fragments from legitimate underlying sources here. On this basis, Morgan (1989) argues for a hybrid analysis of fragments, where case-marked fragments derive via deletion, while caseless fragments are base-generated. However, we do not pursue such an analysis here.

4. HPSG ANALYSIS

We begin by overviewing the theoretical apparatus that we adopt in our analysis and then move on to the constructional constraint that licenses fragments in Section 4.2. The key components of our proposal are based on Ginzburg & Sag (2000), as detailed below.

4.1. *Theoretical apparatus*

We follow Ginzburg & Sag (2000) in assuming that a fragment contains no unpronounced material. The fragment in (47B) therefore is a type of finite clause with the simple structure consisting of the NP *Harvey* shown in (48). This means in turn that fragments can appear in all the contexts that permit finite clauses.[15]

(47) A: Who left the party?
B: Harvey.

(48)

Despite its simple structure, the fragment has propositional semantics supplied by the surrounding context in the form of a Maximal-Question-under-Discussion (see below). The resolution of the fragment is supported by the following additional discourse-based machinery. We adopt the idea common in Construction Grammar frameworks, including in HPSG, that all levels of linguistic description are understood as pairings of form with semantic or discourse functions, and the grammar is a system of constraints that license these pairings (e.g. Goldberg 2006; Sag 2012). Pairings of form and meaning are defined as constructions in some frameworks

[15] Culicover & Jackendoff (2005) treat only sluicing remnants as being of the clausal type, with other fragments treated as nonsentential utterances. See also Kim & Nykiel (2020) and Nykiel & Kim (published online October 6, 2020) for the discussion of HPSG analyses of other elliptical constructions.

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

(Goldberg 2006), and in other frameworks the term constructions is extended to refer to constraints on classes of these pairings (Sag 2012). Regardless of how constructions are understood, it is expected that there may be particular discourse- or processing-based constraints that are limited to only some constructions. The availability of form minimization, as supported by the processing principle of MiF we evoked in Section 3.3, and its limiting case represented by sprouting fragments are examples of just such construction-specific constraints on fragments.

The interpretation of a fragment depends on the notion of 'Question-under-Discussion' (QUD) in the dialogue. Dialogues are described via a Dialogue Game Board (DGB) which provides a record of who said what to whom, and what/who they were referring to (Ginzburg 1996, 2012; Ginzburg & Fernandez 2010). In other words, the DGB monitors which questions are under discussion, constantly updating the value of QUD as the dialogue progresses. This context offers the basis for the interpretation of fragments. The DGB has two attributes, Maximal-Question-under-Discussion (MAX-QUD) and Salient Utterance (SAL-UTT), that are relevant for fragments.

The attribute MAX-QUD, representing the question currently under discussion, takes as its value QUESTIONS defined as propositional abstracts (see Ginzburg & Sag 2000 for a defense of this proposal and a critique of Groenendijk & Stokhof's 1984 early account of the semantics of questions in terms of 'exhaustive answerhood conditions'). We assume, with Ginzburg & Sag (2000), that questions are distinguished from other messages in terms of the feature PARAMS (parameters or variables), whose set value is empty for polar questions and nonempty for wh-questions. Each wh-phrase is associated with one parameter, which means that the PARAMS set of a wh-question will minimally be a singleton set. This can be reinterpreted using the λ-abstraction format, as in Sag (2010). The notations in (49) are sample semantic representations for interrogatives in this format: a unary wh- question with a singleton PARAMS set in (49a) and a polar question with an empty PARAMS set in (49b).

(49) a. Unary *wh* -question: $\lambda\{\pi_x\}$[leave(*x, p*)] (Who left the party?)
b. Polar question: $\lambda\{\}$[leave(*h, p*)] (Did Harvey leave the party?)

As can be seen here, wh-questions are open propositions with a nonempty PARAMS set, while polar questions are closed propositions with an empty PARAMS set (see Sag 2010). A MAX typically arises from an interrogative clause, in which case its content is the content of that interrogative clause, as in (47A), or from a declarative clause hosting an indexed XP, usually an indefinite NP or a quantified NP, as in sluicing (see (38A) and (39A)), so long as the result is a question that remains unresolved.

Although part of MAX-QUD, which is a semantic object, SAL-UTT is also the locus of certain elements of prior syntactic structure. Ginzburg & Sag (2000) define it as an object of type *local*, which encodes both semantic information and morphosyntactic information like syntactic category and case via its CAT ('syntactic category') feature. This information is used by Ginzburg and Sag for licensing the

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

morphosyntax of fragments, as SAL-UTT serves as a fragment's correlate. For instance, uttering the question *Who left the party?* in (47A) will activate the feature structure with the appropriate DGB information in (50).

(50)
$$\begin{bmatrix} \text{FORM} \langle \text{ Who left the party?} \rangle \\ \text{SYN S} \\ \text{SEM } \lambda\{\pi_x\}[\text{leave}(x, p)] \\ \text{DGB} \begin{bmatrix} \text{MAX-QUD } \lambda\{\pi_x\}[\text{leave}(x, p)] \\ \text{SAL-UTT} \left\{ \begin{bmatrix} \text{SYN} [\text{CAT} \quad \text{NP}] \\ \text{SEM } x \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix}$$

The lexical entry for the verb *leave* furthermore includes the ARG-ST list of two ordered NPs, as in (51), in addition to the SPR and COMPS lists. The ARG-ST list is an ordered list of all the arguments of a lexical head, whether realized canonically or non-canonically (e.g. as gaps). The SPR list contains the subject argument, and the COMPS list contains remaining arguments that are realized canonically (the COMPS list thus may be empty if no non-subject arguments are realized canonically). The SPR and COMPS lists are part of the VAL ('valence') feature of a lexical head, a head feature whose values are shared between the mother node and the head daughter node.

(51)
$$\begin{bmatrix} \text{SPR} \quad \langle \boxed{1} \rangle \\ \text{COMPS} \quad \langle \boxed{2} \rangle \\ \text{ARG-ST} \langle \boxed{1}\text{NP}, \boxed{2}\text{NP} \rangle \end{bmatrix}$$

The MAX-QUD in (50) asks for the identity of the person that left the party, and this information functions as the SAL-UTT whose syntactic category is NP (see the SAL-UTT's CAT feature value in (50)) and which is co-indexed with the first NP on the ARG-ST list and the only member of the SPR list in (51). The fragment *Harvey* in (47B) supplies a value for the variable (x) and in this case agrees with the SAL-UTT in syntactic category. What we now need is a better understanding of what can serve as a SAL-UTT, especially given the patterns relating to argument structure that we discussed in Section 3.2.

Note that the value of the SAL-UTT may be determined in a forward- or backward-looking strategy (Ginzburg & Sag 2000; Ginzburg 2012). The forward-looking strategy preselects candidates for the SAL-UTT, while the backward-looking strategy uses a fragment as a guide to what the SAL-UTT is. We will use the backward-looking strategy since it appears to be more consistent with the direct-access mechanism via which fragments are resolved and for reasons addressed in Section 4.2 (but see Ginzburg 2012 for discussion of the viability of both strategies). Settling the value of the SAL-UTT via the backward-looking strategy is also entirely consistent with Culicover & Jackendoff's (2005) indirect licensing mechanism, whereby the fragment is matched to an explicit or implicit phrase in the antecedent clause and

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

inherits the syntactic features assigned to that phrase, without the two being explicitly required to share morphosyntactic features.

It is well known that fragments can pick out overt or implicit adjuncts as their correlates. *Why* remnants are the most frequent among direct sluices (= non-clarificational remnants), and *when* remnants are more frequent than both *who* remnants and *which-NP* remnants, whose correlates are arguments (Ginzburg 2012:244). We must then assume that overt or implicit adjuncts can give rise to MAX-QUDS and are available to serve as the SAL-UTTs within them.[16] The morphosyntactic features of adjunct correlates are not licensed by any lexical heads in the antecedent, and hence it is not necessary to impose any identity constraints, beyond matching index values, on pairs of fragments and their adjunct correlates (for discussion, see Culicover & Jackendoff 2005 and Ginzburg 2012).

However, we need to rule in the set of elements of the ARG-ST list of lexical heads in the antecedent that are permitted to function as the SAL-UTT in case the fragment picks out an argument correlate. We do so in two steps, first by co-indexing the VAL values of the fragment with those of the SAL-UTT, and second by requiring that the SPR and COMPS lists are empty. These steps ensure that the SAL-UTT and the fragment are both maximal categories present in the argument structure of a lexical head, including optional arguments and case and syntactic category alternations, as discussed in Section 3.2. We flesh out this idea in the next section.

4.2. *The head-fragment construction*

Given the current HPSG analyses of sluicing (Ginzburg & Sag 2000; Sag & Nykiel 2012; J.-B. Kim 2015a) and of Korean fragments (J.-B. Kim 2015b), we take the view that there is no syntactic structure at the ellipsis site and fragments are the sole daughters of an S-node. This view has been formalized in the HPSG feature system as follows (see Ginzburg & Sag 2000):

(52) Head-Fragment Construction (to be revised)

$$\left[\begin{array}{ll}\text{SYN} & \text{S} \\ \text{DGB} & \left[\text{SAL-UTT} \left\{\left[\begin{array}{lll}\text{SYN} & \text{[CAT} & \text{[HEAD} \ \boxed{1}\ ] \\ \text{SEM} & \text{[INDEX} & i\end{array}\right]\right\}\right]\end{array}\right] \rightarrow \left[\begin{array}{lll}\text{SYN} & \text{[CAT} & \text{[HEAD} \ \boxed{1}] \\ \text{SEM} & \text{[INDEX} & i\end{array}\right]$$

This constraint allows the mother to be of the same type as finite clauses and the head daughter to serve as a stand-alone XP. The head daughter may be any syntactic category but it must match the syntactic category and case features specified by the SAL-UTT, as indicated by their identical CAT feature values. However, this constraint is too strong, incorrectly ruling out any case or syntactic category alternations, even though they may be licensed by the argument structure of a given lexical head in the antecedent. We propose a way to relax it by means of the constraint given in (53).

[16] It is beyond the scope of this paper to address ways of representing adjuncts in the antecedent structure, but see Bouma, Malouf & Sag (2001) and Ginzburg (2012) for relevant proposals.

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

(53) Head-Fragment Construction

$$\left[\begin{array}{ll}\text{SYN} & \text{S} \\ \text{DGB} & \left[\text{SAL-UTT} \left\{\left[\begin{array}{ll}\text{SYN} \mid \text{VAL} & \boxed{1}\left[\begin{array}{ll}\text{SPR} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle\end{array}\right] \\ \text{SEM} & [\text{INDEX}\ \ i]\end{array}\right]\right\}\right]\end{array}\right] \rightarrow \left[\begin{array}{ll}\text{SYN} \mid \text{VAL} & \boxed{1} \\ \text{SEM} & [\text{INDEX}\ \ i]\end{array}\right]$$

(53) requires that the VAL feature of both the SAL-UTT and the fragment consist of empty SPR and COMPS lists. This ensures that the SAL-UTT and the fragment are saturated phrases, that is, maximal categories present on the ARG-ST lists of lexical heads (usually verbal heads, but see below) in the antecedent. There is no requirement of CAT-feature identity between the fragment and the SAL-UTT, but only of VAL-feature identity, so that we can derive pairs of SAL-UTTs and fragments that may be non-identical in terms of syntactic category or case features but are co-indexed with the same slot in the ARG-ST list of an appropriate lexical head in the antecedent clause. However, the syntactic category and case information that SAL-UTTs and fragments carry as part of their CAT feature provides a means of checking that they have the morphosyntactic features that are appropriate for the slot in the ARG-ST list that they are co-indexed with.

SAL-UTTs and fragments have access to all members of the subcategorization frame of a lexical head, including those realized non-canonically, in a given syntactic arrangement. But they do not have access to alternative syntactic arrangements of the elements on the ARG-ST list that arise from the application of lexical rules and result in different ARG-ST lists. For instance, the alternants that are part of the benefactive alternation (e.g. *Harvey is buying drinks for everyone* vs. *Harvey is buying everyone drinks*) are linked to different ARG-ST lists, as in (54a, b).

(54) a. $\left[\text{ARG-ST}\ \langle\ \text{NP}_i, \text{NP}_j, \text{PP}[\textit{for}]_k\ \rangle\right]$
b. $\left[\text{ARG-ST}\ \langle \text{NP}_i, \text{NP}_k, \text{NP}_j\ \rangle\right]$

This aspect of the theory correctly predicts that fragments attempting to retrieve arguments that are part not of the current ARG-ST list, but of an alternating ARG-ST list (e.g. (54b) if the antecedent features (54a), recall the discussion around Example (24) in Section 3.1) are ill-formed.

By (53), the SAL-UTT may minimally be the content of the PARAMS set of the MAX-QUD. Thus, if the PARAMS set contains an expression co-indexed with an argument derived from the ARG-ST list of a prepositional head in the antecedent, then the SAL-UTT may just correspond to that argument, as in (32B′), repeated here as (55B″). But the SAL-UTT can be a larger maximal category that contains the content of the PARAMS set (e.g. the PP licensed by the verbal head *talk* in (55B′)), and it may include a lexical head and the argument it licenses, as in (55B). All of these possibilities are indicated by the size of the fragments, pointing to the backward-looking strategy, and are available here because all three SAL-UTTs are overt phrases and there is nothing to block the application of MiF to these fragments (they will differ in

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

acceptability and/or usage frequency, though, as a function of the accessibility of the antecedent, see Section 3.3).

(55) A: What the hell are you talking about?
B: Talking about one of your people.
B′: About one of your people.
B″: One of your people.

Furthermore, if there is only one way of realizing a particular argument, then the SAL-UTT will simply carry the morphosyntactic features that are licensed for that argument, with the fragment matching these features. Otherwise, we may see case and/or syntactic category variation between the SAL-UTT and the fragment, as discussed in Section 3.2.

Further reasons to believe that the backward-looking strategy of settling the value of the SAL-UTT is correct come from data like (56)–(59) harvested from COCA.

(56) A: Are you a frustrated performer yourself?
B: Frustrated by what?

(57) A: With trembling fingers the fellow gathered up his scraps from the desk.
B: Trembling with what?

(58) A: But the Fed is an independent agency.
B: Independent of what?

(59) If there is going to be a bridge loan, my question is a bridge to what?

It would be difficult to single out the correct phrases for being the SAL-UTTs using the forward-looking strategy. It seems instead that these sprouting fragments are picking out arguments licensed by nonverbal heads embedded deeper in the structure of the antecedents in a manner consistent with the backward-looking strategy. Note also that these examples speak in favor of the DI approach to fragments in that they cannot be captured by Abels's fit condition (see Section 2.1), which requires that fragments fit into the antecedent structure. These fragments clearly do not do so (e.g. * *Are you a frustrated by something performer?*, * *But the Fed is an independent of something agency*).[17]

Finally, recall the fragments in (41)–(42) repeated here as (60)–(61).

(60) A: What does it look like?
B: Like keys. (Seinfeld)

(61) A: On what does the well-being of the EU depend?
B: A stable currency. (Ginzburg & Sag 2000:301)

[17] However, see Hardt, Anand, & McCloskey (2020) for an attempt to define structural identity between fragments and their antecedents over smaller-sized phrases (e.g. NPs or APs) than an entire antecedent structure, while preserving the mechanism of movement-and-PF-deletion. At first blush, it seems to us to be a plausible non-DI alternative to capturing the data in (56)–(59).

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

In both examples the SAL-UTTS can only be identified on the basis of what the fragments are. In (60A) the SAL-UTT is a PP, and hence a larger phrase than the content of the params set of the MAX-QUD. In (61A) the SAL-UTT is just the value contributed by the wh-phrase despite the fact that the WH value of the wh-phrase is passed on to the fronted PP in cases of pied-piping (see Ginzburg & Sag 2000:195–200) and the entire PP is thus co-indexed with the PARAMS set of the MAX-QUD, being a plausible candidate for the SAL-UTT.

We now turn to ways of licensing sprouting fragments in our account.

4.3. *Licensing sprouting fragments*

We argued in Section 3.3 that fragments that pick out implicit phrases as their correlates must be semantically and syntactically more complex than fragments whose correlates are explicit phrases. This is to ensure that implicit correlates, which are harder to retrieve than explicit phrases due to their low accessibility, are efficiently identified based on the information provided by fragments. Our specific task is to block the generation of fragments smaller than PPs in English and smaller than case-marked NPs in Korean. The PF-deletion accounts designed to specifically deal with sprouting attempt to ensure that all elements that undergo deletion in the IP-clause have counterparts in the antecedent clause. This idea is expressed for sluicing in Chung (2006:11, 2013:30) as stated in (62) and (63), respectively.

(62) *Lexico-syntactic requirement*
Every lexical item in the numeration of the sluice that ends up (only) in the elided IP must be identical to an item in the numeration of the antecedent CP.

(63) *Case condition*
If the interrogative phrase is a DP, it must be case-licensed in the ellipsis site by a head identical to the corresponding head in the antecedent clause.

These conditions guarantee that remnants do not surface as NPs if the corresponding NPs are not part of the overt structure of the antecedent clauses. Hence data like (64)–(65) can be captured correctly if we extend the conditions in (62) or (63) to fragments.

(64) A: Did the letter come?
B: Yes, from Mimi./*Yes, Mimi.

(65) A: Harvey is angry.
B: Yes, with Mimi./*Yes, Mimi.

The DI accounts, on the other hand, cannot establish comparisons between the antecedent structure and the underlying structure of a fragment. A proposal articulated in Kim (2015a: 284) on the basis of Korean sluicing data is that material not present overtly in the antecedent structure has no ability to serve as a correlate for a sluicing remnant. Kim's constraint is given in (66).

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

(66) *Full instantiation constraint*
The syntactic information (e.g. case features or syntactic category) not available at surface, but updated in the DGB needs to be fully specified in the subsequent syntax.

We abandon (66) in favor of NFM, which we motivated in Section 3.3 and which differentiates between fragments with adjunct correlates and fragments with argument correlates. It appears that adjunct sprouting, unlike argument sprouting, permits form minimization, as in (67) from the Santa Cruz Ellipsis Project. The fragment is the minimal NP *which stop* rather than the larger PP *to which stop* (the intended interpretation is provided in square brackets).

(67) A: But I tested at 85 percent, so we're on the right track.
B: Right track, but which stop [are we on the right track to]? (Anand 2019)

To make NFM work correctly we must first ask whether implicit arguments should be represented in the syntax of antecedents. There are two options here. First, implicit arguments could appear on the ARG-ST lists of lexical heads. This idea can be implemented following the proposal of Ruppenhofer & Michaelis (2014) that lexical heads can subcategorize for non-instantiated indefinite null (INI) arguments (see also Kim 2015a). For instance, the lexical entry for the adjective *angry* in (68) would include an implicit PP argument annotated as INI on the ARG-ST list which can optionally be realized as a complement and can be activated in context whether or not realized overtly.[18] This PP is thus defined as a type of non-canonical object on the ARG-ST list.[19]

(68) Lexical entry for *angry*:

$$\begin{bmatrix} \text{FORM} & \langle \text{angry} \rangle \\ \text{ARG-ST} & \left\langle \text{NP}_i, \text{PP}[ini]_j \right\rangle \\ \text{SEM} & angry.with(i, j) \end{bmatrix}$$

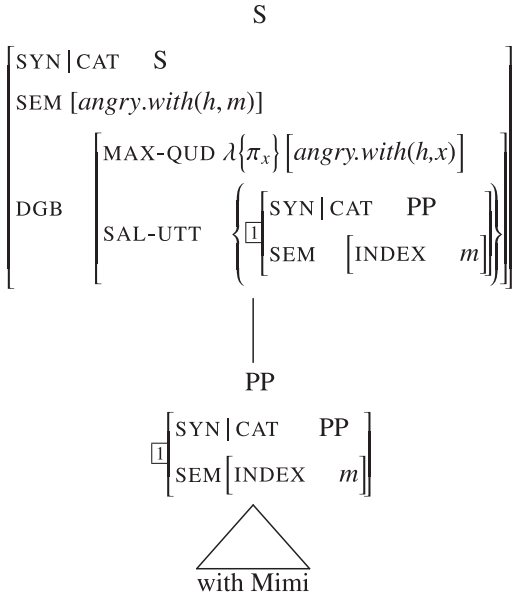
[18] Such examples differ from those like *We arrived [e] at 8*, where the unexpressed goal argument is known to the interlocutors in the given context and the omission of this argument is thus an instance of definite null instantiation (DNI). Ruppenhofer & Michaelis (2014) argue that INI and DNI arguments can be distinguished by the way we reconstruct them, that is, either by an indefinite expression like *something, someone* or a definite expression like *it* or *him.*

[19] Defining implicit phrases as non-canonical objects could help us account for the behavior of P-drop under sprouting without appeal to processing principles. The idea (which we owe to Anne Abeillé) is that an INI argument, unlike a canonical argument, has a specified syntactic category but no internal structure (i.e. no smaller maximal projections within it), which effectively blocks the generation of fragments corresponding to any smaller phrases within that argument (see Nykiel & Kim 2020, for more detail). We leave it open whether this avenue is preferable over our proposal. A potential problem with it is that it runs the risk of being a theory-internal explanation rather than one motivated by independent processing principles that do a good job of explaining how merger and sprouting fragments behave with respect to P-drop, as we argued in Section 3.3.

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

Thus in an exchange like (65), uttering the statement that Harvey is angry can activate the implicit PP argument and introduce a MAX-QUD questioning the value of the variable introduced by the PP argument ($\lambda_x[angry.with(h,x)]$). The fragment *with Mimi* offers a value for this variable. The following structure is thus projected for it in (69):

(69)

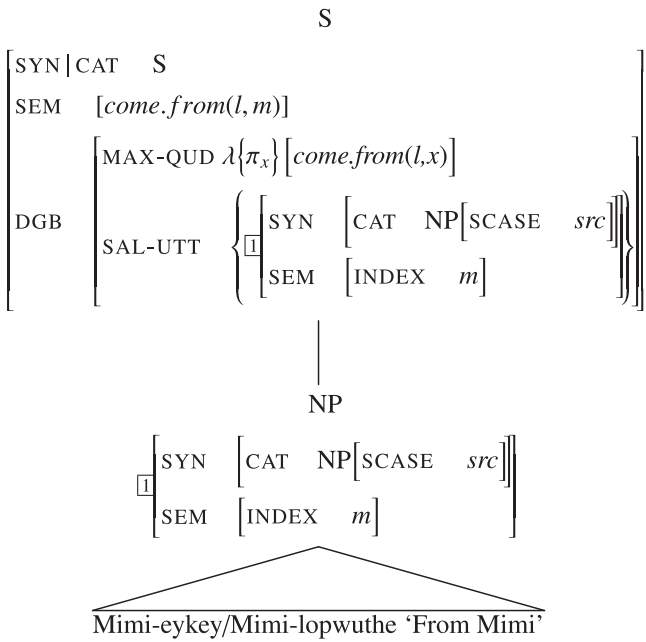


Although featural match is not required between the fragment and the SAL-UTT, they both are PPs since they are co-indexed with the same slot in the ARG-ST list of *angry* (but if this slot in the ARG-ST list of *angry* permitted another syntactic category alternating with a PP, the fragment could optionally be realized as that category, recall Section 3.1).

Korean sprouting fragments are accounted for in a similar manner. We could assume that Korean, like English, allows INI arguments (see Ahn & Cho 2012 for detailed discussion) and that such arguments differ from DNI arguments. The Korean fragment in (22) can function as a response to the statement *A letter came*, licensing a structure like the following:

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

(70)



Uttering the statement that a letter came evokes the MAX-QUD asking who (x) the letter came from. This variable is activated from the ARG-ST of the verb *o-ss-e* 'come-PST-DECL' that licenses an indefinite NP argument bearing the source semantic case (SCASE) and being annotated as INI in the syntactic representation of the statement *A letter came.*[20] The fragment in (70) supplies a value for this variable, which serves as the SAL-UTT. Since two case markers -eykey and *-lopwuthe* can indicate source, either can be selected for the fragment (J.-B. Kim 2015a). However, the INI annotation on the SAL-UTT provided by the antecedent triggers the application of NFM, blocking the fragment from being caseless.

The second option is to allow implicit arguments to be absent from the syntax of antecedents, while recognizing that they are activated in online processing. One implementation of this idea is found in Culicover & Jackendoff's (2005) proposal that fragments may be matched to implicit phrases that are licensed as part of subcategorization frames of lexical heads and activated together with those heads in online processing. The processing of a lexical head will activate its entire set of arguments, including optional ones, which will then be available as correlates for fragments. This means that a fragment's semantic index is not linked to any arguments encoded in its antecedent's syntax but instead points to a grammatical function that is fulfilled by a certain implicit phrase with appropriate morphosyntactic properties. An anonymous referee points out that for a MAX-QUD to arise from

[20] See J.-B. Kim (2016) where case markers in Korean are classified into structural cases and semantic cases (SCASE), and further values are defined to include *nominal* and *verbal.*

https://doi.org/10.1017/S0022226720000420 Published online by Cambridge University Press

such an implicit correlate might require a certain degree of accommodation, which is costly. To see this, consider (65) again. The salient MAX-QUD is the polar question of whether Harvey is angry, but the fragment *with Mimi* requires that the DGB be updated with the new MAX-QUD *Who is Harvey angry with?* This additional step arguably introduces processing difficulty that is absent from examples where no update to the DGB is needed. However, it is equally plausible that processing difficulty arises here from having to retrieve a low-accessibility correlate (one with no linguistic content), as predicted by MiF, and also from retrieving the new MAX-QUD that this correlate gives rise to. All fragments used in this context will be required to be realized as phrases exactly matching their implicit correlates, with no form minimization permitted, as per NFM.

Whichever option we choose, we are dealing with phrases with no linguistic content that make less accessible correlates, and hence phrases that trigger the application of NFM. An advantage of the second option, as noted by an anonymous referee, is that it is more consistent with the backward-looking strategy than the forward-looking one, given that implicit arguments are not represented in the antecedent syntax.

5. Conclusion

We have proposed a structural constraint on fragments that grants them access, where appropriate, to argument structure information derived from the antecedent clauses, including syntactic category and case alternations. Our proposal is couched in terms of the DI approach to clausal ellipsis, based primarily on Ginzburg & Sag (2000) and Culicover & Jackendoff (2005), and incorporates processing-based preferences as a means to motivate the contrast between merger and sprouting fragments. Specifically, we have proposed that phrases which are available to serve as correlates for fragments (i.e. SAL-UTTS) are maximal categories present in the argument structure of lexical heads in the antecedent clauses and that their size is determined by processing preferences, as captured by the principle of Minimize Forms. This proposal successfully predicts form-matching effects that surface under clausal ellipsis and, unlike the existing movement-and-PF-deletion accounts, does not face the challenge of having to derive all possible fragments from legitimate sentential sources.

REFERENCES

Abeillé, Anne & Shrita Hassamal. 2019. Sluicing in Mauritian: A fragment-based approach. *Empirical Issues in Syntax and Semantics* 12.

Abels, Klaus. 2017. On the interaction of P-stranding and sluicing in Bulgarian. In Olav Mueller-Reichau & Marcel Guhl (eds.), *Aspects of Slavic linguistics: Formal grammar, lexicon and communication* (Language, Context and Cognition 16), 1–28. Berlin: De Gruyter Mouton.

Abels, Klaus. 2019. On 'sluicing' with apparent massive pied-piping. *Natural Language and Linguistic Theory* 37, 1205–1271.

Ahn, Hee-Don & Sungeun Cho. 2012. Fragments vs. null arguments in Korean. In Stefan Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar* (HPSG), 369–387. Stanford, CA: CSLI Publications.

Anand, Pranav. 2019. Identity and mismatch: The view from sluicing. *Presented at Experimental and Corpus-based Approaches to Ellipsis 2*. Davis, CA.

Barros, Matthew. 2014. *Sluicing and identity in ellipsis*. Doctoral dissertation, Rutgers University.

Barton, Ellen. 1991. Nonsentential constituents and theories of phrase structure. In Katherine Leffel & Denis Bouchard (eds.), *Views on phrase structure*, 193–214. Dordrecht: Kluwer.

Barton, Ellen. 2006. Toward a nonsentential analysis in generative grammar. In Ljiljana Progovac, Kate Paesani, Eugenia Casielles & Ellen Barton (eds.), *The syntax of nonsententials: Multidisciplinary perspectives*, 11-31. Amsterdam: John Benjamins.

Bechhofer, Robin. 1976. Reduced wh-questions. In Jorge Hankamer & Judith Aissen (eds.), *Harvard studies in syntax and semantics*, vol. 2, 31–67. Cambridge, MA: Harvard University.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19, 1–65.

Caha, Pavel. 2011. Case in adpositional phrases. Ms., lingbuzz/001325, CASTL, Tromsø.

Callahan, Sarah M., Lewis P. Shapiro & Tracy Love. 2010. Parallelism effects and verb activation: The sustained reactivation hypothesis. *Journal of Psycholinguistic Research* 39, 101–118.

Caplan, David & Gloria Waters. 2013. Memory mechanisms supporting syntactic comprehension. *Psychonomic Bulletin & Review* 20, 243–268.

Carlson, Katy. 2002. *Parallelism and prosody in the processing of ellipsis sentences*. New York: Routledge.

Chung, Sandra. 2006. Sluicing and the lexicon: The point of no return. In Rebecca T. Cover & Yuni Kim (eds.), *Proceedings of the annual meeting of the Berkeley Linguistics Society*, vol. 31, 73–91. Berkeley, CA: Berkeley Linguistics Society.

Chung, Sandra. 2013. Syntactic identity in sluicing: How much and why. *Linguistic Inquiry* 44, 1–44.

Chung, Sandra, William A. Ladusaw & James McCloskey. 1995. Sluicing and logical form. *Natural Language and Linguistic Theory* 3, 239–282.

Chung, Sandra, William A. Ladusaw & James McCloskey. 2011. Sluicing(:) between structure and inference. In Rodrigo Gutiérrez-Bravo, Line Mikkelsen & Eric Potsdam (eds.), *Representing language: Essays in honor of Judith Aissen Representing language: Essays in honor of Judith Aissen*, 31–50. Santa Cruz: Department of Linguistics, University of California.

Craik, Fergus I. M. & Endel Tulving. 1975. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 104, 268–294

Culicover, Peter & Ray Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press.

Culicover, Peter & Ray Jackendoff. 2012. Same-except: A domain-general cognitive relation and how language expresses it. *Language* 88, 305–340.

Dickey, Michael Walsh & Ann C. Bunger. 2011. Comprehension of elided structure: Evidence from sluicing. *Language and Cognitive Processes* 26.1, 63–78.

Dubey, Amit, Patrick Sturt & Frank Keller. 2005. Parallelism in coordination as an instance of syntactic priming: Evidence from corpus-based modeling. *Proceedings of Human Language Technology Conference and Empirical Methods in Natural Language Processing*, 827–834.

Ellory, Roger J. 2010. City of lies. Orion.

Fisher, Ronald P. & Fergus I. M. Craik. 1980. The effects of elaboration on recognition memory. *Memory & Cognition* 8, 400–404.

Frazier, Lyn & Charles Clifton Jr. 1998. Comprehension of sluiced sentences. *Language and Cognitive Processes* 13.4, 499–520.

Frazier, Lyn, Lori Taft, Tom Roeper, Charles Clifton Jr. & Kate Ehrlich. 1984. Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition* 12, 421–430.

Gallo, David A., Nathaniel G. Meadow, Elizabeth L. Johnson & Katherine T. Foster. 2008. Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. *Journal of Memory and Language* 58, 1095–1111.

Ginzburg, Jonathan. 1996. Interrogatives: Questions, facts, and dialogue. In Shalom Lappin (ed.), *Handbook of Contemporary Semantic Theory*, 385–422. Oxford: Blackwell.

Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation*. Oxford: Oxford University Press.

Ginzburg, Jonathan & Raquel Fernandez. 2010. Computational models of dialogue. In Alexander Clark, Chris Fox & Shalom Lappin (eds.), *Handbook of computational linguistics and natural language processing*, 429–481. Chichester: Wiley Blackwell.

Ginzburg, Jonathan & Philip Miller. 2018. Ellipsis in head-driven phrase structure grammar. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The Oxford Handbook of Ellipsis*, 75–121. Oxford: Oxford University Press.
Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations*. Stanford, CA: CSLI Publications.
Goldberg, Adele. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago, IL: University of Chicago Press.
Goldberg, Adele. 2006. *Constructions at work: The nature of generalization in language*. New York: Oxford University Press.
Gonzalez-Vilbazo, Kay E. & Sergio E. Ramos. 2014. *A morphosyntactic condition on sluicing: Evidence from Spanish/German code-switching. Ms., lingbuzz/003411*, University of Illinois at Chicago.
Griffiths, James & Anikó Lipták. 2014. Contrast and island sensitivity in clausal ellipsis. *Syntax* 17, 189–234.
Groenendijk, Jeroen & Martin Stokhof. 1984. *Studies on the semantics of questions and the pragmatics of answers*. Doctoral dissertation, University of Amsterdam.
Hardt, Daniel, Pranav Anand & James McCloskey. 2020. Sluicing: Antecedence and the landscape of mismatch. Paper presented at the Experimental and Corpus-based Approaches to Ellipsis 3.
Harris, Jesse A. 2015. Structure modulates similarity-based interference: An eye tracking study. *Frontiers in Psychology* 6, art. 1839.
Hawkins, John. A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
Hawkins, John. A. 2014. *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press.
Hofmeister, Philip. 2011. Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes* 26, 376–405.
Hofmeister, Philip, T. Florian Jaeger, Ivan A. Sag, Inbal Arnon & Neal Snider. 2007. Locality and accessibility in *wh*-questions. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: Linguistics in Search of its Evidential Base*, 185–206. Berlin: Mouton de Gruyter.
Hofmeister, Philip, T. Florian Jaeger, Inbal Arnon, Ivan A. Sag & Neal Snider. 2013. The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes* 28, 48–87.
Jacobson, Pauline. 2016. The short answer: Implications for direct compositionality (and vice versa). *Language* 92, 331–375.
Karimi, Hossein, Kumiko Fukumura, Fernanda Ferreira & Martin J. Pickering. 2014. The effect of noun phrase length on the form of referring expressions. *Memory & Cognition* 42, 993–1009
Kim, Jong-Bok. 2015a. Syntactic and semantic identity in Korean sluicing: A direct interpretation approach. *Lingua* 166, 260–293.
Kim, Jong-Bok. 2015b. Fragments in Korean: A direct interpretation approach. *Studies in Generative Grammar* 25, 703-733.
Kim, Jong-Bok. 2016. *The syntactic structures of Korean: A construction-based perspective*. Cambridge: Cambridge University Press.
Kim, Jong-Bok & Anne Abeillé. 2019. Why-stripping in English. *Linguistic Research* 36, 365–387.
Kim, Jong-Bok & Joanna Nykiel. 2020. The syntax and semantics of elliptical constructions: A direct interpretation perspective. *Linguistic Research* 37.2, 327–358.
Kim, Jungsoo. 2017. *Sluicing and stripping in Korean: A non-ellipsis, anaphoric analysis*. *Doctoral dissertation*, University of Texas at Austin.
Kobele, Gregory M. 2015. LF-copying without LF. *Lingua* 166, 236–259.
Lahiri, Utpal. 2002. *Questions and answers in embedded contexts*. Oxford: Oxford University Press.
Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.
Lee, Hanjung. 2016. Usage probability and subject–object asymmetries in Korean case ellipsis: Experiments with subject case ellipsis. *Journal of Linguistics* 52, 70-110.
Lemke, Robin. 2016. Sentential or not? An experimental investigation on the syntax of fragments. Saarland University. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/77657.
Leung, Tommy. 2014. The preposition stranding generalization and conditions on sluicing: Evidence from Emirati Arabic. *Linguistic Inquiry* 45, 332–340.
Levelt, Willem & Stephanie Kelter. 1982. Surface form and memory in question answering. *Cognitive Psychology* 14, 78–106.
Levin, Lori. 1982. Sluicing: A lexical interpretation procedure. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 590–654. Cambridge, MA: MIT Press.
Lewis, Richard L. & Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29, 375–419.

Lewis, Richard L., Shravan Vasishth & Julie A. Van Dyke. 2006. Computational principles of working memory in sentence comprehension. *Trends in Cognitive Science* 10, 447–454.
Marks, W. 1987. Retrieval constraints on associative elaborations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13.2, 301–309.
Martin, Andrea E. & Brian McElree. 2011. Direct-access retrieval during sentence comprehension: Evidence from sluicing. *Journal of Memory and Language* 64, 327–343.
McElree, Brian. 2000. Sentence comprehension is mediated by content-addressable memory. *Journal of Psycholinguistic Research* 29, 111–123.
McElree, Brian, Stephani Foraker & Lisbeth Dyer. 2003. Memory structures that subserve sentence comprehension. *Journal of Memory and Language* 48, 67–91.
Merchant, Jason. 2001. *The syntax of silence*. Oxford: Oxford University Press.
Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27, 661–738.
Merchant, Jason. 2005. *Revisiting syntactic identity conditions*. Workshop on ellipsis, University of California, Berkeley.
Merchant, Jason. 2013. Voice and ellipsis. *Linguistic Inquiry* 44, 77–108.
Merchant, Jason, Lynn Frazier, Charles Clifton Jr. & Thomas Weskott. 2013. Fragment answers to questions: A case of inaudible syntax. In Laurence Goldstein (ed.), *Brevity*, 21–35. Oxford: Oxford University Press.
Molimpakis, Emilia. 2018. *P-stranding under sluicing: Acceptability and processing. Workshop Relating Elliptical Utterances to Information in Context*. Stuttgart, Germany.
Morgan, Jerry. 1989. Sentence fragments revisited. In Bradley Music, Randolph Graczyk & Caroline Wiltshire (eds.), *Papers from the 25th annual regional meeting of the Chicago Linguistic Society* (CLS25), 228–241. Chicago: Chicago Linguistic Society.
Nykiel, Joanna. 2013. Clefts and preposition omission in sluicing. *Lingua* 123, 74–117.
Nykiel, Joanna. 2015. Constraints on ellipsis alternation: A view from the history of English. *Language Variation and Change* 27, 227–254.
Nykiel, Joanna. 2017. Preposition stranding and ellipsis alternation. *English Language & Linguistics* 21, 27–45.
Nykiel, Joanna & John A. Hawkins. 2020. English fragments, minimize domains, and minimize forms. *Language and Cognition* 12.3, 411–443.
Nykiel, Joanna & Jong-Bok Kim. Ellipsis. In Stefan Müller, Anne Abeillé, Robert Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*. Pre-published online October 6, 2020, by Language Sciences Press. https://hpsg.hu-berlin.de/Projects/HPSG-handbook.
Nykiel, Joanna, Jong-Bok Kim, Rok Sim & Okgi Kim. 2018. Morphosyntactic form of fragments is relevant to their resolution. In P. Farrell (ed.), *Proceedings of the Linguistic Society of America*, 27, 1–14.
Parker, Dan. 2018. A memory-based explanation of antecedent-ellipsis mismatches: New insights from computational modeling. *Glossa* 3.1, art. 129.
Philippova, Tatiana. 2014. P-omission under sluicing, [P clitic] and the nature of P-stranding. In Martin Kohlberger, Kate Bellamy & Eleanor Dutton (eds.), *Proceedings of the 22nd Conference of the Student Organization of Linguistics in Europe* (ConSOLE XXII), 133–155. Leiden: Centre for Linguistics, Leiden University.
Rodrigues, Cillene, Andrew Nevins & Luis Vicente. 2009. Cleaving the interactions between sluicing and P-stranding. In Danièle Torck & Leo Wetzels (eds.), *Romance Languages and Linguistic Theory 2006*, 175–198. Amsterdam: John Benjamins.
Ross, John R. 1969. Guess who? In Robert Binnick, Alice Davison, Georgia Green & Jerry Morgan (eds.), *Papers from the 5th regional meeting of the Chicago Linguistic Society*, 252–286. Chicago: Chicago Linguistic Society.
Ruppenhofer, Josef & Laura A. Michaelis. 2014. Frames and the interpretation of omitted arguments in English. In Stacey Katz Bourns & Lindsy L. Myers (eds.), *Perspectives on linguistic structure and context: Studies in honor of Knud Lambrecht*. 57–86. Amsterdam: John Benjamins.
Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86.3, 486–545.
Sag, Ivan A. 2012. Sign-based construction grammar: An informal synopsis. In Hans Boas & Ivan A. Sag (eds.), *Sign-based construction grammar*, 69–202. Stanford, CA: CSLI Publications.
Sag, Ivan A. & Joanna Nykiel. 2011. Remarks on sluicing. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-driven Phrase Structure Grammar*, 188–208. Stanford, CA: CSLI Publications.
Stjepanović, Sandra. 2008. P-stranding under sluicing in a non-P-stranding language? *Linguistic Inquiry* 39, 179–190.

Stjepanović, Sandra. 2012. Two cases of violation repair under sluicing. In Jason Merchant & Andrew Simpson (eds.), *Sluicing: Cross-linguistic perspectives*, 68–82. Oxford: Oxford University Press.
Szczegielniak, Adam. 2008. Islands in sluicing in Polish. In Natasha Abner & Jason Bishop (eds.), *Proceedings of the 27th West Coast Conference on Formal Linguistics*, 404–412. Somerville, MA: Cascadilla Press.
Theiler, Nadine, Floris Roelofsen & Maria Aloni. 2019. Picky predicates: Why believe doesn't like interrogative complements, and other puzzles. *Natural Language Semantics* 27, 95–134.
Thoms, Gary. 2015. Syntactic identity, parallelism and accommodated antecedents. *Lingua* 166, 172–198.
Turnbull-Sailor, Craig. 2007. *Syntactic patterns of embedded wh-clauses*. Master's dissertation, University of Kansas.
Uegaki, Wataru. 2015. Content nouns and the semantics of question-embedding. *Journal of Semantics* 33.4, 623–660.
Van Craenenbroeck, Jeroen. 2010. Invisible last resort: A note on clefts as the underlying source for sluicing. *Lingua* 120, 1714–1726.
Van Dyke, Julie A. & Brain McElree. 2011. Cue-dependent interference in comprehension. *Journal of Memory and Language* 65, 247–263.
Van Dyke, Julie A. & Clinton L. Johns. 2012. Memory interference as a determinant of language comprehension. *Language and Linguistics Compass* 6, 193–211.
Weir, Andrew. 2014. *Fragments and clausal ellipsis*. Doctoral dissertation, University of Massachusetts Amherst.

Authors' addresses: *(Nykiel)*
Dept. of English Linguistics and Literature, Kyung Hee University,
26, Kyungheedae-ro, Dongdaemun-gu, Seoul, 02447, Korea
joanna.nykiel@khu.ac.kr

(Kim)
Dept. of English Linguistics and Literature, Kyung Hee University,
26, Kyungheedae-ro, Dongdaemun-gu, Seoul, 02447, Korea
jongbok@khu.ac.kr *(corresponding author)*

(Received 23 October 2020)